Series on Concrete and Applicable Mathematics – Vol. 12

Jan A van Casteren

MARKOV PROCESSES, FELLER SEMIGROUPS AND EVOLUTION EQUATIONS

$$\lim_{h\downarrow 0}\int_E \frac{e^{hL}-e^{hL_A}}{h}\int_0^\infty e^{sL_A} f\,ds\,d\pi = \int_{E\setminus A^r} f\,d\pi$$

This page is intentionally left blank

MARKOV PROCESSES, FELLER SEMIGROUPS AND EVOLUTION EQUATIONS

Series on Concrete and Applicable Mathematics – Vol. 12

MARKOV PROCESSES, FELLER SEMIGROUPS AND EVOLUTION EQUATIONS

Jan A van Casteren

University of Antwerp, Belgium

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

MARKOV PROCESSES, FELLER SEMIGROUPS AND EVOLUTION EQUATIONS
Series on Concrete and Applicable Mathematics — Vol. 12

ISBN-13 978-981-4322-18-8
ISBN-10 981-4322-18-0

Printed in Singapore.

SERIES ON CONCRETE AND APPLICABLE MATHEMATICS

ISSN: 1793-1142

Series Editor: Professor George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis, TN 38152, USA

Published

Vol. 3 Introduction to Matrix Theory: With Applications to Business and Economics
by F. Szidarovszky & S. Molnár

Vol. 4 Stochastic Models with Applications to Genetics, Cancers, Aids and Other Biomedical Systems
by Tan Wai-Yuan

Vol. 5 Defects of Properties in Mathematics: Quantitative Characterizations
by Adrian I. Ban & Sorin G. Gal

Vol. 6 Topics on Stability and Periodicity in Abstract Differential Equations
by James H. Liu, Gaston M. N'Guérékata & Nguyen Van Minh

Vol. 7 Probabilistic Inequalities
by George A. Anastassiou

Vol. 8 Approximation by Complex Bernstein and Convolution Type Operators
by Sorin G. Gal

Vol. 9 Distribution Theory and Applications
by Abdellah El Kinani & Mohamed Oudadess

Vol. 10 Theory and Examples of Ordinary Differential Equations
by Chin-Yuan Lin

Vol. 11 Advanced Inequalities
by George A Anastassiou

Vol. 12 Markov Processes, Feller Semigroups and Evolution Equations
by Jan A van Casteren

Dedicated posthumously to my mathematics teacher
Rudi Hirschfeld

Preface

Writing the present book has been a long time project which emerged more than six years ago. One of the main sources of inspiration was a mini-course which the author taught at Monopoli (University of Bari, Italy). This course was based on the text in [Van Casteren (2002)]. The main theorems of the present book (Theorems 2.9 through 2.13), but phrased in the locally compact setting, were a substantial part of that course. The title of the conference was International Summer School on Operator Methods for Evolution Equations and Approximation Problems, Monopoli (Bari), September 15–22, 2002. The mini-course was entitled "Markov processes and Feller semigroups". Other papers which can be considered as predecessors of the present book are [Van Casteren (2000a, 2001, 2008, 2009)]. In this book a Polish state space replaces the locally compact state space in the more classical literature on the subject. A Polish space is separable and complete metrizable. Important examples of such spaces are separable Banach and Frechet spaces. The generators of the Markov processes or diffusions which play a central role in the present book could be associated with stochastic differential equations in a Banach space. In the formulation of our results we avoid the use of the metric which turns the state space into a complete metrizable space; see e.g. the Propositions 4.6 and 9.2. As a rule of thumb we phrase results in terms of (open) subsets rather than using a metric.

For locally compact spaces there is a one-to-one correspondence between Feller-Dynkin semigroups (those are semigroups which send continuous functions, which vanish at infinity, to continuous ones which also are zero at infinity) and certain (strong) Markov processes, which are Hunt processes, and which have the Feller-Dynkin property. This leads to an interaction between stochastic analysis and classical semigroup theory. However, many interesting topological spaces are not locally compact, and in fact are topologically speaking much larger. Nevertheless from the point of view of (stochastic) analysis and possible applications these more general topological spaces are also important. Examples of such spaces are Wiener space, Loop space, Fock space. These spaces are Polish spaces or more general Lusin spaces, which are images of Polish spaces under injective continuous mappings. The present book endeavors to develop an analysis which encompasses Polish spaces. Since, as a rule stochastic differential equations are time-dependent we will consider not only Feller-type semigroups, but also Feller evolutions, or, what is the same, Feller propagators. Our theory works for Feller evolutions acting on the space of bounded continuous functions defined on a Polish space E. The topology of uniform convergence which performs nicely and effectively on locally compact state spaces, is not so appropriate here. One of the main reasons being the fact that the topological dual space of $(C_b(E), \|\cdot\|_\infty)$ consists of bounded Radon measures on the Stone-Čech compactification βE of E, which need not be concentrated on the space E. They may have mass on the "collar" $\beta E \backslash E$. In order to be sure that we are in a setting where the dual space consists of genuine measures on E we replace the unform topology by the strict topology. In the commutative setting this leads to a precise formulation of the relationships which exist between Feller evolution as exhibited in Theorems 2.9 and 2.10. We also bring in the martingale problem, and its relation with Feller processes. The precise results are to be found in Theorems 2.11 and 2.12. In Theorem 2.13 we discuss the problem of operators L which possess a linear extension L_0 which generate a unique Markov process (which in fact is a time-dependent, or non-time-homogeneous, Hunt process).

Included are two chapters on backward stochastic differential equations (BSDE's for short) as well as a chapter on a version of the Hamilton-Jacobi-Bellman equation. Chapter 5 deals with existence and uniqueness of solutions to BSDE's. Conditions on the generator $f(s, x, y, z)$ of the BSDE are phrased in terms of a one-sided Lipschitz condition in the variable y, and a Lipschitz type condition in z. In this condition the squared gradient operator Γ_1, or "opérateur carré du champ" in French, plays a central role. Chapter 6 establishes a relationship between BSDE's and viscosity solutions to more semi-linear classical partial differential equations. It is concluded with a short section on applications to financing (contingent claims and self-financing portfolios). These topics (and presentations) are taken from [El Karoui *et al.* (1997)] and [El Karoui and Quenez (1997)]. In Part 4 we exhibit a number of results pertaining to the long time behavior of recurrent

Markov processes. We present the existence and uniqueness results for stationary (or invariant) measures, also called steady state in case we deal with positive recurrent Markov chains. Chapter 9 also includes a discussion on inequalities of Poincaré and Sobolev type.

Some details

Next we give some more details on the contents of the book. In Chapter 1 we discuss topics related to stochastic differential equations. Results are presented in the finite-dimensional and the infinite-dimensional context. It also contains some standard and not so standard results on martingales and stopping times. This chapter serves as a motivation for the main parts of the book: strong Markov processes, backward stochastic differential equations, long time behavior of solutions. As one of the highlights of the book we mention Theorems 2.9 through 2.13 and everything surrounding it. These theorems give an important relationship between the following concepts: probability transition functions with the (strong) Feller property, strong Markov processes, martingale problems, generators of Markov processes, and uniqueness of Markov extensions. In this approach the classical uniform topology is replaced by the so-called strict topology. A sequence of bounded continuous functions converges for the strict topology if it is uniformly bounded, and if it converges uniformly on compact subsets. It can be described by means of a certain family of semi-norms which turns the space of bounded continuous functions into a sequentially complete locally convex separable vector space. Its topological dual consists of genuine complex measures on the state space. This is the main reason that the whole machinery works. The third chapter contains the proofs of the main theorems. The original proof for the locally compact case, as exhibited in e.g. [Blumenthal and Getoor (1968)], cannot just be copied. Since we deal with a relatively large state space every single step has to be reproved. Many results are based on Proposition 3.1 which ensures that the orbits of our process have the right compactness properties. If we talk about equi-continuity, then we mean equi-continuity relative to the strict topology: see e.g. Theorem 2.2, Definition 2.2, Theorem 2.7, Corollary 2.3, Proposition 3.3, Corollary 3.3, Corollary 3.2, equation (4.114). In §4.4 a general criterion is given in order that the sample paths of the Markov process are almost-surely continuous. In addition this section contains a number of results pertaining to dissipativity properties of its generator: see e.g. Proposition 4.3. A discussion of the maximum principle is found here:

see e.g. Lemma 4.2 and Proposition 4.6. In Section 4.3 we discuss Korovkin properties of generators. This notion is closely related to the range property of a generator. In Section 4.5 we discuss (measurability) properties of hitting times. In Chapters 5 and 6 we discuss backward stochastic differential equations for diffusion processes. A highlight in Chapter 5 is a new way to prove the existence of solutions. It is based on a homotopy argument as explained in Theorem 1 (page 87) in [Crouzeix *et al.* (1983)]: see Proposition 5.7, Corollary 5.3 and Remark 5.19. The connection with the Browder-Minty theorem is mentioned as well: see Theorem 5.10. A martingale which plays an important role in Chapter 6 is depicted in formula (6.3). Basic results are Theorems 6.1 and 6.2. These theorems compare solutions to BSDE's for different generating functions $f(s, x, y, z)$. An interesting consequence of these stopping time and martingale techniques is the fact that the solution (candidate) to the corresponding classical semi-linear partial differential equation of parabolic type is a viscosity solution; for details see Theorem 6.3. In Chapter 7 we discuss for a time-homogeneous process a version of the Hamilton-Jacobi-Bellmann equation. Interesting theorems are the Noether theorems 7.5 and 7.6. In Chapters 8, 9, and 10 the long time behavior of a recurrent time-homogeneous Markov process is investigated. Chapter 8 is analytic in nature; it is inspired by the Ph.-D. thesis of Katilova [Katilova (2004)]. Chapter 9 describes a coupling technique from Chen and Wang [Chen and Wang (2003)]: see Theorem 9.1 and Corollary 9.1. The problem raised by Chen and Wang (see §9.5) about the boundedness of the diffusion matrix can be partially solved by using a Γ_2-condition instead of condition (9.5) in Theorem 9.1 without violating the conclusion in (9.6): see Theorem 9.18 and Example 9.1, Proposition 9.18 and the formulas (9.269) and (9.270). For more details see Remark 9.9 and inequality (9.171) in Remark 9.13. Furthermore Chapter 9 contains a number of results related to the existence of an invariant σ-additive measure for our recurrent Markov process. For example in Theorem 9.2 conditions are given in order that there exist compact recurrent subsets. This property has far-reaching consequences: see e.g. Proposition 9.4, Theorem 9.4, and Proposition 9.6. Results about uniqueness of invariant measures are obtained: see Corollary 9.3. The results about recurrent subsets and invariant measures are due to Seidler [Seidler (1997)]. Poincaré type inequalities are proved: see the propositions 9.10 and 9.16, and Theorem 9.4. The results on the Γ_2-condition are taken from Bakry [Bakry (1994, 2006)], and Ledoux [Ledoux (2000)]. For recent applications of the Γ_2-condition to problems related to the theory of transportation costs see e.g. [Gozlan (2008)]. In Chapter 10 we prove the existence and uniqueness of a σ-finite invariant measure for an irreducible time-homogeneous Markov process: see Theorem 10.5 and the results in §10.1. In Theorem 10.7 we follow Kaspi and Mandelbaum [Kaspi and Mandelbaum (1994)] to give a precise relationship between Harris recurrence and recurrence phrased in terms of hitting times. Theorem 10.12 is the most important one for readers interested in an existence proof of a σ-additive invariant measure which is unique up to a multiplicative constant. Assertion (e) of Proposition 10.8 together with Orey's theorem for Markov chains (see Theorem 10.2) yields the interesting consequence that, up to multiplicative constants, σ-finite invariant measures are unique. In §10.3 Orey's theorem is proved for recurrent Markov chains. In the proof we use a version of the bivariate linked forward recurrence time chain as explained in Lemma 10.14. We also use Nummelin's splitting technique: see Meyn and Tweedie [Meyn and Tweedie (1993b)], §5.1 (and §17.3.1). The proof of Orey's theorem is based on Theorems 10.14 and 10.17. Results in Chapter 10 go back to Meyn and Tweedie [Meyn and Tweedie (1993b)] for time-homogeneous Markov chains and Seidler [Seidler (1997)] for time-homogeneous Markov processes.

Interdependence

From the above discussion it is clear how the chapters in this book are related. Chapter 2 is a prerequisite for all the others except Chapter 8. Chapter 3 contains the proofs of the main results in Chapter 2; it can be skipped at a first reading. Chapter 4 contains material very much related to the contents of Chapter 2. Chapter 6 is a direct continuation of Chapter 5, and is somewhat difficult to read and comprehend without the knowledge of the contents of Chapter 5. Chapter 7 is more or less independent of the other chapters in Part 3. For a big part Chapter 8 is independent of the other chapters: most of the results are phrased and proved for a finite-dimensional state space. The chapters 9 and 10 are very much interrelated. Some results in Chapter 9 are based on results in Chapter 10. In particular this is true for those results which use the existence of an invariant measure. A complete proof of existence and uniqueness is given in Chapter 10 Theorem 10.12. As a general prerequisite for understanding and appreciating this book a thorough knowledge of probability theory, in particular the concept of the Markov property, combined with a comprehensive notion of functional analysis is very helpful. On the other hand most topics are explained from scratch.

Acknowledgement

Part of this work was presented at a Colloquium at the University of Gent, October 14, 2005, at the occasion of the 65th birthday of Richard Delanghe and appeared in a very preliminary form in [Van Casteren (2005b)]. Some results were also presented at the University of Clausthal, at the occasion of Michael Demuth's 60th birthday September 10–11, 2006, and at a Conference in Marrakesh, Morocco, "Marrakesh World Conference on Differential Equations and Applications", June 15–20, 2006. Some of this work was also presented at a Conference on "The Feynman Integral and Related Topics in Mathematics and Physics: In Honor of the 65th Birthdays of Gerry Johnson and David Skoug", Lincoln, Nebraska, May 12–14, 2006. Finally, another preliminary version was delivered as a lecture during a Conference on Evolution Equations, in memory of G. Lumer, at the Universities of Mons and Valenciennes, August 28–September 1, 2006. The author also has presented some of this material during a colloquium at the University of Amsterdam (December 21, 2007), and at the AMS Special Session on the Feynman Integral in Mathematics and Physics, II, on January 9, 2008, in the Convention Center in San Diego, CA.

The author is obliged to the University of Antwerp (UA) and FWO Flanders (Grant number 1.5051.04N) for their financial and material support. He was also very fortunate to have discussed part of this material with Karel in't Hout (University of Antwerp), who provided some references with a crucial result about a surjectivity property of one-sided Lipschitz mappings: see Theorem 1 in Croezeix et al [Crouzeix *et al.* (1983)]. Some aspects concerning this work, like backward stochastic differential equations, were at issue during a conservation with Étienne Pardoux (CMI, Université de Provence, Marseille); the author is grateful for his comments and advice. The author is indebted to J.-C. Zambrini (Lisboa) for interesting discussions on the subject and for some references. In addition, the information and explanation given by Willem Stannat (Technical University Darmstadt) while he visited Antwerp are gratefully acknowledged. In particular this is true for topics related to asymptotic stability: see Chapter 9. The author is very much obliged to Natalia Katilova who has given the ideas of Chapter 8; she is to be considered as a co-author of this chapter. Moreover, the author is grateful for several anonymous referees who read a preliminary version of the manuscript and who have supplied a number of remarks and references which improved the quality of the book considerably. He also wants to express his sincere gratitude to Jerry Goldstein for his support and practical assistance. In addition, the author is grateful to the editorial staff of WSPC who agreed to publish this work. Finally, this work was part of the ESF program "Global".

Key words and phrases, subject classification

Some key words and phrases are: backward stochastic differential equation, parabolic equations of second order, Markov processes, Markov chains, ergodicity conditions, Orey's theorem, theorem of Chacon-Ornstein, invariant measure, long time behavior, Korovkin properties, maximum principle, Kolmogorov operator, squared gradient operator, martingale theory, bounded analytic semigroups.
AMS Subject classification [2010]: 35K20, 37L40, 46E10, 46E27, 47D07, 47D08, 60F05, 60G46, 60H99, 60J25.

Antwerp, *Jan A. Van Casteren*
June 2010

This page is intentionally left blank

Contents

PART 1
Introduction

This page is intentionally left blank

Chapter 1

Introduction: Stochastic differential equations

Some pertinent topics in the present chapter consist of a discussion on martingale theory, and a few relevant results on stochastic differential equations in spaces of finite as well as infinite dimension. This chapter also services as a motivation for the remaining part of the book. In particular unique weak solutions to stochastic differential equations give rise to strong Markov processes whose one-dimensional distributions are governed by the corresponding second order parabolic type differential equation. Some attention is paid to stochastic differential equations in infinite dimensions: see §1.2.

1.1 Weak and strong solutions to stochastic differential equations

In this section we discuss weak and strong solutions to stochastic differential equations. Basically, the material in this section is taken from [Ikeda and Watanabe (1998)]. We begin with a martingale characterization of Brownian motion. First we give a definition of Brownian motion. In the sequel $p_{0,d}(t,x,y)$ stands for the classical Gaussian kernel:

$$p_{0,d}(t,x,y) = \frac{1}{\left(\sqrt{2\pi t}\right)^d} \exp\left(-\frac{|x-y|^2}{2t}\right). \tag{1.1}$$

Definition 1.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $(\mathcal{F}_t)_{t\geqslant 0}$. A d-dimensional Brownian motion is a $\mathbb{P}$-almost surely continuous process $\{B(t) = (B_1(t), \dots, B_d(t)) : t \geqslant 0\}$, which is adapted to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, such that for $0 < t_1 < t_2 < \cdots < t_n < \infty$ and for C any Borel subset of $\left(\mathbb{R}^d\right)^n$ the following equality holds:

$$\begin{aligned}
&\mathbb{P}\left[\left(B\left(t_1\right)-B(0), \ldots, B\left(t_n\right)-B(0)\right) \in C\right] \\
&=\int \cdots \int_C p_{0,d}\left(t_n-t_{n-1}, x_{n-1}, x_n\right) \cdots p_{0,d}\left(t_2-t_1, x_1, x_2\right) p_{0,d}\left(t_1, 0, x_1\right) \\
&\qquad dx_1 \ldots dx_n .
\end{aligned} \tag{1.2}$$

This process is called a d-dimensional Brownian motion with initial distribution μ if for $0<t_1<t_2<\cdots<t_n<\infty$ and every Borel subset of $\left(\mathbb{R}^d\right)^{n+1}$ the following equality holds:

$$\begin{aligned}
&\mathbb{P}\left[\left(B(0), B\left(t_1\right), \ldots, B\left(t_n\right)\right) \in C\right] \\
&=\int \cdots \int_C p_{0,d}\left(t_n-t_{n-1}, x_{n-1}, x_n\right) \cdots p_{0,d}\left(t_2-t_1, x_1, x_2\right) p_{0,d}\left(t_1, x_0, x_1\right) \\
&\qquad d\mu\left(x_0\right)\, dx_1 \ldots dx_n .
\end{aligned} \tag{1.3}$$

For the definition of $p_{0,d}(t, x, y)$ see formula (1.1) above. By definition a filtration $\left(\mathcal{F}_t\right)_{t \geqslant 0}$ is an increasing family of σ-fields, i.e. $0 \leqslant t_1 \leqslant t_2<\infty$ implies $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$. The process of Brownian motion $\{B(t): t \geqslant 0\}$ is said to be adapted to the filtration $\left(\mathcal{F}_t\right)_{t \geqslant 0}$ if for every $t \geqslant 0$ the variable $B(t)$ is $\mathcal{F}_t$-measurable. It is assumed that the $\mathbb{P}$-negligible sets belong to $\mathcal{F}_0$. The following result we owe to Lévy.

Theorem 1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration (or reference system) $\left(\mathcal{F}_t\right)_{t \geqslant 0}$. Suppose $\mathcal{F}$ is the σ-field generated by $\bigcup_{t \geqslant 0} \mathcal{F}_t$ augmented with the $\mathbb{P}$-zero stes, and suppose $\mathcal{F}_t$ is continuous from the right: $\mathcal{F}_t=\bigcap_{s>t} \mathcal{F}_s$ for all $t \geqslant 0$. Let $\left\{M(t)=\left(M_1(t), \ldots, M_d(t)\right): t \geqslant 0\right\}$ be an $\mathbb{R}^d$-valued local $\mathbb{P}$-almost surely continuous martingale with the property that the quadratic covariation processes $t \mapsto\left\langle M_i, M_j\right\rangle(t)$ satisfy*

$$\left\langle M_i, M_j\right\rangle(t)=\delta_{i,j} t, \quad 1 \leqslant i, j \leqslant d. \tag{1.4}$$

Then $\{M(t): t \geqslant 0\}$ is d-dimensional Brownian motion with initial distribution given by $\mu(B)=\mathbb{P}[M(0) \in B]$, $B \in \mathcal{B}_{\mathbb{R}^d}$, the Borel field of $\mathbb{R}^d$.

It follows that the finite-dimensional distributions of the process $t \mapsto M(t)$ are given by:

$$\begin{aligned}
&\mathbb{P}\left[M\left(t_1\right) \in B_1, \ldots, M\left(t_n\right) \in B_n\right] \\
&=\int\left(\int_{B_1} \cdots \int_{B_n} p_{0,d}\left(t_n-t_{n-1}, x_{n-1}, x_n\right) \cdots p_{0,d}\left(t_2-t_1, x_1, x_2\right)\right. \\
&\qquad \left. p_{0,d}\left(t_1, x, x_1\right)\, dx_n \cdots dx_1\right) d\mu(x).
\end{aligned}$$

Proof. [Proof of Theorem 1.1.] Let $\xi \in \mathbb{R}^d$ be arbitrary. First we show that it suffices to establish the equality:

$$\mathbb{E}\left[e^{-i\langle\xi, M(t)-M(s)\rangle} \mid \mathcal{F}_s\right] = e^{-\frac{1}{2}|\xi|^2(t-s)}, \quad t > s \geqslant 0. \tag{1.5}$$

For suppose that (1.5) is true for all $\xi \in \mathbb{R}^d$. Then, by standard approximation arguments, it follows that the variable $M(t) - M(s)$ is $\mathbb{P}$-independent of $\mathcal{F}_s$. In other words the process $t \mapsto M(t)$ possesses independent increments. Since the Fourier transform of the function $y \mapsto p_{0,d}(t-s, 0, y)$ is given by

$$\int_{\mathbb{R}^d} e^{-i\langle\xi, y\rangle} p_{0,d}(t-s, 0, y)\, dy = e^{-\frac{1}{2}|\xi|^2(t-s)}$$

it also follows that the distribution of $M(t) - M(s)$ is given by

$$\mathbb{P}[M(t) - M(s) \in B] = \int_B p_{0,d}(t-s, 0, y)\, dy. \tag{1.6}$$

Moreover, for $0 < t_1 < \cdots < t_n$ we also have

$$\begin{aligned}
&\mathbb{P}[M(0) \in B_0,\ M(t_1) - M(0) \in B_1, \ldots, M(t_n) - M(t_{n-1}) \in B_n] \\
&= \mathbb{P}[M(0) \in B_0]\, \mathbb{P}[M(t_1) - M(0) \in B_1] \cdots \mathbb{P}[M(t_n) - M(t_{n-1}) \in B_n] \\
&= \int_{B_0} \int_{B_1} \cdots \int_{B_n} p_{0,d}(t_1, 0, y_1) \cdots p_{0,d}(t_n - t_{n-1}, 0, y_n)\, d\mu(y_0)\, dy_1 \cdots dy_n.
\end{aligned}$$

Here $B_0, \ldots, B_n$ are Borel subsets of $\mathbb{R}^d$. Hence, if B is a Borel subset of $\underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{n+1\text{times}}$, then it follows that

$$\begin{aligned}
&\mathbb{P}[(M(0), M(t_1) - M(0), \ldots, M(t_n) - M(t_{n-1})) \in B] \\
&= \int \cdots \int_B p_{0,d}(t_1, 0, y_1) \cdots p_{0,d}(t_n - t_{n-1}, 0, y_n)\, d\mu(y_0)\, dy_1 \cdots dy_n.
\end{aligned} \tag{1.7}$$

Next we compute the joint distribution of $(M(0), M(t_1), \ldots, M(t_n))$ by employing (1.7). Define the linear map $\ell : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \to \mathbb{R}^d \times \cdots \times \mathbb{R}^d$ by $\ell(x_0, x_1, \ldots, x_n) = (x_0, x_1 - x_0, x_2 - x_1, \ldots, x_n - x_{n-1})$. Let B be a Borel subset of $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$. By (1.7) we get

$$\begin{aligned}
&\mathbb{P}[(M(0), \ldots, M(t_n)) \in B] \\
&= \mathbb{P}[\ell(M(0), \ldots, M(t_n)) \in \ell(B)] \\
&= \mathbb{P}[(M(0), M(t_1) - M(0), \ldots, M(t_n) - M(t_{n-1})) \in \ell(B)] \\
&= \int \ldots \int_{\ell(B)} p_{0,d}(t_1, 0, y_1) \cdots p_{0,d}(t_n - t_{n-1}, 0, y_n)\, d\mu(y_0)\, dy_1 \cdots dy_n
\end{aligned}$$

(change of variables: $(y_0, y_1, \ldots, y_n) = \ell(x_0, x_1, \ldots, x_n)$)

$$= \int \cdots \int_B p_{0,d}(t_1, x_0, x_1) \cdots p_{0,d}(t_n - t_{n-1}, x_{n-1}, x_n)\, d\mu(x_0)\, dx_1 \cdots dx_n. \tag{1.8}$$

In order to complete the proof of Theorem 1.1 from equality (1.8) it follows that it is sufficient to establish the equality in (1.5). Therefore, fix $\xi \in \mathbb{R}^d$ and $t > s \geqslant 0$. An application of Itô's lemma to the function $x \mapsto e^{-i\langle \xi, x\rangle}$ yields

$$e^{-i\langle \xi, M(t)\rangle} - e^{-i\langle \xi, M(s)\rangle}$$

$$= -i \sum_{j=1}^{d} \xi_j \int_s^t e^{-i\langle \xi, M(\tau)\rangle} dM_j(\tau) - \frac{1}{2} \sum_{j,k=1}^{d} \xi_j \xi_k \int_s^t e^{-i\langle \xi, M(\tau)\rangle} d\langle M_j, M_k\rangle(\tau)$$

(formula (1.4))

$$= -i \sum_{j=1}^{d} \xi_j \int_s^t e^{-i\langle \xi, M(\tau)\rangle} dM_j(\tau) - \frac{1}{2} |\xi|^2 \int_s^t e^{-i\langle \xi, M(\tau)\rangle}\, d\tau. \tag{1.9}$$

Hence, from (1.9) it follows that

$$e^{-i\langle \xi, M(t)-M(s)\rangle} - 1 \tag{1.10}$$

$$= -i \sum_{j=1}^{d} \xi_j \int_s^t e^{-i\langle \xi, M(\tau)-M(s)\rangle} dM_j(\tau) - \frac{1}{2} |\xi|^2 \int_s^t e^{-i\langle \xi, M(\tau)-M(s)\rangle}\, d\tau.$$

Since the processes

$$t \mapsto \int_s^t e^{-i\langle \xi, M(\tau)-M(s)\rangle} dM_j(s),\ t \geqslant s,\ 1 \leqslant j \leqslant d,$$

are local martingales, from (1.10) we infer by (possibly) using a stopping time argument that

$$\mathbb{E}\left[e^{-i\langle \xi, M(t)-M(s)\rangle} \mid \mathcal{F}_s\right] = 1 - \frac{1}{2} |\xi|^2 \int_s^t \mathbb{E}\left[e^{-i\langle \xi, M(\tau)-M(s)\rangle} \mid \mathcal{F}_s\right] d\tau. \tag{1.11}$$

Next, let $v(t)$, $t \geqslant s$, be given by

$$v(t) = \int_s^t \mathbb{E}\left[e^{-i\langle \xi, M(\tau)-M(s)\rangle} \mid \mathcal{F}_s\right] d\tau.$$

Then $v(s) = 0$, and (1.11) implies

$$v'(t) + \frac{1}{2} |\xi|^2 v(t) = 1. \tag{1.12}$$

From (1.12) we infer

$$\frac{d}{dt}\left(e^{\frac{1}{2}(t-s)|\xi|^2}v(t)\right) = \left(\frac{1}{2}\left|\xi\right|^2 v(t) + v'(t)\right) e^{\frac{1}{2}(t-s)|\xi|^2}$$
$$= e^{\frac{1}{2}(t-s)|\xi|^2}. \tag{1.13}$$

The equality in (1.13) implies:

$$e^{\frac{1}{2}(t-s)|\xi|^2}v(t) - v(s) = \frac{2}{|\xi|^2}\left(e^{\frac{1}{2}(t-s)|\xi|^2} - 1\right),$$

and thus we see

$$v'(t) + \frac{1}{2}v(s)e^{-\frac{1}{2}(t-s)|\xi|^2} = e^{-\frac{1}{2}(t-s)|\xi|^2}. \tag{1.14}$$

Since $v(s) = 0$ (1.14) results in

$$\mathbb{E}\left[e^{-i\langle\xi, M(\tau)-M(s)\rangle} \mid \mathcal{F}_s\right] = v'(t) = e^{-\frac{1}{2}(t-s)|\xi|^2}. \tag{1.15}$$

The equality in (1.15) is the same as the one in (1.5). By the above arguments this completes the proof of Theorem 1.1. □

As a corollary to Theorem 1.1 we get the following one-dimensional result due to Lévy.

Corollary 1.1. *Let $\{M(t) : t \geqslant 0\}$ be an almost surely continuous local martingale in $\mathbb{R}$ such that the process $t \mapsto M(t)^2 - t$ is a local martingale as well. Then the process $\{M(t) : t \geqslant 0\}$ is a Brownian motion with initial distribution given by $\mu(B) = \mathbb{P}[M(0) \in B]$, $B \in \mathcal{B}_{\mathbb{R}}$.*

Proof. Since $M(t)^2 - t$ is a local martingale, it follows that the quadratic variation process $t \mapsto \langle M, M\rangle(t)$ satisfies $\langle M, M\rangle(t) = t$, $t \geqslant 0$. So the result in Corollary 1.1 follows from Theorem 1.1. □

The following result contains a d-dimensional version of Corollary 1.1.

Theorem 1.2. *Let $\{M(t) = (M_1(t), \ldots, M_{d'}(t)) : t \geqslant 0\}$ be a continuous local martingale with covariation process given by*

$$\langle M_j, M_k\rangle(t) = \int_0^t \Phi_{j,k}(s)ds, \ 1 \leqslant j, \, k \leqslant d'. \tag{1.16}$$

Let the $d' \times d$-matrix process $\{\chi(t) : t \geqslant 0\}$ be such that $\chi(t)\Phi(t)\chi(t)^* = I$, where I is the $d \times d$ identity matrix. Put $B(t) = \int_0^t \chi(s)\,dM(s)$. This integral should be interpreted in Itô sense. Then the process $t \mapsto B(t)$ is d-dimensional Brownian motion. Put $\Psi(t) = \Phi(t)\chi(t)^*$, and suppose that $\Psi(t)\chi(t) = I$, the $d' \times d'$ identity matrix. Then $M(t) - M(0) = \int_0^t \Psi(s)\,dB(s)$.

Remark 1.1. Since

$$\chi(t)\left(\Phi(t)\chi(t)^*\chi(t)-I\right)=\left(\chi(t)\Phi(t)\chi(t)^*-I\right)\chi(t)=0$$

we see that the second equality in $\Psi(t)\chi(t)=\Phi(t)\chi(t)^*\chi(t)=I$ is only possible if we assume $d=d'$. Of course here we take the dimensions of the null and range space of the matrix $\chi(t)$ into account.

Proof. [Proof of Theorem 1.2.] Fix $1\leqslant i,\ j\leqslant d$. We shall calculate the quadratic covariation process

$$\begin{aligned}\langle B_i,B_j\rangle(t)&=\left\langle\sum_{k=1}^{d'}\int_0^{(\cdot)}(\chi(s))_{i,k}\,dM_k(s),\sum_{l=1}^{d'}\int_0^{(\cdot)}(\chi(s))_{j,l}\,dM_l(s)\right\rangle(t)\\&=\sum_{k=1}^{d'}\sum_{l=1}^{d'}\int_0^t(\chi(s))_{i,k}(\chi(s))_{j,l}\,\Phi(s)_{i,j}ds\\&=\int_0^t\left(\chi(s)\Phi(s)\chi(s)^*\right)_{i,j}ds=t\delta_{i,j}.\end{aligned}\tag{1.17}$$

From Theorem 1.1 and (1.17) we see that the process $t\mapsto B(t)$ is a Brownian motion. This proves the first part of Theorem 1.2. Next we calculate

$$\int_0^t\Psi(s)\,dB(s)=\int_0^t\Psi(s)\chi(s)\,dM(s)=\int_0^t dM(s)=M(t)-M(0).\tag{1.18}$$

This completes the proof of Theorem 1.2. □

In the following theorem the symbols $\sigma_{i,j}$ and b_j, $1\leqslant i,\ j\leqslant d$, stand for real-valued locally bounded Borel measurable functions defined on $[0,\infty)\times\mathbb{R}^d$. The matrix $\left(a_{i,j}(s,x)\right)_{i,j=1}^d$ is defined by

$$a_{j,k}(s,x)=\sum_{k=1}^d\sigma_{i,k}(s,x)\sigma_{j,k}(s,x)=\left(\sigma(s,x)\sigma^*(s,x)\right)_{i,j}.$$

For $s\geqslant 0$, the operator $L(s)$ is defined on $C^2\left(\mathbb{R}^d\right)$ with values in the space of locally bounded Borel measurable functions:

$$L(s)f(x)=\frac{1}{2}\sum_{i,j=1}^d a_{i,j}(s,x)\,D_iD_jf(x)+\sum_{j=1}^d b_j(s,x)D_jf(x).\tag{1.19}$$

The following theorem shows the close relationship between weak solutions and solutions to the martingale problem.

Theorem 1.3. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space with a right-continuous filtration $(\mathcal{F}_t)_{t\geqslant 0}$. Let $\{X(t)=(X_1(t),\dots,X_d(t)):\ t\geqslant 0\}$ be a d-dimensional continuous adapted process. Then the following assertions are equivalent:*

(i) *For every* $f \in C^2\left(\mathbb{R}^d\right)$ *the process*

$$t \mapsto f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L(s) f\left(X(s)\right)\, ds \tag{1.20}$$

is a local martingale.

(ii) *The processes*

$$t \mapsto M_j(t) := X_j(t) - \int_0^t b_j\left(s, X(s)\right)\, ds, \; t \geqslant 0, \quad 1 \leqslant j \leqslant d, \tag{1.21}$$

are local martingales with covariation processes

$$t \mapsto \left\langle M_i, M_j \right\rangle(t) = \int_0^t a_{i,j}\left(s, X(s)\right) ds, \; t \geqslant 0, \quad 1 \leqslant i, j \leqslant d. \tag{1.22}$$

(iii) *On an extended probability space* $(\Omega \times \Omega', \mathcal{F}_t \otimes \mathcal{F}'_t, \mathbb{P} \times \mathbb{P}')$ *there exists a Brownian motion* $\{B(t) : t \geqslant 0\}$ *starting at 0 such that*

$$X(t) = X(0) + \int_0^t b\left(s, X(s)\right)\, ds + \int_0^t \sigma\left(s, X(s)\right)\, dB(s), \quad t \geqslant 0. \tag{1.23}$$

Notice that under the conditions of Theorem 1.3 the martingale problem need not be uniquely solvable: for some more details the reader is referred to Remark 2.12 in Chapter 2.

The following corollary easily follows from Theorem 1.3. It establishes a close relationship between unique weak solutions to stochastic differential equations and unique solutions to the martingale problem.

Corollary 1.2. *Let the notation and hypotheses be as in Theorem 1.3. Put* $\Omega = C\left([0,\infty), \mathbb{R}^d\right)$, *and* $X(t)(\omega) = \omega(t)$, $t \geqslant 0$, $\omega \in \Omega$. *Fix* $x \in \mathbb{R}^d$. *Then the following assertions are equivalent:*

(i) *There exists a unique probability measure* $\mathbb{P}$ *on* $\mathcal{F}$ *such that* $\mathbb{P}\left[X(0) = x\right] = 1$, *and the process*

$$f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L(s) f\left(X(s)\right)\, ds$$

is a $\mathbb{P}$*-martingale for all* C^2*-functions* f *with compact support.*

(ii) *The stochastic integral equation*

$$X(t) = x + \int_0^t \sigma\left(s, X(s)\right)\, dB(s) + \int_0^t b\left(s, X(s)\right) ds$$

has unique weak solutions.

Proof. [Proof of Theorem 1.3.] (i) $\Longrightarrow$ (ii). With $f_j(x_1,\dots,x_d) = x_j$, $1 \leqslant j \leqslant d$, assertion (i) implies that the process

$$M_j(t) = X_j(t) - \int_0^t b_j(s, X(s))\, ds = f_j(X(t)) - \int_0^t L(s) f_j(X(s))\, ds \tag{1.24}$$

is a local martingale. We will show that the processes

$$\left\{ M_i(t) M_j(t) - \int_0^t a_{i,j}(s, X(s))\, ds : t \geqslant 0 \right\}, \quad 1 \leqslant i,\, j \leqslant d,$$

are local martingales as well. To this end fix $1 \leqslant i,\, j \leqslant d$, and define the function $f_{i,j} : \mathbb{R}^d \to \mathbb{R}$ by $f_{i,j}(x_1,\dots,x_d) = x_i x_j$. From (i) it follows that the process

$$\left\{ X_i(t) X_j(t) - \int_0^t \left(a_{i,j}(s, X(s)) + b_i(s, X(s)) X_j(s) + b_j(s, X(s)) X_i(s) \right) ds \right\}$$

is a local martingale. For brevity we write

$$\begin{aligned} \alpha_{i,j}(s) &= a_{i,j}(s, X(s)), \quad \beta_j(s) = b_j(s, X(s)), \quad \beta_i(s) = b_j(s, X(s)), \\ M_i(s) &= X_i(s) - \int_0^s \beta_i(\tau)\, d\tau, \quad M_j(s) = X_i(s) - \int_0^s \beta_j(\tau)\, d\tau, \\ M_{i,j}(s) &= X_i(s) X_j(s) - \int_0^s \left(\beta_i(\tau) X_j(\tau) + \beta_j(\tau) X_i(\tau) + \alpha_{i,j}(\tau) \right) d\tau. \end{aligned} \tag{1.25}$$

Then the processes M_i and $M_{i,j}$ are local martingales. Moreover, we have

$$\begin{aligned} &\left(M_i(t) + \int_0^t \beta_i(s)\, ds \right) \left(M_j(t) + \int_0^t \beta_j(s)\, ds \right) = X_i(t) X_j(t) \\ &= \int_0^t \left(\beta_i(\tau) X_j(\tau) + \beta_j(\tau) X_i(\tau) + \alpha_{i,j}(\tau) \right) d\tau + M_{i,j}(t) \\ &= \int_0^t \left(\beta_i(\tau) \left(X_j(\tau) - M_j(\tau) \right) + \beta_j(\tau) \left(X_i(\tau) - M_i(\tau) \right) + \alpha_{i,j}(\tau) \right) d\tau \\ &\qquad + \int_0^t \left(\beta_i(\tau) M_j(\tau) + \beta_j(\tau) M_i(\tau) \right) d\tau + M_{i,j}(t) \\ &= \int_0^t \beta_i(\tau) \left(X_j(\tau) - M_j(\tau) \right) d\tau + \int_0^t \beta_j(\tau) \left(X_i(\tau) - M_i(\tau) \right) d\tau \\ &\qquad + \int_0^t \alpha_{i,j}(\tau)\, d\tau + \int_0^t \left(\beta_i(\tau) M_j(\tau) + \beta_j(\tau) M_i(\tau) \right) d\tau + M_{i,j}(t) \\ &= \int_0^t \beta_i(\tau) \int_0^\tau \beta_j(s)\, ds\, d\tau + \int_0^t \beta_j(\tau) \int_0^\tau \beta_i(s)\, ds\, d\tau \end{aligned}$$

$$+\int_0^t \alpha_{i,j}(\tau)\,d\tau + \int_0^t \left(\beta_i(\tau)M_j(\tau) + \beta_j(\tau)M_i(\tau)\right)\,d\tau + M_{i,j}(t)$$

(in the second integral the integration over s and τ are exchanged)

$$\begin{aligned}
&= \iint_{0<s<\tau<t} \beta_i(\tau)\beta_j(s)\,d\tau\,ds + \iint_{0<\tau<s<t} \beta_i(\tau)\beta_j(s)\,d\tau\,ds \\
&\qquad + \int_0^t \alpha_{i,j}(\tau)\,d\tau + \int_0^t \left(\beta_i(\tau)M_j(\tau) + \beta_j(\tau)M_i(\tau)\right)\,d\tau + M_{i,j}(t) \\
&= \int_0^t \beta_i(\tau)\,d\tau \int_0^t \beta_j(s)\,ds + \int_0^t \alpha_{i,j}(s)\,ds + M_{i,j}(t) \\
&\qquad + \int_0^t \left(\beta_i(s)M_j(s) + \beta_j(s)M_i(s)\right)\,ds. \qquad (1.26)
\end{aligned}$$

Consequently, from (1.26) we see

$$\begin{aligned}
&M_i(t)M_j(t) - \int_0^t \alpha_{i,j}(s)\,ds \\
&= M_{i,j}(t) - \int_0^t \left(\beta_i(s)\left(M_j(t) - M_j(s)\right) + \beta_j(s)\left(M_i(t) - M_i(s)\right)\right)\,ds.
\end{aligned} \qquad (1.27)$$

It is readily verified that the processes

$$\int_0^t \beta_i(s)\left(M_j(t) - M_j(s)\right)\,ds \quad \text{and} \quad \int_0^t \beta_j(s)\left(M_i(t) - M_i(s)\right)\,ds$$

are local martingales. It follows that the process

$$\left\{M_i(t)M_j(t) - \int_0^t \alpha_{i,j}(s)\,ds : \; t \geqslant 0\right\}$$

is a local martingale. So that the covariation process $\langle M_i, M_j\rangle$ is given by $\langle M_i, M_j\rangle(t) = \int_0^t \alpha_{i,j}(s)\,ds$.

(ii) $\Longrightarrow$ (iii). This implication follows from an application of Theorem 1.2 with $\Phi_{i,j}(t) = a_{i,j}(t, X(t))$, and $\chi(t) = \sigma(t, X(t))^{-1}$. If the matrix process $\sigma(t, X(t))$ is not invertible we proceed as follows. First we choose a Brownian motion which is independent of $(\Omega, \mathcal{F}_t, \mathbb{P})$ and which lives on the probability space $(\Omega', \mathcal{F}'_t, \mathbb{P}')$. The probability spaces $(\Omega, \mathcal{F}_t, \mathbb{P})$ and $(\Omega', \mathcal{F}'_t, \mathbb{P}')$ are coupled by employing a standard extension of the original probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$. This extension is denoted by $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$, where $\widetilde{\Omega} = \Omega \times \Omega'$, $\widetilde{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'_t$, and $\widetilde{\mathbb{P}} = \mathbb{P} \times \mathbb{P}'$. Finally, $\widetilde{B}'(\omega, \omega') = B'(\omega')$, $(\omega, \omega') \in \Omega \times \Omega'$. We have a martingale $M(s)$, $0 \leqslant s \leqslant t$, on $(\Omega, \mathcal{F}_t, \mathbb{P})$ with the properties of Assertion (ii). We introduce the matrix processes $\widetilde{\psi}_\varepsilon(s)$, $\varepsilon > 0$, $E_R(s)$, and $E_N(s)$ as follows

$$\widetilde{\psi}_\varepsilon(s) = \sigma^*\left(s, X(s)\right)\left(\sigma\left(s, X(s)\right)\sigma^*\left(s, X(s)\right) + \varepsilon I\right)^{-1}$$

$$E_R(s) = \lim_{\varepsilon\downarrow 0} \sigma^*\left(s, X(s)\right)\left(\sigma\left(s, X(s)\right)\sigma^*\left(s, X(s)\right) + \varepsilon I\right)^{-1} \sigma\left(s, X(s)\right), \quad \text{and}$$

$$E_N(s) = I - E_R(s).$$

The matrix $E_R(s)$ can be considered as an orthogonal projection on the range of the matrix $\sigma^*\left(s, X(s)\right)\sigma\left(s, X(s)\right)$, and $E_N(s)$ as an orthogonal projection on its null space. More precisely, $E_R(s)\sigma^*\left(s, X(s)\right) = \sigma^*\left(s, X(s)\right)$, and $\sigma\left(s, X(s)\right)E_N(s) = 0$. In terms of these processes we define the following process:

$$B(s) = \lim_{\varepsilon\downarrow 0}\int_0^s \widetilde{\psi}_\varepsilon(\tau)\, dM(\tau) + \int_0^s E_N(\tau)\, dB'(\tau).$$

Next we will prove that the process $s \mapsto B(s)$ is a Brownian motion, and that $M(s) = \int_0^s \sigma\left(\tau, X(\tau)\right)\, dB(\tau)$. Put

$$B_\varepsilon(s) = \int_0^s \widetilde{\psi}_\varepsilon(\tau)\, dM(\tau) + \int_0^s E_N(\tau)\, dB'(\tau).$$

Then we have:

$$\begin{aligned}
&\left\langle B_{\varepsilon,j_1}, B_{\varepsilon,j_2}\right\rangle(s) \\
&= \sum_{k_1,k_2,\ell=1}^{d} \int_0^s \widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)\widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)\sigma_{k_1,\ell}\left(\tau, X(\tau)\right)\sigma_{k_2,\ell}\left(\tau, X(\tau)\right)\, d\tau \\
&\quad + \sum_{k=1}^{d}\int_0^s \widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)E_{N,j_2,K_1}(\tau)\, d\left\langle M_{k_1}, B'_k\right\rangle(\tau) \\
&\quad + \sum_{k=1}^{d}\int_0^s \widetilde{\psi}_{\varepsilon,j_2,k_1}(\tau)E_{N,j_1,K_1}(\tau)\, d\left\langle M_{k_1}, B'_k\right\rangle(\tau) \\
&\quad + \sum_{k=1}^{d}\int_0^s E_{N,j_1,k}(\tau)E_{N,j_2,k}(\tau)\, d\tau
\end{aligned}$$

(the processes M and B' are $\widetilde{P}$-independent)

$$\begin{aligned}
&= \int_0^s \left(\widetilde{\psi}_\varepsilon(\tau)\sigma\left(\tau, X(\tau)\right)\sigma^*\left(\tau, X(\tau)\right)\widetilde{\psi}^*_\varepsilon(\tau)\right)_{j_1,j_2} d\tau \\
&\quad + \int_0^s \left(E_N(\tau)E_N^*(\tau)\right)_{j_1,j_2} d\tau.
\end{aligned} \tag{1.28}$$

From (1.28) we infer by continuity and the definition of $E_R(\tau)$ that

$$\begin{aligned}\langle B_{j_1}, B_{j_2}\rangle(s) &= \lim_{\varepsilon\downarrow 0}\langle B_{\varepsilon,j_1}, B_{\varepsilon,j_2}\rangle(s)\\ &= \int_0^s \left(E_R(\tau)E_R^*(\tau)\right)_{j_1,j_2} d\tau + \int_0^s \left(E_N(\tau)E_N^*(\tau)\right)_{j_1,j_2} d\tau\\ &= \int_0^s \left(E_R(\tau)E_R^*(\tau) + E_N(\tau)E_N^*(\tau)\right)_{j_1,j_2} d\tau\end{aligned}$$

(the processes $E_R(\tau)$ and $E_N(\tau)$ are orthogonal projections such that $E_R(\tau) + E_N(\tau) = I$)

$$= \delta_{j_1,j_2} s. \tag{1.29}$$

From Lévy's theorem 1.1 it follows that the process $s \mapsto B(s)$, $0 \leqslant s \leqslant t$, is a Brownian motion. In order to finish the proof of the implication (ii) $\Longrightarrow$ (iii) we still have to prove the equality $M(s) = \int_0^s \sigma(\tau, X(\tau))\, dB(\tau)$. For brevity we write $\sigma(\tau) = \sigma(\tau, X(\tau))$. Then by definition and standard calculations with martingales we obtain:

$$\begin{aligned}&M(s) - \int_0^s \sigma(\tau)\, dB_\varepsilon(\tau)\\ &= M(s) - \int_0^s \sigma(\tau)\widetilde{\psi}_\varepsilon(\tau)\, dM(\tau) - \int_0^s \sigma(\tau)\, E_N(\tau)\, dB'(\tau)\\ &= \int_0^s \left(I - \sigma(\tau)\sigma^*(\tau)\left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1}\right) dM(\tau)\\ &= \varepsilon \int_0^s \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} dM(\tau).\end{aligned} \tag{1.30}$$

From (1.30) together with the fact that covariation process of the local martingale $M(s)$ is given by $\int_0^s \sigma(\tau)\sigma^*(\tau)\, d\tau$, it follows that the covariation matrix of the local martingale

$$M(s) - \int_0^s \sigma(\tau)\, dB_\varepsilon(\tau)$$

is given by

$$\varepsilon^2 \int_0^s \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} \sigma(\tau)\sigma^*(\tau)\left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} d\tau. \tag{1.31}$$

In addition, in spectral sense we have:

$$0 \leqslant \varepsilon^2 \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} \sigma(\tau)\sigma^*(\tau)\left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} \leqslant \frac{\varepsilon}{4} I, \tag{1.32}$$

and thus in L^2-sense we see

$$M(s) - \int_0^s \sigma(\tau)\, dB(\tau) = L^2\text{-}\lim_{\varepsilon\downarrow} \left(M(s) - \int_0^s \sigma(\tau) B_\varepsilon(\tau)\right) = 0. \tag{1.33}$$

The equality in (1.33) completes the proof of the implication (ii) $\longrightarrow$ (iii).

(iii) $\Longrightarrow$ (i). Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice continuously differentiable function. By Itô's lemma we get

$$
\begin{aligned}
& f(X(t)) - f(X(0)) - \int_0^t L(s) f(X(s))\, ds \\
&= \int_0^t \nabla f(X(s)) \cdot dX(s) + \frac{1}{2} \sum_{i,j=1}^d \int_0^t D_i D_j f(X(s))\, d\langle X_i, X_j\rangle(s) \\
&\quad - \int_0^t L(s) f(X(s))\, ds \\
&= \sum_{j=1}^d \int_0^t b_j(s, X(s))\, D_j f(X(s))\, ds \\
&\quad + \frac{1}{2} \sum_{i,j=1}^d \sum_{k=1}^d \int_0^t \sigma_{i,k}(s, X(s))\, \sigma_{j,k}(s, X(s))\, D_i D_j f(X(s))\, ds \\
&\quad + \int_0^t \nabla f(X(s))\, \sigma(s, X(s))\, dB(s) - \int_0^t L(s) F(X(s))\, ds \\
&= \int_0^t \nabla f(X(s))\, \sigma(s, X(s))\, dB(s).
\end{aligned}
\tag{1.34}
$$

The final expression in (1.34) is a local martingale. Hence (iii) implies (i).

This completes the proof of Theorem 1.3. □

Remark 1.2. The implication (ii) $\Longrightarrow$ (i) in Theorem 1.2 can also be proved directly by using Itô calculus. Let f be a C^2-function defined on $\mathbb{R}^d$. Then we have:

$$
\begin{aligned}
& f(X(t)) - f(X(0)) - \int_0^t L(s) f(X(s))\, ds \\
&= \int_0^t \nabla f(X(s))\, dX(s) + \frac{1}{2} \sum_{i,j=1}^d \int_0^t D_i D_j f(X(s))\, d\langle X_i, X_j\rangle(s) \\
&\quad - \int_0^t L(s) f(X(s))\, ds \\
&= \int_0^t \nabla f(X(s))\, dM(s) + \int_0^t \nabla f(X(s))\, b(s, X(s))\, ds \\
&\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t D_i D_j f(X(s))\, d\langle M_i, M_j\rangle(s) - \int_0^t L(s) f(X(s))\, ds
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^t \nabla f\left(X(s)\right)\, dM(s) + \int_0^t \nabla f\left(X(s)\right)\, b\left(s, X(s)\right)\, ds \\
&\qquad + \frac{1}{2}\sum_{i,j=1}^{d} \int_0^t D_i D_j f\left(X(s)\right) a_{i,j}\left(s, X(s)\right)\, ds - \int_0^t L(s) f\left(X(s)\right)\, ds \\
&= \int_0^t \nabla f\left(X(s)\right)\, dM(s). \qquad (1.35)
\end{aligned}
$$

Assertion (i) is a consequence of equality (1.35).

We also want to discuss the Cameron-Martin-Girsanov transformation of Wiener measure. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t\geqslant 0}$. In addition, let $\{B(t) : t \geqslant 0\}$ be a d-dimensional Brownian motion. Let b_j, c_j, $\sigma_{i,j}$ be Borel measurable locally bounded functions on $[0,\infty) \times \mathbb{R}^d$. Suppose that the stochastic differential equation

$$
X(t) = x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right)\, ds \qquad (1.36)
$$

has unique weak solutions. For more information on transformations of measures on Wiener space see e.g. [Üstünel and Zakai (2000a)].

Definition 1.2. The equation in (1.36) is said to have unique weak solutions, also called unique distributional solutions, provided that the finite-dimensional distributions of the process $X(t)$ which satisfy (1.36) do not depend on the particular Brownian motion $B(t)$ which occurs in (1.36). This is the case if and only if for any pair of Brownian motions

$$
\left\{\left(B(t) : t \geqslant 0\right), \left(\Omega, \mathcal{F}, \mathbb{P}\right)\right\} \quad \text{and} \quad \left\{\left(B'(t) : t \geqslant 0\right), \left(\Omega', \mathcal{F}', \mathbb{P}'\right)\right\}
$$

and any pair of adapted processes $\{X(t) : t \geqslant 0\}$ and $\{X'(t) : t \geqslant 0\}$ for which

$$
\begin{aligned}
X(t) &= x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right)\, ds \quad \text{and} \\
X'(t) &= x + \int_0^t \sigma\left(s, X'(s)\right) dB'(s) + \int_0^t b\left(s, X'(s)\right)\, ds
\end{aligned}
$$

it follows that the finite-dimensional distributions of the process $\{X(t) : t \geqslant 0\}$ relative to $\mathbb{P}$ coincide with the finite-dimensional distributions of the process $\{X'(t) : t \geqslant 0\}$ relative to $\mathbb{P}'$.

In particular this means that if in equation (1.37) below (for the process $Y(t)$) the process $B'(t)$ is a Brownian motion relative to a probability measure $\mathbb{P}'$, then the $\mathbb{P}'$-distribution of the process $Y(t)$ coincides with the

$\mathbb{P}$-distribution of the process $X(t)$ which satisfies (1.36). Next we will elaborate on this item. Suppose that the process $t \mapsto Y(t)$ satisfies the equation:

$$\begin{aligned} Y(t) &= x + \int_0^t \sigma(s, Y(s))\, dB(s) + \int_0^t \left(b(s, Y(s)) + \sigma(s, Y(s))\, c(s, Y(s))\right) ds \\ &= x + \int_0^t \sigma(s, Y(s))\, dB'(s) + \int_0^t b(s, Y(s))\, ds, \end{aligned} \tag{1.37}$$

where $B'(t) = B(t) + \int_0^t c(s, Y(s))\, ds$. The following proposition says that relative to a martingale transformation $\mathbb{P}'$ of the measure $\mathbb{P}$ (Girsanov or Cameron-Martin transformation) the process $t \mapsto B'(t)$ is a $\mathbb{P}'$-Brownian motion. More precisely, we introduce the local martingale $M'(t)$ and the corresponding measure $\mathbb{P}'$ by

$$M'(t) = \exp\left(-\int_0^t c(s, Y(s))\, dB(s) - \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds\right), \tag{1.38}$$

and

$$\mathbb{P}'[A] = \mathbb{E}\left[M'(t)\mathbf{1}_A\right], \quad A \in \mathcal{F}_t. \tag{1.39}$$

We also need the process $Z'(t)$ defined by

$$Z'(t) = -\int_0^t c(s, Y(s))\, dB(s) - \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds. \tag{1.40}$$

In addition, we have a need for a vector-valued function $c_1(t, y)$ satisfying $c(t, y) = c_1(t, y)\sigma(t, y)$. We assume that such a vector function $c_1(t, y)$ exists.

Proposition 1.1. *Suppose that the process $Y(t)$ satisfies the equation in (1.37). Let the processes $M'(t)$ and $Z'(t)$ be defined by (1.38) and (1.40) respectively. Then the following assertions are true:*

(1) The process $t \mapsto M'(t)$ is a local $\mathbb{P}$-martingale. It is a martingale provided that $\mathbb{E}[M'(t)] = 1$ for all $t \geqslant 0$.
(2) Fix $t > 0$. The variable $M'(t)$ only depends on the process $s \mapsto Y(s)$, $0 \leqslant s \leqslant t$.
(3) Suppose that the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale, and not just a local $\mathbb{P}$-martingale. Then $\mathbb{P}'$ can be considered as a probability measure on the σ-field generated by $\bigcup_{t>0} \mathcal{F}_t$.
(4) Suppose that the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale. Then the process $t \mapsto B'(t)$ is a Brownian motion relative to $\mathbb{P}'$.

Proof. (1). From Itô calculus we get

$$M'(t) - M'(0) = -\int_0^t M'(s)c\,(s, Y(s))\,dB(s),$$

and hence Assertion (1) follows, because stochastic integrals with respect to Brownian motion are local martingales. Next we choose a sequence of stopping times τ_n which increase to ∞ $\mathbb{P}$-almost surely, and which are such that the processes $t \mapsto M'(t \wedge \tau_n)$ are genuine martingales. Then we see $\mathbb{E}[M'(t \wedge \tau_n)] = 1$ for all $n \in \mathbb{N}$ and $t \geqslant 0$. Fix $t_2 > t_1$. Since the processes $t \mapsto M'(t \wedge \tau_n)$, $n \in \mathbb{N}$, are $\mathbb{P}$-martingales, we see that

$$\mathbb{E}\left[M'(t_2 \wedge \tau_n) \,\middle|\, \mathcal{F}_{t_1}\right] = M'(t_1 \wedge \tau_n) \quad \mathbb{P}\text{-almost surely.} \tag{1.41}$$

In (1.41) we let $n \to \infty$, and apply Scheffé's theorem to conclude that

$$\mathbb{E}\left[M'(t_2) \,\middle|\, \mathcal{F}_{t_1}\right] = M'(t_1) \quad \mathbb{P}\text{-almost surely.} \tag{1.42}$$

The equality in (1.42) shows that the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale provided that $\mathbb{E}[M'(t)] = 1$ for all $t \geqslant 0$. This completes the proof of Assertion (1).

(2). This assertion follows from the following calculation:

$$\begin{aligned}
&Z'(t)\\
&= -\int_0^t c(s, Y(s))\,dB(s) - \frac{1}{2}\int_0^t |c(s, Y(s))|^2\,ds\\
&= -\int_0^t c(s, Y(s))\,dB'(s) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\,ds
\end{aligned}$$

$(c(s,y) = c_1(s,y)\sigma(s,y))$

$$\begin{aligned}
&= -\int_0^t c_1(s, Y(s))\,\sigma(s, Y(s))\,dB'(s) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\,ds\\
&= -\int_0^t c_1(s, Y(s))\,d\left(\int_0^s \sigma(\tau, Y(\tau))\,dB'(\tau)\right) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\,ds\\
&= -\int_0^t c_1(s, Y(s))\,d\left(Y(s) - \int_0^s b(\tau, Y(\tau))\,d\tau\right) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\,ds.
\end{aligned} \tag{1.43}$$

From (1.43), (1.38), and (1.40) it is plain that $M'(t)$ only depends on the path $\{Y(s) : 0 \leqslant s \leqslant t\}$.

(3). This assertion is a consequence of Kolmogorov's extension theorem. The measure is $\mathbb{P}'$ is well defined on $\bigcup_{t>0} \mathcal{F}_t$. Here we use the martingale property. By Kolmogorov's extension theorem, it extends to the σ-field generated by this union.

(4). The equality $B'(t) = B(t) + \int_0^t c\,(s, Y(s))\, ds$ entails the following equality for the quadratic covariation of the processes B'_i and B'_j:

$$\left\langle B'_i, B'_j \right\rangle (t) = \langle B_i, B_j \rangle (t) = t\delta_{i,j}. \tag{1.44}$$

From Itô calculus we also infer

$$\begin{aligned} & M'(t)B'_i(t) \\ &= \int_0^t M'(s)B'_i(s)\, dZ'(s) + \int_0^t M'(s)\, dB'_i(s) \\ &\quad + \frac{1}{2}\int_0^t M'(s)B'(s)\, d\left\langle Z', Z'\right\rangle(s) + \int_0^t M'(s) d\left\langle Z', B'_i\right\rangle(s) \\ &= -\int_0^t M'(s)B'_i(s)c\,(s, Y(s))\; dB(s) - \frac{1}{2}\int_0^t M'(s)B'_i(s) \left|c\,(s, Y(s))\right|^2 ds \\ &\quad + \frac{1}{2}\int_0^t M'(s)B'_i(s) \left|c\,(s, Y(s))\right|^2 ds + \int_0^t M'(s)\, dB_i(s) \\ &\quad + \int_0^t M'(s)c_i\,(s, Y(s))\; ds - \int_0^t M'(s)c_i\,(s, Y(s))\; ds \\ &= -\int_0^t M'(s)B'_i(s)c\,(s, Y(s))\; dB(s) + \int_0^t M'(s)\, dB_i(s). \end{aligned} \tag{1.45}$$

Upon invoking Theorem 1.1 and employing (1.44) and (1.45) Assertion (4) follows.

This concludes the proof of Proposition 1.1. □

Let the process $X(t)$ solve the equation in (1.36), and put

$$M(t) = \exp\left(\int_0^t c\,(s, X(s))\, dB(s) - \frac{1}{2}\int_0^t \left|c\,(s, X(s))\right|^2\, ds\right), \tag{1.46}$$

and assume that the process $M(t)$ is not merely a local martingale, but a genuine $\mathbb{P}$-martingale.

Theorem 1.4. *Fix $T > 0$, and let the functions*

$$b(s,y), \quad \sigma(s,y), \quad c(s,y), \quad \text{and} \quad c_1(s,y), \quad 0 \leqslant s \leqslant T,$$

be locally bounded Borel measurable vector or matrix functions such that $c(s,y) = c_1(s,y)\sigma\,(s,y)$, $0 \leqslant s \leqslant T$, $y \in \mathbb{R}^d$. Suppose that the equation in (1.36) possesses unique weak solutions on the interval $[0,T]$.

Uniqueness. If weak solutions to the stochastic differential equation in (1.37) exist, then they are unique in the sense as explained next. In fact, let the couple $(Y(s), B(s))$, $0 \leqslant s \leqslant t$, *be a solution to the equation in (1.37) with the property that the local martingale* $M'(t)$ *given by*

$$M'(t) = \exp\left(-\int_0^t c\,(s, Y(s))\, dB(s) - \frac{1}{2}\int_0^t |c\,(s, Y(s))|^2\, ds\right)$$

satisfies $\mathbb{E}\left[M'(t)\right] = 1$. *Then the finite-dimensional distributions of the process* $Y(s)$, $0 \leqslant s \leqslant t$, *are given by the Girsanov or Cameron-Martin transform:*

$$\mathbb{E}\left[f\left(Y\left(t_1\right), \dots, Y\left(t_n\right)\right)\right] = \mathbb{E}\left[M'(t) f\left(X\left(t_1\right), \dots, X\left(t_n\right)\right)\right], \tag{1.47}$$

$t \geqslant t_n > \cdots > t_1 \geqslant 0$, *where* $f : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \to \mathbb{R}$ *is an arbitrary bounded Borel measurable function.*

Existence. Conversely, let the process $s \mapsto (X(s), B(s))$ *be a solution to the equation in (1.36). Suppose that the local martingale* $s \mapsto M(s)$, *defined by*

$$M(s) = \exp\left(\int_0^s c\,(\tau, X(\tau))\, dB(\tau) - \frac{1}{2}\int_0^s |c\,(\tau, X(\tau))|^2\, d\tau\right), \quad 0 \leqslant s \leqslant t,$$

is a martingale, i.e. $\mathbb{E}\left[M(t)\right] = 1$. *Then there exists a couple* $\left(\widetilde{Y}(s), \widetilde{B}(s)\right)$, $0 \leqslant s \leqslant t$, *where* $s \mapsto \widetilde{B}(s)$, $0 \leqslant s \leqslant t$, *is a Brownian motion on a probability space* $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *such that*

$$\begin{aligned}\widetilde{Y}(s) = x &+ \int_0^s \sigma\left(\tau, \widetilde{Y}(s)\right)\, d\widetilde{B}(s) \\ &+ \int_0^s \sigma\left(\tau, \widetilde{Y}(s)\right) c\left(\tau, \widetilde{Y}(\tau)\right)\, d\tau + \int_0^s b\left(\tau, \widetilde{Y}(s)\right)\, d\tau,\end{aligned} \tag{1.48}$$

and such that

$$\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s, \widetilde{Y}(s)\right)\, d\widetilde{B}(s) - \frac{1}{2}\int_0^t \left|c\left(s, \widetilde{Y}(s)\right)\right|^2 ds\right)\right] = 1. \tag{1.49}$$

Remark 1.3. The formula in (1.47) is known as the Girsanov transform or Cameron-Martin transform of the measure $\mathbb{P}$. It is a martingale measure. Suppose that the process $t \mapsto M'(t)$, as defined in (1.38) is a $\mathbb{P}$-martingale. Then the proof of Theorem 1.4 shows that the process $t \mapsto M(t)$, as defined in (1.46) is a $\mathbb{P}$-martingale. By Assertion (1) in Proposition 1.1 the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale if and only $\mathbb{E}\left[M'(t)\right] = 1$ for all $T \geqslant t \geqslant 0$, and a similar statement holds for the process $t \mapsto M(t)$. If the process $t \mapsto M'(t)$ is a martingale, then taking $G \equiv \mathbf{1}$ in (1.65) shows that $\mathbb{E}[M(t)] = 1$, and hence by 1 in Proposition 1.1 the process $t \mapsto M(t)$ is a $\mathbb{P}$-martingale. Conversely, if the process $t \mapsto M(t)$ is a $\mathbb{P}$-martingale, then we reverse the implications in the proof of Theorem 1.4 and take $F \equiv \mathbf{1}$ in (1.64) to conclude that $\mathbb{E}[M'(t)] = 1$ for all $\geqslant 0$. But then the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale.

Notice that the process $t \mapsto M(t)$ is a $\mathbb{P}$-martingale provided Novikov's condition is satisfied, i.e. if $\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t |c(s, X(s))|^2\, ds\right)\right] < \infty$. For a precise formulation see Corollary 1.3 below. Novikov's result is a consequence of Theorem 1.6. For a closely related Novikov condition on an exponential (local) martingale see item (5) in the beginning of §1.3.

Remark 1.4. Let $s \mapsto c(s)$ be a process which is adapted to a Brownian motion $(B(t))_{t\geqslant 0}$ starting at 0 in $\mathbb{R}^d$, and let $\rho > 0$ be such that Novikov's condition is satisfied: $\mathbb{E}\left[\exp\left(\frac{1}{2}\rho^2\int_0^t |c(s)|^2\, ds\right)\right] < \infty$. From Assertion (4) in Proposition 1.1 and Theorem 1.4 we see that the following identity holds for all bounded Borel measurable functions F defined on $\left(\mathbb{R}^d\right)^n$:

$$\begin{aligned}
&\mathbb{E}\left[F\left(Y_\rho(t_1), \dots, Y_\rho(t_n)\right)\right] \\
&= \mathbb{E}\left[\exp\left(\rho\int_0^t c(s)dB(s) - \frac{1}{2}\rho^2\int_0^t |c(s)|^2\, ds\right) F\left(B(t_1), \dots, B(t_n)\right)\right]
\end{aligned} \tag{1.50}$$

where $0 \leqslant t_1 < \cdots < t_n \leqslant t$, and $Y_\rho(\tau) = B(\tau) + \rho\int_0^\tau c(s)\, ds$, $0 \leqslant \tau \leqslant t$. In particular, if $n = 1$ we get

$$\begin{aligned}
&\mathbb{E}\left[F\left(B(t) + \rho\int_0^t c(s)\, ds\right)\right] \\
&\quad = \mathbb{E}\left[\exp\left(\rho\int_0^t c(s)dB(s) - \frac{1}{2}\rho^2\int_0^t |c(s)|^2\, ds\right) F\left(B(t)\right)\right].
\end{aligned} \tag{1.51}$$

Assume that the gradient DF of the function F exists and is bounded. The equality in (1.51) can be differentiated with respect to ρ to obtain:

$$\begin{aligned}
&\mathbb{E}\left[\left\langle DF\left(B(t) + \rho\int_0^t c(s)\, ds\right), \int_0^t c(s)\, ds\right\rangle\right] \\
&= \mathbb{E}\left[\exp\left(\rho\int_0^t c(s)\, dB(s) - \frac{1}{2}\rho^2\int_0^t |c(s)|^2\, ds\right)\right. \\
&\qquad \left.\times\left(\int_0^t c(s)\, dB(s) - \rho\int_0^t |c(s)|^2\, ds\right) F\left(B(t)\right)\right].
\end{aligned} \tag{1.52}$$

The bracket in the left-hand side of (1.52) indicates the inner-product in $\mathbb{R}^d$. In (1.52) we put $\rho = 0$ and we obtain the first order version of the famous integration by parts formula:

$$\mathbb{E}\left[\left\langle DF\left(B(t)\right), \int_0^t c(s)\, ds\right\rangle\right] = \mathbb{E}\left[\int_0^t c(s)\, dB(s)\, F\left(B(t)\right)\right]. \qquad (1.53)$$

We mention that the Cameron-Martin-Girsanov transformation is a cornerstone for integration by parts formulas with higher derivatives than in (1.53), which is a central issue in Malliavin calculus, also called stochastic variation calculus. For details on this subject see e.g. [Nualart (1998, 2006)], [Malliavin (1978)], [Sanz-Solé (2005)], [Kusuoka and Stroock (1985, 1987, 1984)], [Stroock (1981)], and [Norris (1986)].

For a proof of Theorem 1.4 we will need the Skorohod-Dudley-Wichura representation theorem: see Theorem 11.7.2 in [Dudley (2002)]. It will be applied with $S = C\left([0,t], \mathbb{R}^d\right)$ and can be formulated as follows.

Theorem 1.5. *Let (S,d) be a complete separable metric space (i.e. a Polish space), and let $\mathbb{P}_k$, $k \in \mathbb{N}$, and $\mathbb{P}$ be probability measures on the Borel field $\mathcal{B}_S$ of S such that the weak limit $\text{w-}\lim_{k\to\infty} \mathbb{P}_k = \mathbb{P}$, i.e. $\lim_{k\to\infty} \int F d\mathbb{P}_k = \int F d\mathbb{P}$ for all bounded continuous functions of $F \in C_b(S)$. Then there exist a probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ and S-valued random variables $\widetilde{Y}_k$, $k \in \mathbb{N}$, and $\widetilde{Y}$, defined on $\widetilde{\Omega}$ with the following properties:*

(1) $\mathbb{P}_k[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y}_k \in B\right]$, $k \in \mathbb{N}$, and $\mathbb{P}[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y} \in B\right]$, $B \in \mathcal{B}_S$.
(2) The sequence $\widetilde{Y}_k$, $k \in \mathbb{N}$, converges to $\widetilde{Y}$ $\widetilde{\mathbb{P}}$-almost surely.

Remark 1.5. An analysis of the existence part of the proof of Theorem 1.4 shows that the invertibility of the matrix $\sigma(s,y)$ is not needed. Let $\widetilde{N}(s)$, $0 \leqslant s \leqslant t$, be a local martingale on a filtered probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}_s, \widetilde{\mathbb{P}}\right)$, where the σ-field $\widetilde{\mathcal{F}}_s$ is generated by $\left(\widetilde{Y}(\tau) : 0 \leqslant \tau \leqslant s\right)$. Suppose that the covariation process of $\widetilde{N}(s)$ is given by

$$\langle N_{j_1}, N_{j_2}\rangle(s) = \int_0^s \left(\sigma\left(\tau, \widetilde{Y}(\tau)\right)\sigma^*\left(\tau, \widetilde{Y}(\tau)\right)\right)_{j_1,j_2} d\tau, \quad 1 \leqslant j_1,\ j_2 \leqslant d.$$

Here $\widetilde{Y}$ is a local martingale on $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$. Then by assertion (iii) in Theorem 1.3 there exists a Brownian motion $\widetilde{B}(s)$, $0 \leqslant s \leqslant t$, on this space such that

$$\int_0^s c_1\left(\tau,\widetilde{Y}(\tau)\right)\, d\widetilde{N}(\tau) = \int_0^s c_1\left(\tau,\widetilde{Y}(\tau)\right)\sigma\left(\tau,\widetilde{Y}(\tau)\right)\, d\widetilde{B}(\tau)$$
$$= \int_0^s c\left(\tau,\widetilde{Y}(\tau)\right)\, d\widetilde{B}(\tau). \tag{1.54}$$

Proof. [Proof of Theorem 1.4.] *Uniqueness.* Let the process $Y(s)$, $0 \leqslant s \leqslant t$, be a solution to equation (1.37). So that

$$Y(s) = x + \int_0^s \sigma(\tau, Y(\tau))\, dB(\tau)$$
$$+ \int_0^s \left(b(\tau, Y(\tau)) + \sigma(\tau, Y(\tau))\, c(\tau, Y(\tau))\right) d\tau$$
$$= x + \int_0^s \sigma(\tau, Y(\tau))\, dB'(\tau) + \int_0^s b(\tau, Y(\tau))\, d\tau. \tag{1.55}$$

Let $F\left((Y(s))_{0\leqslant s\leqslant t}\right)$ be a bounded random variable which depends on the path $Y(s)$, $0 \leqslant s \leqslant t$. As observed in 4 of Proposition 1.1 the process $B'(t)$ is a $\mathbb{P}'$-Brownian motion, provided $\mathbb{E}\left[M'(t)\right] = 1$. Uniqueness of weak solutions to equation (1.36) implies that the P'-distribution of the process $s \mapsto Y(s)$, $0 \leqslant s \leqslant t$, coincides with the $\mathbb{P}$-distribution of the process $s \mapsto X(s)$, $0 \leqslant s \leqslant t$. In other words we have

$$\mathbb{E}'\left[F\left((Y(s))_{0\leqslant s\leqslant t}\right)\right]$$
$$= \mathbb{E}\left[\exp\left(-\int_0^t c_1(s, Y(s))\, dN^Y(s) - \frac{1}{2}\left|c(s, Y(s))\right|^2 ds\right) F\left((Y(s))_{0\leqslant s\leqslant t}\right)\right]$$
$$= \mathbb{E}\left[F\left((X(s))_{0\leqslant s\leqslant t}\right)\right], \tag{1.56}$$

where

$$N^Y(s) = Y(s) - \int_0^s \sigma(\tau, Y(\tau))\, c(\tau, Y(\tau))\, d\tau - \int_0^s b(\tau, Y(\tau))\, d\tau$$
$$= \int_0^s \sigma(\tau, Y(\tau))\, dB(\tau). \tag{1.57}$$

With

$$G\left((Y(s))_{0\leqslant s\leqslant t}\right)$$
$$= \exp\left(-\int_0^t c_1(s, Y(s))\, dN^Y(s) - \frac{1}{2}\left|c(s, Y(s))\right|^2 ds\right) F\left((Y(s))_{0\leqslant s\leqslant t}\right)$$

we have

$$F\left((Y(s))_{0\leqslant s\leqslant t}\right)$$

$$= \exp\left(\int_0^t c_1\left(s, Y(s)\right)\, dN^Y(s) + \frac{1}{2}\left|c\left(s, Y(s)\right)\right|^2 ds\right) G\left(\left(Y(s)\right)_{0\leqslant s\leqslant t}\right).$$

So, since

$$\begin{aligned} dN^X(s) &= dX(s) - \sigma\left(s, X(s)\right) c\left(s, X(s)\right)\, ds - b\left(s, X(s)\right)\, ds \\ &= \sigma\left(s, X(s)\right)\, \left(dB(s) - c\left(s, X(s)\right)\, ds\right) \end{aligned} \tag{1.58}$$

it follows that

$$\begin{aligned} &F\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right) \\ &= \exp\left(\int_0^t c_1\left(s, X(s)\right)\, dN^X(s) + \frac{1}{2}\left|c\left(s, X(s)\right)\right|^2 ds\right) G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right) \\ &= \exp\left(\int_0^t c\left(s, X(s)\right)\, dB(s) - \frac{1}{2}\left|c\left(s, X(s)\right)\right|^2 ds\right) G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right). \end{aligned} \tag{1.59}$$

From (1.56) and (1.59) we infer:

$$\begin{aligned} &\mathbb{E}'\left[G\left(\left(Y(s)\right)_{0\leqslant s\leqslant t}\right)\right] \\ &= \mathbb{E}\left[\exp\left(\int_0^t c\left(s, X(s)\right)\, ds - \frac{1}{2}\int_0^t \left|c\left(s, X(s)\right)\right|^2 ds\right) G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right]. \end{aligned} \tag{1.60}$$

By inserting $G \equiv \mathbf{1}$ in (1.60) we see that

$$\mathbb{E}\left[\exp\left(\int_0^t c\left(s, X(s)\right)\, ds - \frac{1}{2}\int_0^t \left|c\left(s, X(s)\right)\right|^2 ds\right)\right] = 1$$

in case there is a unique solution to the equation in (1.48). This proves the uniqueness part of Theorem 1.4.

Existence. Therefore we will approximate the solution Y by a sequence Y_k, $k \in \mathbb{N}$, which are solutions to equations of the form:

$$\begin{aligned} Y_k(s) = x &+ \int_0^s \sigma\left(\tau, Y_k(\tau)\right)\, dB(\tau) \\ &+ \int_0^s \left(b\left(\tau, Y_k(\tau)\right) + \sigma\left(\tau, Y_k(\tau)\right) c_k\left(\tau, Y_k(\tau)\right)\right)\, d\tau \\ = x &+ \int_0^s \sigma\left(\tau, Y_k(\tau)\right) dB_k'(\tau) + \int_0^s b\left(\tau, Y_k(\tau)\right) d\tau. \end{aligned} \tag{1.61}$$

Here $B_k'(s) = B_k(s) + \int_0^t c_k\left(\tau, Y_k(\tau)\right)\, d\tau$, and the coefficients $c_k(s,y) = c_{1,k}(s,y)\sigma\left(s,y\right)$ are chosen in such a way that they are bounded and that

$c(s,y) = \lim_{k\to\infty} c_k(s,y)$ for all $s \in [0,t]$ and $y \in \mathbb{R}^d$. By Novikov's theorem the corresponding local martingales M_k', given by

$$M_k'(s) = \exp\left(-\int_0^s c_k(\tau, Y_k(\tau))\, dB(\tau) - \frac{1}{2}\int_0^s |c_k(\tau, Y_k(\tau))|^2\, d\tau\right), \quad k \in \mathbb{N},$$

are then automatically genuine martingales: see Corollary 1.3. From the uniqueness of weak solutions to equations in $X(t)$ of the form (1.36) (and thus to equations in $Y_k(s)$ of the form (1.61) we infer

$$\mathbb{E}_k'\left[F\left((Y_k(s))_{0\leqslant s\leqslant t}\right)\right] = \mathbb{E}\left[F\left((X(s))_{0\leqslant s\leqslant t}\right)\right]. \tag{1.62}$$

In equality (1.62) the process $Y_k(s)$, $0 \leqslant s \leqslant t$, solves the equation in (1.61). The equality in (1.62) can be rewritten as

$$\mathbb{E}\left[M_k'(t)F\left((Y_k(s))_{0\leqslant s\leqslant t}\right)\right] = \mathbb{E}\left[F\left((X(s))_{0\leqslant s\leqslant t}\right)\right]. \tag{1.63}$$

By (1.43) the equality in (1.63) can be rewritten as

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(-\int_0^t c_k(s, Y_k(s))\, dB(s) - \frac{1}{2}\int_0^t |c_k(s, Y_k(s))|^2\, ds\right) F\left((Y_k(s))_{0\leqslant s\leqslant t}\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\int_0^t c_{1,k}(s, Y_k(s))\ d\left(Y_k(s) - \int_0^s b(\tau, Y_k(\tau))\ d\tau\right)\right.\right. \\
&\qquad \left.\left. + \frac{1}{2}\int_0^t |c_k(s, Y_k(s))|^2\ ds\right) F\left((Y_k(s))_{0\leqslant s\leqslant t}\right)\right] \\
&= \mathbb{E}\left[F\left((X(s))_{0\leqslant s\leqslant t}\right)\right].
\end{aligned} \tag{1.64}$$

Let $G\left((Y_k(s))_{0\leqslant s\leqslant t}\right)$ be a (bounded) random variable which depends on the path $Y_k(s)$, $0 \leqslant s \leqslant t$. From the equality in (1.64) we infer

$$\begin{aligned}
&\mathbb{E}\left[G\left((Y_k(s))_{0\leqslant s\leqslant t}\right)\right] \\
&= \mathbb{E}\left[\exp\left(\int_0^t c_{1,k}(s, X(s))\ d\left(X(s) - \int_0^s b(\tau, X(\tau))\ d\tau\right)\right.\right. \\
&\qquad \left.\left. - \frac{1}{2}\int_0^t |c_k(s, X(s))|^2\ ds\right) G\left((X(s))_{0\leqslant s\leqslant t}\right)\right] \\
&= \mathbb{E}\left[\exp\left(\int_0^t c_{1,k}(s, X(s))\, \sigma(s, X(s))\ dB(s) - \frac{1}{2}\int_0^t |c_k(s, X(s))|^2\ ds\right)\right. \\
&\qquad \left. G\left((X(s))_{0\leqslant s\leqslant t}\right)\right] \\
&= \mathbb{E}\left[M_k(t) G\left((X(s))_{0\leqslant s\leqslant t}\right)\right].
\end{aligned} \tag{1.65}$$

Here the martingales $M_k(s)$ are given by

$$M_k(s) = \exp\left(\int_0^s c_k(\tau, X(\tau))\, dB(\tau) - \frac{1}{2}\int_0^s |c_k(\tau, X(\tau))|^2\, d\tau\right), \quad k \in \mathbb{N}.$$

This fact together with the pointwise convergence of $M_k(s)$ to $M(s)$, as $k \to \infty$, and invoking the hypothesis that $\mathbb{E}[M(t)] = 1$, shows that the right-hand side of (1.65) converges to $\mathbb{E}\left[M(t)G\left((X(s))_{0\leqslant s\leqslant t}\right)\right]$. In other words the distribution $\mathbb{P}^{Y_k}$ of Y_k converges weakly to the measure $\mathbb{P}^{M,X}$ defined by $\mathbb{P}^{M,X}(A) = \mathbb{E}[M(t), X \in A]$, where A is a Borel subset of the space $C\left([0,t],\mathbb{R}^d\right)$. By the Skorohod-Dudley-Wichura representation theorem (Theorem 1.5) there exist a probability space $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}}\right)$ and $C\left([0,t],\mathbb{R}^d\right)$-valued random variables $\widetilde{Y}_k$, $k \in \mathbb{N}$, and $\widetilde{Y}$, defined on $\widetilde{\Omega}$ with the following properties:

(1) $\mathbb{P}^{Y_k}[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y}_k \in B\right]$, $k \in \mathbb{N}$, and $\mathbb{P}^{M,X}[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y} \in B\right]$, $B \in \mathcal{B}_{C([0,t],\mathbb{R}^d)}$.
(2) The sequence $\widetilde{Y}_k$, $k \in \mathbb{N}$, converges to $\widetilde{Y}$ $\widetilde{\mathbb{P}}$-almost surely.

By taking the limit in (1.65) for $k \to \infty$ and using the theorem of Skorohod-Dudley-Wichura we obtain

$$\mathbb{E}\left[G\left(\left(\widetilde{Y}(s)\right)_{0\leqslant s\leqslant t}\right)\right] = \mathbb{E}\left[M(t)G\left((X(s))_{0\leqslant s\leqslant t}\right)\right] \tag{1.66}$$

where G is a bounded continuous function on $C\left([0,t],\mathbb{R}^d\right)$. Then we consider the process $\widetilde{N}(s)$, $0 \leqslant s \leqslant t$, defined by

$$\widetilde{N}(s) = \widetilde{Y}(s) - \int_0^s \sigma\left(\tau,\widetilde{Y}(\tau)\right) c\left(\tau,\widetilde{Y}(\tau)\right) d\tau - \int_0^s b\left(\tau,\widetilde{Y}(\tau)\right) d\tau. \tag{1.67}$$

If $\widetilde{Y}(s)$ were $Y(s)$, then by (1.55) $\widetilde{N}(s)$ would be $N^Y(s)$, given by the formula in (1.57). Hence the process $s \mapsto N^Y(s)$, $s \in [0,t]$, is a stochastic integral relative to Brownian motion on the space $(\Omega,\mathcal{F}_t,\mathbb{P})$. We want to do same for the process $s \mapsto \widetilde{N}(s)$, $0 \leqslant s \leqslant t$, on the probability space $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}}\right)$. Let $\mathbb{P}^{M(t)}$ be the probability measure on $(\Omega,\mathcal{F}_t)$ defined by $\mathbb{P}^{M(t)}[A] = \mathbb{E}[M(t), A]$, $A \in \mathcal{F}_t$. Then like in item (4) of Proposition 1.1 we see that the process $s \mapsto B(s) - \int_0^s \sigma(\tau, X(\tau))\, d\tau$ is a $\mathbb{P}^{M(t)}$-Brownian motion. In addition, from (1.66) and (1.67) we infer that the $\widetilde{\mathbb{P}}$-distribution of the process $\widetilde{N}(s)$, $0 \leqslant s \leqslant t$, is given by the $\mathbb{P}^{M(t)}$-distribution of the process

$$\begin{aligned} s \mapsto & X(s) - \int_0^s \sigma(\tau, X(\tau))\, c(\tau, X(\tau))\, d\tau - \int_0^s b(\tau, X(\tau))\, d\tau \\ &= \int_0^s \sigma(\tau, X(\tau))\, (dB(\tau) - c(\tau, Y(\tau))\, d\tau) \end{aligned}$$

$$= \int_0^s \sigma\left(\tau, X(\tau)\right)\, dB^{M(t)}(\tau), \tag{1.68}$$

where $B^{M(t)}(s)$ is a $\mathbb{P}^{M(t)}$-Brownian motion: see Proposition 1.1 item (4). It also follows that the process in (1.68) has covariation process given by the square matrix process

$$s \mapsto \int_0^s \sigma\left(\tau, X(\tau)\right) \sigma^*\left(\tau, X(\tau)\right)\, d\tau, \quad 0 \leqslant s \leqslant t.$$

Consequently, the process $s \mapsto \widetilde{N}(s)$, $0 \leqslant s \leqslant t$, is a local $\widetilde{P}$-martingale with covariation process given by

$$s \mapsto \int_0^s \sigma\left(\tau, \widetilde{Y}(\tau)\right) \sigma^*\left(\tau, \widetilde{Y}(\tau)\right)\, d\tau, \quad 0 \leqslant s \leqslant t. \tag{1.69}$$

In order to prove (1.69) we must show that the process

$$s \mapsto \widetilde{N}_{j_1}(s)\widetilde{N}_{j_2}(s) - \sum_{k=1}^d \int_0^s \sigma_{j_1,k}\left(\tau, \widetilde{Y}(\tau)\right) \sigma_{j_2,k}\left(\tau, \widetilde{Y}(\tau)\right)\, d\tau$$

is a local $\widetilde{\mathbb{P}}$-martingale. The latter can be achieved by appealing to the fact that the $\widetilde{\mathbb{P}}$-distribution of the process $s \mapsto \widetilde{Y}(s)$, $0 \leqslant s \leqslant t$, coincides with the $\mathbb{P}^{M(t)}$-distribution of the process $s \mapsto X(s)$, $0 \leqslant s \leqslant t$. Then we choose a Brownian motion $\widetilde{B}(s)$, possibly on an extension of the probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$, which we call again $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ such that $\widetilde{N}(s) = \int_0^s \sigma\left(\tau, \widetilde{Y}(\tau)\right)\, d\widetilde{B}(\tau)$. For details see the proof of the implication (ii) $\Longrightarrow$ (iii) of Theorem 1.3. With such a Brownian motion we obtain:

$$\begin{aligned}\widetilde{Y}(s) = x &+ \int_0^s \sigma\left(\tau, \widetilde{Y}(\tau)\right)\, d\widetilde{B}(\tau) + \int_0^s \sigma\left(\tau, \widetilde{Y}(\tau)\right) c\left(\tau, \widetilde{Y}(\tau)\right)\, d\tau \\ &+ \int_0^s b\left(\tau, \widetilde{Y}(\tau)\right)\, d\tau.\end{aligned} \tag{1.70}$$

Since

$$\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s, \widetilde{Y}(s)\right)\, d\widetilde{B}(s) - \frac{1}{2}\int_0^t \left|c\left(s, \widetilde{Y}(s)\right)\right|^2 ds\right)\right] = 1 \tag{1.71}$$

it follows that the process $s \mapsto \widetilde{B}(s) + \int_0^s c\left(\tau, \widetilde{Y}(\tau)\right)\, d\tau$ is a Brownian motion relative to the measure

$$A \mapsto \widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s, \widetilde{Y}(s)\right)\, d\widetilde{B}(s) - \frac{1}{2}\int_0^t \left|c\left(s, \widetilde{Y}(s)\right)\right|^2 ds\right), A\right],$$

$A \in \widetilde{\mathcal{F}}$. The equalities in (1.70) and (1.71) complete the proof of Theorem 1.4. □

We include a proof of a result due to [Novikov (1973)]. In fact we will insert a proof established by Krylov [Krylov (2002)]. In fact the result is somewhat more general than the original result by Novikov; it also improves a result which we owe to [Kazamaki (1978)]. For the significance of the covariation process $t \mapsto \langle M, M\rangle (t)$ see items (4) and (5) in §1.3.

Theorem 1.6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and let $t \mapsto M(t)$ be a continuous local martingale on $(\Omega, \mathcal{F}, \mathbb{P})$ relative to a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ such that $\langle M, M\rangle = \langle M, M\rangle(\infty) := \sup_{t\geqslant 0} \langle M, M\rangle (t) < \infty$ ($\mathbb{P}$-almost surely). Define*

$$\mathcal{E}(M)(t) = e^{M(t)-\frac{1}{2}\langle M,M\rangle(t)}. \tag{1.72}$$

Then the following assertions are true:

(1) If $\liminf_{\varepsilon\downarrow 0} \varepsilon \log \mathbb{E}\left[e^{\frac{1}{2}(1-\varepsilon)\langle M,M\rangle(\infty)}\right] < \infty$, then

$$\mathbb{E}\left[\exp(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty))\right] = 1. \tag{1.73}$$

Consequently, the process $t \mapsto \mathcal{E}(M)(t)$ is a $\mathbb{P}$-martingale relative to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, where $\mathcal{F}_t = \sigma(M(s) : 0 \leqslant s \leqslant t)$, the σ-field generated by the variables $M(s)$, $0 \leqslant s \leqslant t$.

(2) If $\liminf_{\varepsilon\downarrow 0} \varepsilon \log \sup_{t\geqslant 0} \mathbb{E}\left[e^{\frac{1}{2}(1-\varepsilon)M(t)}\right] < \infty$, then again the equality in (1.73) holds. So that the process $t \mapsto \mathcal{E}(M)(t)$ is a $\mathbb{P}$-martingale relative to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ determined by the local martingale $t \mapsto M(t)$.

We mention the following corollaries. Corollary 1.3 is due to [Novikov (1973)]. Corollary 1.4 is a result by [Kazamaki (1978)]. In the corollaries 1.3 and 1.4, and in the lemmas 1.2 and 1.3 it is assumed that the process $t \mapsto M(t)$ is a continuous local martingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, the notation is as in Theorem 1.6.

Corollary 1.3. *If $\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M, M\rangle(\infty)\right)\right] < \infty$, then*

$$\mathbb{E}\left[\exp\left(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty)\right)\right] = 1,$$

and consequently the process $t \mapsto \mathcal{E}(M)(t)$ is a $\mathbb{P}$-martingale relative to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, where $\mathcal{F}_t = \sigma(M(s) : 0 \leqslant s \leqslant t)$, the σ-field generated by the variables $M(s)$, $0 \leqslant s \leqslant t$.

Proof. If $\mathbb{E}[\mathcal{E}(M)(\infty)] = 1$, then $\mathcal{E}(M)(t) = \mathbb{E}\left[\mathcal{E}(M)(\infty) \mid \mathcal{F}_t\right]$, and hence the process $t \mapsto \mathcal{E}(M)(t)$ is a $\mathbb{P}$-martingale. □

The same argument shows the following corollary.

Corollary 1.4. *If* $\sup_{t\geqslant 0} \mathbb{E}\left[\exp\left(\frac{1}{2}M(t)\right)\right] < \infty$, *then*

$$\mathbb{E}\left[\exp\left(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty)\right)\right] = 1.$$

Again the process $t \mapsto \mathcal{E}(M)(t)$ *is not just a local martingale, but a* $\mathbb{P}$-*martingale.*

The proof of Theorem 1.6 is based on Lemma 1.1 below. Doob's martingale inequality for moments, which is also needed, reads as follows. Let $k \mapsto Y_k$ be a discrete martingale on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $\delta > 0$, then

$$\mathbb{E}\left[\max_{1\leqslant k\leqslant n} |Y_k|^{1+\delta}\right] \leqslant \left(\frac{1+\delta}{\delta}\right)^{1+\delta} \mathbb{E}\left[|Y_n|^{1+\delta}\right], \quad n \in \mathbb{N}. \tag{1.74}$$

For details see e.g. [Cox (1984)]. If $\delta = 0$, then the inequality in (1.74) should be replaced with:

$$\mathbb{E}\left[\sup_{1\leqslant k\leqslant n} |Y_k|\right] \leqslant \frac{e}{e-1}\left(1 + \mathbb{E}\left[|Y_n| \log^+ |Y_n|\right]\right). \tag{1.75}$$

Similar inequalities hold for right-continuous local submartingales. In particular we have

$$\mathbb{E}\left[\sup_{0\leqslant s\leqslant t} \mathcal{E}(N)(s)^{1+\delta}\right] \leqslant \left(\frac{1+\delta}{\delta}\right)^{1+\delta} \mathbb{E}\left[\mathcal{E}(N)(t)^{1+\delta}\right] \tag{1.76}$$

provided that the process $t \mapsto N(t)$ is a continuous local martingale. The inequality in (1.76) follows from (1.74) by taking a discretization of the form $j \mapsto N\left(j2^{-n}t\right)$, $1 \leqslant j \leqslant 2^n$, and then letting n tend to ∞. In addition, in general a stopping time argument (or localization argument) is required. In such a case we replace $N(t)$ by $N\left(\min\left(t, \tau_m\right)\right)$, where $\tau_m = \inf\{t > 0 : |N(t)| > m\}$. Then first we let n tend to ∞, and then m.

Lemma 1.1. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, and let* $t \mapsto N(t)$ *be a continuous local martingale for which there exists* $\varepsilon_0 > 0$ *such that*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)^2 \langle N, N\rangle(\infty)\right)\right] < \infty. \tag{1.77}$$

Then

$$\sup_{t\geqslant 0} \mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right) N(t)\right)\right] \leqslant \mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)^2 \langle N, N\rangle(\infty)\right)\right] < \infty, \tag{1.78}$$

and $\mathbb{E}\left[\mathcal{E}(N)(\infty)\right] = 1$.

Proof. The inequality in (1.78) follows from the Cauchy-Schwarz inequality. In fact we write

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)\right)\right] \\
&= \mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)-\frac{1}{4}\left(1+\varepsilon_0\right)^2\left\langle N,N\right\rangle(t)\right)\right. \\
&\qquad \left.\times\exp\left(\frac{1}{4}\left(1+\varepsilon_0\right)^2\left\langle N,N\right\rangle(t)\right)\right] \\
&\leqslant \mathbb{E}\left[\exp\left(\left(1+\varepsilon_0\right)N(t)-\frac{1}{2}\left(1+\varepsilon_0\right)^2\left\langle N,N\right\rangle(t)\right)\right]^{1/2} \\
&\qquad \times\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)^2\left\langle N,N\right\rangle(t)\right)\right]^{1/2} \\
&= \mathbb{E}\left[\mathcal{E}\left(\left(1+\varepsilon_0\right)N\right)(t)\right]^{1/2}\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)^2\left\langle N,N\right\rangle(t)\right)\right]^{1/2}. \qquad (1.79)
\end{aligned}$$

The process $t\mapsto N(t)$ is a continuous local martingale, and so is the process $t\mapsto\left(1+\varepsilon_0\right)N(t)$. A stopping time argument, which in fact is a localization technique, then shows that

$$\mathbb{E}\left[\mathcal{E}\left(\left(1+\varepsilon_0\right)N\right)(t)\right]\leqslant 1. \qquad (1.80)$$

A combination of (1.79), (1.80) and (1.77) then shows

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)\right)\right]\leqslant\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)^2\left\langle N,N\right\rangle(\infty)\right)\right]^{1/2}<\infty. \qquad (1.81)$$

For brevity we write

$$\delta=\frac{\varepsilon_0^2}{1+2\varepsilon_0},\quad \gamma=\frac{1}{1+\varepsilon_0},\quad p=1+2\varepsilon_0,\quad q=\frac{1+2\varepsilon_0}{2\varepsilon_0}. \qquad (1.82)$$

Notice that $\frac{1}{p}+\frac{1}{q}=1$. Then with the notation of (1.82) we have by (1.76) the following estimates:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\mathcal{E}(N)(s)\right]^{1+\delta}\leqslant\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\mathcal{E}(N)(s)^{1+\delta}\right] \\
&\leqslant\left(\frac{1+\delta}{\delta}\right)^{1+\delta}\mathbb{E}\left[\mathcal{E}(N)(t)^{1+\delta}\right] \\
&=\left(\frac{1+\delta}{\delta}\right)^{1+\delta}\mathbb{E}\left[\exp\left(\gamma(1+\delta)N(t)-\frac{1}{2}(1+\delta)\left\langle N,N\right\rangle(t)\right)\right.
\end{aligned}$$

$$\exp\left((1-\gamma)(1+\delta)N(t)\right)\Big]$$

(apply Hölder's inequality)

$$\leqslant \left(\frac{1+\delta}{\delta}\right)^{1+\delta}\left(\mathbb{E}\left[\exp\left(p\gamma(1+\delta)N(t)-\frac{1}{2}p(1+\delta)\left\langle N,N\right\rangle(t)\right)\right]\right)^{1/p}$$
$$\times\left(\mathbb{E}\left[\exp\left((1-\gamma)(1+\delta)qN(t)\right)\right]\right)^{1/q}$$
$$=\left(\frac{1+\delta}{\delta}\right)^{1+\delta}\left(\mathbb{E}\left[\mathcal{E}\left((1+\varepsilon_0)N\right)(t)\right]\right)^{1/p}\left(\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)\right)\right]\right)^{1/q}$$

(apply (1.80))

$$\leqslant \left(\frac{1+\delta}{\delta}\right)^{1+\delta}\left(\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)\right)\right]\right)^{1/q}. \tag{1.83}$$

From (1.83) we infer

$$\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\mathcal{E}(N)(s)\right]\leqslant \frac{\left(1+\varepsilon_0\right)^2}{\varepsilon_0^2}\left(\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)\right)\right]\right)^{2\varepsilon_0/(1+\varepsilon_0)^2}, \tag{1.84}$$

and hence, since

$$\sup_{t\geqslant 0}\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1+\varepsilon_0\right)N(t)\right)\right]<\infty \tag{1.85}$$

from (1.84) we infer

$$\mathbb{E}\left[\sup_{0\leqslant s<\infty}\mathcal{E}(N)(s)\right]<\infty. \tag{1.86}$$

From (1.86) we obtain that the continuous local martingale $t\mapsto\mathcal{E}(N)(t)$ is in fact a martingale. By writing $\mathcal{E}(N)(\infty)=\lim_{n\to\infty}\mathcal{E}(N)\left(\tau_n\right)$, where τ_n is a sequence of stopping times which increases to ∞ $\mathbb{P}$-almost surely, and which is such that $\mathbb{E}\left[\mathcal{E}(N)\left(\tau_n\right)\right]=1$, $n\in\mathbb{N}$, we obtain by dominated convergence that $\mathbb{E}\left[\mathcal{E}(N)\left(\infty\right)\right]=1$.

This completes the proof of Lemma 1.1. □

In order to prove Assertion (1) in Theorem 1.6 it will be convenient to formulate and prove the following weaker lemma first.

Lemma 1.2. *If* $\liminf_{\varepsilon\downarrow 0}\varepsilon\log\mathbb{E}\left[e^{\frac{1}{2}(1-\varepsilon)\langle M,M\rangle(\infty)}\right]=0$, *then the equality in (1.73) holds.*

Proof. By assumption there exists a sequence of positive real numbers $(\varepsilon_n)_{n\in\mathbb{N}}$ with $0 < \varepsilon_{n+1} < \varepsilon_n \leqslant 1$ such that $\lim_{n\to\infty} \varepsilon_n = 0$, and such that

$$\lim_{n\to\infty} \varepsilon_n \log \mathbb{E}\left[\exp\left(\frac{1}{2}(1-\varepsilon_n)\langle M, M\rangle(\infty)\right)\right] = 0. \tag{1.87}$$

In particular it follows that for every $n \in \mathbb{N}$ we have $1 - \varepsilon_{n+1} = \left(1+\delta_n^2\right)(1-\varepsilon_n)$, for some $\delta_n > 0$, and

$$\mathbb{E}\left[\exp\left(\frac{1}{2}(1-\varepsilon_{n+1})\langle M, M\rangle(\infty)\right)\right] < \infty. \tag{1.88}$$

An application of Lemma 1.1 with $N(t) = (1-\varepsilon_n) M(t)$ and using (1.87) yields the equality $1 = \mathbb{E}\left[\mathcal{E}\left((1-\varepsilon_n) M\right)(\infty)\right]$. Consequently, we see

$$\begin{aligned} 1 &= \mathbb{E}\left[\mathcal{E}\left((1-\varepsilon_n) M\right)(\infty)\right] \\ &= \mathbb{E}\left[\exp\left((1-\varepsilon_n)\left(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty)\right)\right)\right. \\ &\qquad \left.\exp\left(\frac{1}{2}(1-\varepsilon_n)\varepsilon_n\langle M, M\rangle(\infty)\right)\right] \\ &\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{1-\varepsilon_n}\left(\mathbb{E}\left[\exp\left(\frac{1}{2}(1-\varepsilon_n)\langle M, M\rangle(\infty)\right)\right]\right)^{\varepsilon_n}. \end{aligned} \tag{1.89}$$

In (1.89) we let $n \to \infty$ to obtain $1 \leqslant \mathbb{E}\left[\mathcal{E}(M)(\infty)\right]$. Since the process $t \mapsto \mathcal{E}(M)(t)$ is a nonnegative local martingale we also have $\mathbb{E}\left[\mathcal{E}(M)(\infty)\right] \leqslant 1$. As a consequence we see that $\mathbb{E}\left[\mathcal{E}(M)(\infty)\right] = 1$.

This completes the proof of Lemma 1.2. □

Similarly for the proof of (2) in Theorem 1.6 the following weaker lemma turns out to be convenient.

Lemma 1.3. *If* $\liminf_{\varepsilon\downarrow 0} \varepsilon \log \sup_{t\geqslant 0} \mathbb{E}\left[e^{\frac{1}{2}(1-\varepsilon)M(t)}\right] = 0$, *then the equality in (1.73) holds.*

Proof. By assumption there exists a sequence of positive real numbers $(\varepsilon_n)_{n\in\mathbb{N}}$ with $0 < \varepsilon_n < 1$ such that $\lim_{n\to\infty} \varepsilon_n = 0$, and such that

$$\lim_{n\to\infty} \frac{\varepsilon_n}{2-\varepsilon_n} \log \mathbb{E} \sup_{t\geqslant 0}\left[\exp\left(\frac{1}{2}\left(1 - \frac{\varepsilon_n}{2-\varepsilon_n}\right) M(t)\right)\right] = 0. \tag{1.90}$$

In particular it follows that for every $n \in \mathbb{N}$ we have

$$\sup_{t\geqslant 0} \mathbb{E}\left[\exp\left(\frac{1}{2}\left(1 - \frac{\varepsilon_n}{2-\varepsilon_n}\right) M(t)\right)\right] < \infty. \tag{1.91}$$

An application of Lemma 1.1 with $N(t) = (1-\varepsilon_n)\, M(t)$ and using (1.91) yields the equality $1 = \mathbb{E}\left[\mathcal{E}\left((1-\varepsilon_n)\, M\right)(\infty)\right]$. Consequently, we see

$$\begin{aligned}
1 &= \mathbb{E}\left[\mathcal{E}\left((1-\varepsilon_n)\,(M)\,(\infty)\right)\right] \\
&= \mathbb{E}\left[\exp\left((1-\varepsilon_n)^2\left(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty)\right)\right)\right. \\
&\qquad \left.\exp\left((1-\varepsilon_n)\,\varepsilon_n M(\infty)\right)\right] \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{(1-\varepsilon_n)^2}\left(\mathbb{E}\left[\exp\left(\frac{1-\varepsilon_n}{2-\varepsilon_n} M(\infty)\right)\right]\right)^{\varepsilon_n(2-\varepsilon_n)} \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{(1-\varepsilon_n)^2}\left(\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1-\frac{\varepsilon_n}{2-\varepsilon_n}\right) M(\infty)\right)\right]\right)^{\varepsilon_n(2-\varepsilon_n)} \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{(1-\varepsilon_n)^2}\left(\sup_{t\geqslant 0}\mathbb{E}\left[\exp\left(\frac{1}{2}\left(1-\frac{\varepsilon_n}{2-\varepsilon_n}\right) M(t)\right)\right]\right)^{\varepsilon_n(2-\varepsilon_n)}.
\end{aligned} \tag{1.92}$$

In the final step in (1.92) we applied Fatou's lemma. In (1.92) we let $n \to \infty$ to obtain $1 \leqslant \mathbb{E}\left[\mathcal{E}(M)(\infty)\right]$, where we used (1.90). Since the process $t \mapsto \mathcal{E}(N)(t)$ is a nonnegative local martingale we also have $\mathbb{E}\left[\mathcal{E}(M)(\infty)\right] \leqslant 1$. As a consequence we see that $\mathbb{E}\left[\mathcal{E}(M)(\infty)\right] = 1$.

This completes the proof of Lemma 1.3. □

Proof. [Proof of Theorem 1.6.] Assertion (1). Let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,1)$ such that $\varepsilon_n \downarrow 0$, as $n \to \infty$, and such that

$$C_1 := \sup_{n\in\mathbb{N}} \varepsilon_n \log \mathbb{E}\left[\exp\left(\frac{1}{2}(1-\varepsilon_n)\langle M, M\rangle(\infty)\right)\right] < \infty. \tag{1.93}$$

As in the proof of Lemma 1.2 we have with $0 < T < \infty$ fixed

$$\begin{aligned}
1 &= \mathbb{E}\left[\mathcal{E}\left((1-\varepsilon_n)\, M\right)(\infty)\right] \\
&= \mathbb{E}\left[\exp\left((1-\varepsilon_n)\left(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty)\right)\right)\right. \\
&\qquad \left.\times \exp\left(\frac{1}{2}(1-\varepsilon_n)\,\varepsilon_n\langle M, M\rangle(\infty)\right), \langle M, M\rangle(\infty) \leqslant T\right] \\
&\quad + \mathbb{E}\left[\exp\left((1-\varepsilon_n)\left(M(\infty) - \frac{1}{2}\langle M, M\rangle(\infty)\right)\right)\right. \\
&\qquad \left.\exp\left(\frac{1}{2}(1-\varepsilon_n)\,\varepsilon_n\langle M, M\rangle(\infty)\right), \langle M, M\rangle(\infty) > T\right] \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{1-\varepsilon_n} \\
&\qquad \times \left(\mathbb{E}\left[\exp\left(\frac{1}{2}(1-\varepsilon_n)\langle M, M\rangle(\infty)\right), \langle M, M\rangle(\infty) \leqslant T\right]\right)^{\varepsilon_n}
\end{aligned}$$

$$
\begin{aligned}
&+ \left(\mathbb{E}\left[\mathcal{E}(M)(\infty),\, \langle M,M\rangle(\infty) > T\right]\right)^{1-\varepsilon_n} \\
&\quad\times \left(\mathbb{E}\left[\exp\left(\frac{1}{2}(1-\varepsilon_n)\langle M,M\rangle(\infty)\right)\right]\right)^{\varepsilon_n} \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{1-\varepsilon_n} \exp\left(\frac{1}{2}(1-\varepsilon_n)\varepsilon_n T\right) \\
&\quad+ \left(\mathbb{E}\left[\mathcal{E}(M)(\infty),\, \langle M,M\rangle(\infty) > T\right]\right)^{1-\varepsilon_n} \exp(C_1). \qquad (1.94)
\end{aligned}
$$

In (1.94) we let n tend to ∞ to obtain

$$
1 \leqslant \mathbb{E}\left[\mathcal{E}(M)(\infty)\right] + \mathbb{E}\left[\mathcal{E}(M)(\infty),\, \langle M,M\rangle(\infty) > T\right]\exp(C_1). \qquad (1.95)
$$

In (1.95) we let $T \to \infty$ and deduce

$$
1 \leqslant \mathbb{E}\left[\mathcal{E}(M)(\infty)\right] + \mathbb{E}\left[\mathcal{E}(M)(\infty),\, \langle M,M\rangle(\infty) = \infty\right]\exp(C_1). \qquad (1.96)
$$

Since $\mathbb{E}\left[\mathcal{E}(M)(\infty)\right] \leqslant 1$, and $\langle M,M\rangle(\infty) < \infty$ $\mathbb{P}$-almost surely, (1.96) implies $1 = \mathbb{E}\left[\mathcal{E}(M)(\infty)\right]$. This completes the proof of Assertion (1).

Assertion (2). Let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,1)$ be such that $\varepsilon_n \downarrow 0$, as $n \to \infty$, and such that

$$
C_2 := \sup_{n\in\mathbb{N}} \varepsilon_n (2-\varepsilon_n) \log \sup_{t\geqslant 0} \mathbb{E}\left[\exp\left(\frac{1-\varepsilon_n}{2-\varepsilon_n} M(t)\right)\right] < \infty. \qquad (1.97)
$$

As in the proof of Lemma 1.1 (see also Lemma 1.3) we have with $0 < T < \infty$ fixed

$$
\begin{aligned}
1 &= \mathbb{E}\left[\mathcal{E}\left((1-\varepsilon_n)M\right)(\infty)\right] \\
&= \mathbb{E}\left[\exp\left((1-\varepsilon_n)^2\left(M(\infty) - \frac{1}{2}\langle M,M\rangle(\infty)\right)\right)\right. \\
&\qquad \left.\times \exp\left((1-\varepsilon_n)\varepsilon_n M(\infty)\right),\, M(\infty) \leqslant T\right] \\
&\quad + \mathbb{E}\left[\exp\left((1-\varepsilon_n)^2\left(M(\infty) - \frac{1}{2}\langle M,M\rangle(\infty)\right)\right)\right. \\
&\qquad \left.\exp\left((1-\varepsilon_n)\varepsilon_n M(\infty)\right),\, M(\infty) > T\right] \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{(1-\varepsilon_n)^2} \\
&\qquad \times \left(\mathbb{E}\left[\exp\left(\frac{1-\varepsilon_n}{2-\varepsilon_n} M(\infty)\right),\, M(\infty) \leqslant T\right]\right)^{2\varepsilon_n - \varepsilon_n^2} \\
&\quad + \left(\mathbb{E}\left[\mathcal{E}(M)(\infty),\, M(\infty) > T\right]\right)^{(1-\varepsilon_n)^2} \\
&\qquad \times \left(\mathbb{E}\left[\exp\left(\frac{1-\varepsilon_n}{2-\varepsilon_n} M(\infty)\right)\right]\right)^{2\varepsilon_n - \varepsilon_n^2} \\
&\leqslant \left(\mathbb{E}\left[\mathcal{E}(M)(\infty)\right]\right)^{(1-\varepsilon_n)^2} \exp\left((1-\varepsilon_n)\varepsilon_n T\right)
\end{aligned}
$$

$$+ \left(\mathbb{E}\left[\mathcal{E}(M)(\infty),\, M(\infty) > T\right]\right)^{(1-\varepsilon_n)^2} \exp\left(C_2\right). \tag{1.98}$$

In (1.98) we let n tend to ∞ to obtain

$$1 \leqslant \mathbb{E}\left[\mathcal{E}(M)(\infty)\right] + \mathbb{E}\left[\mathcal{E}(M)(\infty),\, M(\infty) > T\right] \exp\left(C_2\right). \tag{1.99}$$

In (1.99) we let $T \to \infty$ and deduce

$$1 \leqslant \mathbb{E}\left[\mathcal{E}(M)(\infty)\right] + \mathbb{E}\left[\mathcal{E}(M)(\infty),\, M(\infty) = \infty\right] \exp\left(C_2\right). \tag{1.100}$$

Since $\mathbb{E}\left[\mathcal{E}(M)(\infty)\right] \leqslant 1$, and $M(\infty) < \infty$ $\mathbb{P}$-almost surely, (1.100) implies $1 = \mathbb{E}\left[\mathcal{E}(M)(\infty)\right]$. This completes the proof of Assertion (2).

Altogether this completes the proof of Theorem 1.6. □

In [Krylov (2002)] Krylov shows by way of an example that his results are really stronger than those of Novikov [Novikov (1973)] and Kazamaki [Kazamaki (1978)].

Definition 1.3. The equation in (1.36) is said to have unique pathwise solutions, if for any Brownian motion $\{(B(t) : t \geqslant 0), (\Omega, \mathcal{F}, \mathbb{P})\}$ and any pair of adapted processes $\{X(t) : t \geqslant 0\}$ and $\{X'(t) : t \geqslant 0\}$ for which

$$X(t) = x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right) ds \quad \text{and} \tag{1.101}$$

$$X'(t) = x + \int_0^t \sigma\left(s, X'(s)\right) dB(s) + \int_0^t b\left(s, X'(s)\right) ds \tag{1.102}$$

it follows that $X(t) = X'(t)$ $\mathbb{P}$-almost surely for all $t \geqslant 0$.

Pathwise solutions are also called strong solutions. A version of the following result (Itô's theorem) can be found in many books on stochastic differential equations: see e.g. [Ikeda and Watanabe (1998); Øksendal and Reikvam (1998); Revuz and Yor (1999)].

Theorem 1.7. *Let $\sigma_{j,k}(s,x)$ and $b_j(s,x)$, $1 \leqslant j,\, k \leqslant d$ be continuous functions defined on $[0,\infty) \times \mathbb{R}^d$ such that for all $t > 0$ there exists a constant $K(t)$ with the property that*

$$\sum_{j,k=1}^{d} \left|\sigma_{j,k}(s,x) - \sigma_{j,k}(s,y)\right|^2 + \sum_{j=1}^{d} \left|b_j(s,x) - b_j(s,y)\right|^2 \leqslant K(t)\, |x-y|^2 \tag{1.103}$$

for all $0 \leqslant s \leqslant t$, and all $x,\, y \in \mathbb{R}^d$. Fix $x \in \mathbb{R}^d$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t\geqslant 0}$. Moreover, let $\{B(t) : t \geqslant 0\}$ be *a Brownian motion on the filtered probability space* $(\Omega, \mathcal{F}_t, \mathbb{P})$. *Then there exists an* $\mathbb{R}^d$*-valued process* $\{X(t) : t \geqslant 0\}$ *such that*

$$X(t) = x + \int_0^t \sigma(s, X(s))\, dB(s) + \int_0^t b(s, X(s))\, ds, \quad t \geqslant 0.$$

This process is pathwise unique in the sense of Definition 1.3.

The following theorem shows that stochastic differential equations having unique strong solutions also possess unique weak solutions.

Theorem 1.8. *Let the vector and matrix functions* $b(s,x)$ *and* $\sigma(s,x)$ *be as in Theorem 1.4. Fix* $x \in \mathbb{R}^d$. *Suppose that the stochastic (integral) equation*

$$X(t) = x + \int_0^t \sigma(s, X(s))\, dB(s) + \int_0^t b(s, X(s))\, ds \tag{1.104}$$

possesses unique pathwise solutions. Then this equation has unique weak solutions.

In the proof we employ a certain coupling argument. In fact weak solutions to the equations in (1.101) and (1.102) are recast as two pathwise solutions on the same probability space.

Proof. Let

$$\{(B(t) : t \geqslant 0), (\Omega, \mathcal{F}, \mathbb{P})\} \quad \text{and} \quad \{(B'(t) : t \geqslant 0), (\Omega', \mathcal{F}', \mathbb{P}')\}$$

be two Brownian motions. Let $\{X(t) : t \geqslant 0\}$ be an adapted process which satisfies (1.101), and let $\{X'(t) : t \geqslant 0\}$ be an adapted process which satisfies (1.102). Suppose $0 \leqslant t_1 < t_2 < \cdots < t_n < \infty$, and let $C_1, \ldots, C_n$ be Borel subsets of $\mathbb{R}^d$. We have to prove the equality:

$$\mathbb{P}'\left[X'(t_1) \in C_1, \ldots, X'(t_n) \in C_n\right] = \mathbb{P}\left[X(t_1) \in C_1, \ldots, X(t_n) \in C_n\right]. \tag{1.105}$$

Let $\Omega_0 = C\left([0,\infty), \mathbb{R}^d\right)$ be the space of $\mathbb{R}^d$-valued continuous functions defined on $[0,\infty)$. This space is equipped with its standard filtration, which originates from the coordinate mappings: $\omega \mapsto \omega(t)$, $t \geqslant 0$, and its Borel field. Define the $\mathbb{R}^d$-valued processes $Y(t)$, $Y'(t)$, and $B_0(t)$ on $\Omega \times \Omega' \times \Omega_0$ as follows:

$$\begin{cases} Y(t)(\omega, \omega', \omega_0) = \omega(t), & (\omega, \omega', \omega_0) \in \Omega \times \Omega' \times \Omega_0; \\ Y'(t)(\omega, \omega', \omega_0) = \omega'(t), & (\omega, \omega', \omega_0) \in \Omega \times \Omega' \times \Omega_0; \\ B_0(t)(\omega, \omega', \omega_0) = \omega_0(t), & (\omega, \omega', \omega_0) \in \Omega \times \Omega' \times \Omega_0. \end{cases} \tag{1.106}$$

In fact we use the notation Ω_0 instead of Ω to distinguish the third component of the space $\Omega \times \Omega' \times \Omega_0$ from the first. The role of the first two components are very similar; the third component is related to the driving Brownian motion $\{B_0(t) : t \geqslant 0\}$. The processes $Y(t)$ and $Y'(t)$ are going to be the pathwise solutions on the same probability space $\left(\Omega \times \Omega' \times \Omega_0, \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{F}^0, \widetilde{\mathbb{P}}_x\right)$: see (1.116) and (1.117) below. On Ω_0 the probability measure $\mathbb{P}_0$ is determined by prescribing its finite-dimensional distributions by the equality:

$$\begin{aligned} &\mathbb{P}_0\left[\left(\omega_0\left(t_1\right), \ldots, \omega_0\left(t_n\right)\right) \in D\right] = \mathbb{P}\left[\left(B\left(t_1\right), \ldots, B\left(t_n\right)\right) \in D\right] \\ &\quad = \mathbb{P}'\left[\left(B'\left(t_1\right), \ldots, B'\left(t_n\right)\right) \in D\right]. \end{aligned} \tag{1.107}$$

Here $0 \leqslant t_1 < \cdots < t_n < \infty$, and D is a Borel subset of $\left(\mathbb{R}^d\right)^n$. Let C be another Borel subset of $\left(\mathbb{R}^d\right)^n$. On $\Omega \times \Omega_0$ and $\Omega' \times \Omega_0$ we define the probability measures $\mathbb{Q}_x$ respectively $\mathbb{Q}'_x$ by the equalities:

$$\begin{aligned} &\mathbb{Q}_x\left[\left(\omega\left(t_1\right), \ldots, \omega\left(t_n\right)\right) \in C,\ \left(\omega_0\left(t_1\right), \ldots, \omega_0\left(t_n\right)\right) \in D\right] \\ &= \mathbb{P}\left[\left(B\left(t_1\right), \ldots, B\left(t_n\right)\right) \in D,\ \left(B\left(t_1\right), \ldots, B\left(t_n\right)\right) \in D\right] \\ &= \mathbb{P}'\left[\left(B\left(t_1\right), \ldots, B\left(t_n\right)\right) \in D,\ \left(B'\left(t_1\right), \ldots, B'\left(t_n\right)\right) \in D\right]. \end{aligned} \tag{1.108}$$

Notice that $\mathbb{P}_0\left[A_0\right] = 0$ implies $\mathbb{Q}_x\left[\Omega \times A_0\right] = \mathbb{Q}'_x\left[\Omega' \times A_0\right] = 0$. Consequently, by Radon-Nikodym's theorem there are (measurable) functions

$$Q_x, \text{ and } Q'_x : \mathcal{F} \times \Omega_0 \to [0,1]$$

such that

$$\begin{aligned} \mathbb{Q}_x\left[A \times A_0\right] &= \int_{A_0} Q_x\left(A, \omega_0\right) \mathbb{P}_0\left(\omega_0\right), \quad A \in \mathcal{F},\ A_0 \in \mathcal{F}^0, \quad \text{and} \\ \mathbb{Q}'_x\left[A' \times A_0\right] &= \int_{A_0} Q'_x\left(A, \omega_0\right) \mathbb{P}_0\left(\omega_0\right), \quad A' \in \mathcal{F}',\ A_0 \in \mathcal{F}^0. \end{aligned} \tag{1.109}$$

Here $Q_x\left(\Omega, \omega_0\right) = Q'_x\left(\Omega', \omega_0\right) = 1$ for all $\omega_0 \in \Omega_0$. Moreover, the functions

$$\omega_0 \mapsto Q_x\left(A, \omega_0\right), \quad \text{and} \quad \omega_0 \mapsto Q'_x\left(A, \omega_0\right) \tag{1.110}$$

are measurable relative to the $\mathbb{P}_0$-completion of $\mathcal{F}$. Finally, we define the measure

$$\widetilde{\mathbb{Q}}_x : \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{F}^0 \to [0,1] \quad \text{by}$$

$$\widetilde{\mathbb{Q}}_x\left[A \times A' \times A_0\right] = \int_{A_0} Q_x\left(A, \omega_0\right) Q'_x\left(A', \omega_0\right) d\mathbb{P}_0\left(\omega_0\right). \tag{1.111}$$

Here A, A', and A_0 belong to $\mathcal{F}$, $\mathcal{F}'$, and $\mathcal{F}^0$ respectively. First we prove that the process $\{B_0(t) : t \geqslant 0\}$ is a Brownian motion with respect to the measure $\widetilde{\mathbb{Q}}_x$. From the proof of Theorem 1.1 (Lévy's theorem) it follows that it suffices to show that the following equality holds:

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\exp\left(-i\left\langle \xi, B_0(t)-B_0(s)\right\rangle\right) \,\middle|\, \mathcal{F}_s\otimes\mathcal{F}'_s\otimes\mathcal{F}^0_s\right]\\ &=\exp\left(-\frac{1}{2}\left|\xi\right|^2(t-s)\right),\quad t>s\geqslant 0,\ \xi\in\mathbb{R}^d. \end{aligned}\tag{1.112}$$

In order to prove (1.112) we pick A, A', and A_0 in $\mathcal{F}_s$, $\mathcal{F}'_s$, and $\mathcal{F}^0_s$ respectively. Then by (1.111) we get

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\exp\left(-i\left\langle \xi, B_0(t)-B_0(s)\right\rangle\right)\mathbf{1}_{A\times A'\times A_0}\right]\\ &=\int_{A\times A'\times A_0}\exp\left(-i\left\langle \xi, B_0(t)-B_0(s)\right\rangle\right)\,d\widetilde{\mathbb{Q}}_x\\ &=\int_{A_0}\exp\left(-i\left\langle \xi, \omega_0(t)-\omega_0(s)\right\rangle\right)Q_x\left(A,\omega_0\right)Q'_x\left(A',\omega_0\right)\,d\mathbb{P}_0\left(\omega_0\right). \end{aligned}\tag{1.113}$$

The process $(\omega_0, t)\mapsto\omega_0(t)$ is a Brownian motion relative to $\mathbb{P}_0$, and the events A, A', and A_0 belong to $\mathcal{F}_s$, $\mathcal{F}'_s$, and $\mathcal{F}^0_s$ respectively, and hence $B_0(t)-B_0(s)$ is $\mathbb{P}_0$-independent of $A\times A'\times A_0$. Therefore (1.113) implies

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\exp\left(-i\left\langle \xi, B_0(t)-B_0(s)\right\rangle\right)\mathbf{1}_{A\times A'\times A_0}\right]\\ &=\int_{A_0}Q_x\left(A,\omega_0\right)Q'_x\left(A',\omega_0\right)\,d\mathbb{P}_0\left(\omega_0\right)\int\exp\left(-i\left\langle \xi, \omega_0(t)-\omega_0(s)\right\rangle\right)\,d\mathbb{P}_0\left(\omega_0\right)\\ &=\widetilde{\mathbb{Q}}_x\left[A\times A'\times A_0\right]\,\exp\left(-\frac{1}{2}\left|\xi\right|^2(t-s)\right). \end{aligned}\tag{1.114}$$

The equality in (1.112) is a consequence of (1.114). Since, by definition (see (1.107))

$$\mathbb{P}_0\left[\left(\omega_0\left(t_1\right),\dots,\omega_0\left(t_n\right)\right)\in C\right]=\mathbb{P}\left[\left(B\left(t_1\right),\dots,B\left(t_n\right)\right)\in C\right]\tag{1.115}$$

for $0\leqslant t_1<\cdots<t_n<\infty$, C Borel subset of $\left(\mathbb{R}^d\right)^n$, and since the process $\{B(t):t\geqslant 0\}$ is Brownian motion relative to $\mathbb{P}$, the same is true for the process $(\omega_0,t)\mapsto\omega_0(t)$ relative to $\mathbb{P}_0$. Next we compute the quantity:

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\left|Y(t)-x-\int_0^t\sigma\left(s,Y(s)\right)\,dB_0(s)-\int_0^t b\left(s,Y(s)\right)ds\right|\right]\\ &=\int\left|\omega(t)-x-\int_0^t\sigma\left(s,\omega(s)\right)\,d\omega_0(s)-\int_0^t b\left(s,\omega(s)\right)ds\right|\,\widetilde{\mathbb{Q}}\left(d\omega,d\omega',d\omega_0\right)\\ &=\int\left|\omega(t)-x-\int_0^t\sigma\left(s,\omega(s)\right)\,d\omega_0(s)-\int_0^t b\left(s,\omega(s)\right)ds\right|\,\mathbb{Q}\left(d\omega,d\omega_0\right)\\ &=\int\left|X(t)-x-\int_0^t\sigma\left(s,X(s)\right)\,dB(s)-\int_0^t b\left(s,X(s)\right)ds\right|\,d\mathbb{P}=0. \end{aligned}\tag{1.116}$$

Similarly we have

$$\widetilde{\mathbb{E}}_x\left[\left|Y'(t)-x-\int_0^t \sigma\left(s,Y'(s)\right)\,dB_0(s)-\int_0^t b\left(s,Y'(s)\right)\,ds\right|\right]=0. \quad (1.117)$$

From (1.116) and (1.117) we infer that the following equalities hold $\widetilde{\mathbb{Q}}_x$-almost surely:

$$Y(t)=x+\int_0^t \sigma\left(s,Y(s)\right)dB_0(s)+\int_0^t b\left(s,Y(s)\right)\,ds \quad \text{and} \quad (1.118)$$

$$Y'(t)=x+\int_0^t \sigma\left(s,Y'(s)\right)dB_0(s)+\int_0^t b\left(s,Y'(s)\right)\,ds. \quad (1.119)$$

Moreover, the process $\{B_0(t) : t \geqslant 0\}$ is a Brownian motion relative to $\widetilde{\mathbb{Q}}_x$. From the pathwise uniqueness and the equalities (1.118) and (1.119) we see that, $\widetilde{\mathbb{Q}}_x$-almost surely,

$$Y(t)=Y'(t), \quad t \geqslant 0. \quad (1.120)$$

Let $0 \leqslant 0 < t_1 < \cdots < t_n < \infty$, and let C be a Borel subset of $\left(\mathbb{R}^d\right)^n$. From (1.120) it follows that

$$\widetilde{\mathbb{Q}}_x\left[\left(Y\left(t_1\right),\ldots,Y\left(t_n\right)\right)\in C\right]=\widetilde{\mathbb{Q}}_x\left[\left(Y'\left(t_1\right),\ldots,Y'\left(t_n\right)\right)\in C\right]. \quad (1.121)$$

Using (1.121) and the definition of the measure $\widetilde{\mathbb{Q}}_x$ show that the following identities are self-explanatory:

$$\begin{aligned}&\widetilde{\mathbb{Q}}_x\left[\left(Y\left(t_1\right),\ldots,Y\left(t_n\right)\right)\in C\right]\\&=\mathbb{Q}_x\left[\left(\omega\left(t_1\right),\ldots,\omega\left(t_n\right)\right)\in C,\ \left(\omega_0\left(t_1\right),\ldots,\omega_0\left(t_n\right)\right)\in\Omega_0\right]\\&=\mathbb{P}\left[\left(X\left(t_1\right),\ldots,X\left(t_n\right)\right)\in C\right].\end{aligned} \quad (1.122)$$

The definition of the measure $\widetilde{\mathbb{Q}}_x$ is given in (1.111). Similarly we conclude

$$\widetilde{\mathbb{Q}}_x\left[\left(Y\left(t_1\right),\ldots,Y\left(t_n\right)\right)\in C\right]=\mathbb{P}\left[\left(X'\left(t_1\right),\ldots,X'\left(t_n\right)\right)\in C\right]. \quad (1.123)$$

From (1.122), (1.123), and (1.121) we obtain

$$\mathbb{P}\left[\left(X\left(t_1\right),\ldots,X\left(t_n\right)\right)\in C\right]=\mathbb{P}\left[\left(X'\left(t_1\right),\ldots,X'\left(t_n\right)\right)\in C\right]. \quad (1.124)$$

The equality in (1.124) implies that the finite-dimensional distributions of the solution in equation in (1.101) are the same as those of the solution of equation (1.102). So that stochastic differential equations with unique pathwise solutions also possess unique weak (or distributional) solutions.

This concludes the proof of Theorem 1.8. □

The following result is often very useful.

Theorem 1.9. *Let $M(s)$, $t \leqslant s \leqslant T$, be a continuous local L^2-martingale taking values in $\mathbb{R}^k$. Put $M^*(s) = \sup_{t\leqslant\tau\leqslant s} |M(\tau)|$. Fix $0 < p < \infty$. The Burkholder-Davis-Gundy inequality says that there exist universal finite and strictly positive constants c_p and C_p such that*

$$c_p\mathbb{E}\left[(M^*(s))^{2p}\right] \leqslant \mathbb{E}\left[\langle M(\cdot), M(\cdot)\rangle^p(s)\right] \leqslant C_p\mathbb{E}\left[(M^*(s))^{2p}\right], \quad t \leqslant s \leqslant T. \tag{1.125}$$

If $p = 1$, then $c_p = \frac{1}{4}$, $C_1 = 1$, and if $p = \frac{1}{2}$, then $c_p = \frac{1}{8}\sqrt{2}$, $C_p = 2$. For more details and a proof using stochastic calculus see e.g. [Ikeda and Watanabe (1998)]. A proof based on good λ-inequalities can be found in [Rogers and Williams (2000)] or [Durrett (1984, 1996)].

Another result we need is the following one on tightness.

Theorem 1.10. *Let $\{X^n(t) : t \geqslant 0\}$ be a sequence of continuous $\mathbb{R}^d$-valued processes satisfying the following the following two conditions:*

(a) $\lim_{N\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}\left[|X^n(0)| > N\right] = 0$;
(b) For every $T > 0$ and $\varepsilon > 0$ the following equality holds:

$$\lim_{h\downarrow 0} \sup_{n\in\mathbb{N}} \mathbb{P}\left[\max_{s,t\in[0,T],\,|t-s|\leqslant h} |X^n(t) - X^n(s)| > \varepsilon\right] = 0.$$

Then there exist a subsequence $n_1 < n_2 < \cdots$, a probability space $\left(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}\right)$, d-dimensional continuous stochastic processes $\left\{\widehat{X}^{n_k}(t) : t \geqslant 0\right\}$, $k \in \mathbb{N}$, and $\left\{\widehat{X}(t) : t \geqslant 0\right\}$ defined on this probability space with the following properties:

(1) The finite-dimensional $\widehat{\mathbb{P}}$-distributions of the process $\left\{\widehat{X}^{n_k}(t) : t \geqslant 0\right\}$ coincide with the finite-dimensional $\mathbb{P}$-distributions of $\{X^{n_k}(t) : t \geqslant 0\}$ for $k = 1, 2, \ldots$.
(2) The sequence $\left\{\widehat{X}^{n_k}(t) : t \geqslant 0\right\}_{k\in\mathbb{N}}$ converges to the process $\left\{\widehat{X}(t) : t \geqslant 0\right\}$ in the sense that

$$\widehat{\mathbb{P}}\left[\widehat{\omega} \in \widehat{\Omega} : \lim_{k\to\infty} d\left(\widehat{X}^{n_k}(\widehat{\omega}), \widehat{X}(\widehat{\omega})\right) = 0\right] = 1.$$

Here

$$d(w, w') = \sum_{n=1}^{\infty} 2^{-n} \min\left(1, \max_{0\leqslant s\leqslant n} \left|w(s) - w'(s)\right|\right), \quad w,\ w' \in C\left([0,\infty), \mathbb{R}^d\right).$$

Moreover, if every finite-dimensional distribution of the image measures $\mathbb{P}^{X_n}$ converges as $n \to \infty$, then there is no need to take subsequences: the sequence $n_k = k$ will do.

The conditions in Theorem 1.10 can be verified by appealing to the results in the following theorem.

Theorem 1.11. *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of d-dimensional processes satisfying the following two conditions:*

(a) There exist strictly positive finite constants M and γ such that

$$\mathbb{E}\left[|X^n(0)|^\gamma\right] \leqslant M < \infty, \quad n \in \mathbb{N}.$$

(b) There exist strictly positive finite constants α, β, M_k, $k = 1, 2, \dots$, such that for all $n \in \mathbb{N}$ and for all s, $t \in [0, k]$ the inequality

$$\mathbb{E}\left[|X^n(t) - X^n(s)|^\alpha\right] \leqslant M_k \left|t - s\right|^{1+\beta}$$

holds for $k = 1, 2, \dots$.

Then the sequence $\{X^n(t) : t \geqslant 0\}_{n\in\mathbb{N}}$ satisfies the conditions (a) and (b) of Theorem 1.10.

As a corollary we have the following result.

Corollary 1.5. *Let $\{X(t) : t \geqslant 0\}$ be a family of d-dimensional random variables such that for some finite strictly positive constants α, β, and M_k, $k = 1, 2, \dots$, the following inequalities are valid:*

$$\mathbb{E}\left[|X(t) - X(s)|^\alpha\right] \leqslant M_k \left|t - s\right|^{1+\beta}, \quad s, \in [0, k], \ k = 1, 2, \dots .$$

Then there exists a d-dimensional continuous process $\left\{\widehat{X}(t) : t \geqslant 0\right\}$ such that $X(t) = \widehat{X}(t)$ $\mathbb{P}$-almost surely for all $t \geqslant 0$.

We conclude this section with a result of Skorohod [Skorokhod (1965)].

Theorem 1.12. *Let $\sigma_{j,k}(s, x)$, $1 \leqslant j$, $k \leqslant d$, and $b_j(s, x)$, $1 \leqslant j \leqslant d$, be be bounded continuous real-valued functions on $[0, \infty) \times \mathbb{R}^d$, and let $x \in \mathbb{R}^d$. Then there exists a probability measure $\mathbb{P}$ on the Borel field of $C\left([0, \infty), \mathbb{R}^d\right)$ and a Brownian motion relative to this measure $\mathbb{P}$ such that the process defined $\{X(t) : t \geqslant 0\}$ defined by $X(t)(\omega) = \omega(t)$, $t \geqslant 0$, $\omega \in C\left([0, \infty), \mathbb{R}^d\right)$, satisfies the equality*

$$X(t) = x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right) ds, \quad \mathbb{P}\text{-almost surely.} \tag{1.126}$$

Here $\sigma(s,y) = (\sigma_{j,k}(s,y))_{j,k=1}^d$, and $b(s,y)$ is the column vector with entries $b_j(s,y)$, $1 \leqslant j \leqslant d$.

Proof. [Outline of a proof of Theorem 1.12.] Define the differential operators $L(s)$, $s \geqslant 0$, by

$$L(s)f(s) = \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(s,x)\frac{\partial^2 f}{\partial x_j \partial x_k}(x) + \sum_{i=1}^d b_j(s,x)\frac{\partial f}{\partial x_i}(x) \tag{1.127}$$

where the function f is twice continuously differentiable, and where the coefficients $a_{i,j}(s,x)$ are given by

$$a_{i,j}(s,x) = \sum_{k=1}^d \sigma_{i,k}(s,x)\sigma_{j,k}(s,x) = (\sigma(s,x)\sigma^*(s,x))_{i,j}\,.$$

Fix $x \in \mathbb{R}^d$. From assertion (iii) in Theorem 1.3 we see that it suffices that there exists a probability measure $\mathbb{P}$ on the space $W = C\left([0,\infty),\mathbb{R}^d\right)$ and a function $X \in W$ such that $\mathbb{P}[X(0) = x] = 1$, and such that for every $f \in C_{00}^2\left(\mathbb{R}^d\right)$ (i.e. f is twice continuously differentiable and has compact support in $\mathbb{R}^d$) the process

$$f(X(t)) - f(X(0)) - \int_0^t L(s)f(X(s))\,ds$$

is a $\mathbb{P}$-martingale. On W we take the filtration generated by the coordinate functions: $\omega \mapsto \omega(t)$, $\omega \in W$, $t \geqslant 0$. Let $(\Omega', \mathcal{F}'_t, \mathbb{P}')_{t\geqslant 0}$ be a filtered probability space, and let $\{B'(t) : t \geqslant 0\}$ be a Brownian motion with respect to $\mathbb{P}'$. Define for $\ell \in \mathbb{N}$, the function $\varphi^\ell : [0,\infty) \to [0,\infty)$ by

$$\varphi^\ell(t) = \sum_{k=0}^\infty k2^{-\ell}\mathbf{1}_{[k2^{-\ell},(k+1)2^{-\ell})},$$

and put $\sigma^\ell(t,y) = \sigma\left(\varphi^\ell(t), y\right)$, $b^\ell(t,y) = b\left(\varphi^\ell(t), y\right)$. Define the processes $Y^\ell(t)$, $\ell \in \mathbb{N}$, by the equality:

$$\begin{aligned}
Y^\ell(t) &= x + \sum_{k=0}^\infty \left\{\sigma\left(k2^{-\ell}, Y^\ell\left(\min\left(t,k2^{-\ell}\right)\right)\right)\left(B'(t) - B'\left(\min\left(t,k2^{-\ell}\right)\right)\right)\right.\\
&\qquad\left. + b\left(k2^{-\ell}, Y^\ell\left(\min\left(t,k2^{-\ell}\right)\right)\right)\left(t - \min\left(t,k2^{-\ell}\right)\right)\right\}\\
&= x + \int_0^t \sigma\left(\varphi^\ell(\tau), Y^\ell(\tau)\right) dB'(\tau) + \int_0^t b\left(\varphi^\ell(\tau), Y^\ell(\tau)\right)\,d\tau\\
&= x + \int_0^t \sigma^\ell\left(\tau, Y^\ell(\tau)\right) dB'(\tau) + \int_0^t b^\ell\left(\tau, Y^\ell(\tau)\right)\,d\tau.
\end{aligned} \tag{1.128}$$

Notice that the equalities in (1.128) yield a genuine definition of the process $t \mapsto Y^\ell(t)$, $t \geqslant 0$, because (1.128) can be considered as a recursive definition of the process $Y^\ell(t)$, $k2^{-\ell} \leqslant t < (k+1)2^{-\ell}$ where recursion is done with respect to k, $k = 0,\ 1,\ ,\ldots$. In principle we want to take the limit in (1.128) for $\ell \to \infty$, and obtain an equality of the form:

$$Y(t) = x + \int_0^t \sigma\left(\tau, Y(\tau)\right) dB'(\tau) + \int_0^t b\left(\tau, Y(\tau)\right) d\tau. \tag{1.129}$$

However, this cannot be done directly. We need some results on moment inequalities for continuous martingales, like the Burkholder-Davis-Gundy inequality (see Theorem 1.9), and on weak convergence of continuous adapted stochastic processes, like the Skorohod-Dudley-Wichura representation theorem (see Theorem 1.5), which we essentially speaking used in Theorem 1.10. Using moment inequalities for martingale it is shown that the sequence $\left(Y^\ell\right)_{\ell\in\mathbb{N}}$ converges weakly. By an application of the Skorohod-Dudley-Wichura representation theorem we may assume that, possibly after changing the filtered probability space that the sequence Y^ℓ converges almost surely to some random variable Y which is defined on $C\left([0,\infty),\mathbb{R}^d\right)$. Then Y can be considered as a weak solution to the equation in (1.126).

From the equalities in (1.128) it follows that

$$Y^\ell(t) - Y^\ell(s) = \int_s^t \sigma^\ell\left(\tau, Y^\ell(\tau)\right)\, dB'(\tau) + \int_s^t b^\ell\left(\tau, Y^\ell(\tau)\right)\, d\tau, \quad 0 \leqslant s \leqslant t. \tag{1.130}$$

So that we have

$$\begin{aligned}
&\left|Y^\ell(t) - Y^\ell(s)\right|^{2m} \\
&\leqslant 2^{2m}\left(\left|\int_s^t \sigma^\ell\left(\tau, Y^\ell(\tau)\right)\, dB'(\tau)\right|^{2m} + \left|\int_s^t b^\ell\left(\tau, Y^\ell(\tau)\right)\, d\tau\right|^{2m}\right) \\
&\leqslant 4^m\left(\left|\int_s^t \sigma^\ell\left(\tau, Y^\ell(\tau)\right)\, dB'(\tau)\right|^{2m} + (t-s)^{2m}\left\|b\right\|_\infty^{2m}\right).
\end{aligned} \tag{1.131}$$

From the Burkholder-Davis-Gundy inequality (1.125) in Theorem 1.9 with $p = m$ and (1.131) we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|Y^\ell(t) - Y^\ell(s)\right|^{2m}\right] \\
&\leqslant 2^{2m}\left(\mathbb{E}\left[\left|\int_s^t \sigma^\ell\left(\tau, Y^\ell(\tau)\right)\, dB'(\tau)\right|^{2m}\right] + (t-s)^{2m}\left(\sum_{i=1}^d \left\|b_i\right\|_\infty^2\right)^m\right)
\end{aligned}$$

$$
\begin{aligned}
&\leqslant 4^m d^{m-1} C_m \mathbb{E}\left[\left(\sum_{i=1}^{d}\int_s^t a_{i,i}^{\ell}\left(\tau, Y^{\ell}(\tau)\right)\, d\tau\right)^m\right] \\
&\quad + 4^m (t-s)^{2m}\left(\sum_{i=1}^{d}\|b_i\|_{\infty}^2\right)^m \\
&\leqslant 4^m (t-s)^m\left(d^{m-1}C_m\left(\sum_{i=1}^{d}\|a_{i,i}\|_{\infty}\right)^m + (t-s)^m\left(\sum_{i=1}^{d}\|b_i\|_{\infty}^2\right)^m\right).
\end{aligned}
\tag{1.132}
$$

Here, of course, $a_{i,j}^{\ell}(s,y) = \sum_{k=1}^{d}\sigma_{i,k}^{\ell}(s,y)\sigma_{j,k}^{\ell}(s,y)$, and the covariation process of the martingale (or more precise the martingale after time s) $\left\{\int_s^t \sigma^{\ell}\left(\tau, Y^{\ell}(\tau)\right) dB'(\tau) : t \geqslant s\right\}$ is given by the matrix process:

$$
\left\{\left(\int_s^t a_{i,j}^{\ell}\left(\tau, Y^{\ell}(\tau)\right)\, d\tau\right)_{i,j=1}^{d} : t \geqslant s\right\}.
$$

Hence we may apply Theorem 1.11 (with $\alpha = 4$, $\beta = 1$, which corresponds to $m = 2$) to infer that the sequence $\left\{Y^{\ell}(t) : t \geqslant 0\right\}$, $\ell = 1, 2, \ldots$ satisfies conditions (a) and (b) of Theorem 1.10. From Theorem 1.10 it follows that there exists a probability space $\left(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}\right)$ together with processes $\left\{\widehat{Y}^{n_k}(t) : t \geqslant 0\right\}$, $k \in \mathbb{N}$, and $\left\{\widehat{Y}(t) : t \geqslant 0\right\}$ defined on this probability space with the following properties:

(1) The finite-dimensional $\widehat{\mathbb{P}}$-distributions of the process $\left\{\widehat{Y}^{n_k}(t) : t \geqslant 0\right\}$ coincides with the finite-dimensional $\mathbb{P}'$-distributions of the process $\{Y^{n_k}(t) : t \geqslant 0\}$ for $k = 1, 2, \ldots$.
(2) The sequence $\left\{\widehat{Y}^{n_k}(t) : t \geqslant 0\right\}_{k \in \mathbb{N}}$ converges on compact subsets of $[0, \infty)$ to the process $\left\{\widehat{Y}(t) : t \geqslant 0\right\}$ in the sense that

$$
\widehat{\mathbb{P}}\left[\widehat{\omega} \in \widehat{\Omega} : \lim_{k\to\infty} d\left(\widehat{Y}^{n_k}(\widehat{\omega}), \widehat{Y}(\widehat{\omega})\right) = 0\right] = 1.
$$

Next let f be a bounded C^2-function on $\mathbb{R}^d$, let g be bounded continuous functions defined on $\left(\mathbb{R}^d\right)^n$, and let $0 \leqslant s_1 < s_2 < \cdots < s_n \leqslant s < t$. Then we have

$$
\begin{aligned}
&\widehat{\mathbb{E}}\left[\left(f\left(\widehat{Y}(t)\right) - f\left(\widehat{Y}(s)\right) - \int_s^t L(\tau) f\left(\widehat{Y}(\tau)\right) d\tau\right) g\left(\widehat{Y}(s_1), \ldots, \widehat{Y}(s_n)\right)\right] \\
&= \lim_{k\to\infty}\widehat{\mathbb{E}}\left[\left(f\left(\widehat{Y}^{n_k}(t)\right) - f\left(\widehat{Y}^{n_k}(s)\right) - \int_s^t L_{\tau}^{(n_k)} f\left(\widehat{Y}^{n_k}(\tau)\right) d\tau\right)\right. \\
&\qquad \left.\times g\left(\widehat{Y}^{n_k}(s_1), \ldots, \widehat{Y}^{n_k}(s_n)\right)\right]
\end{aligned}
$$

(the finite-dimensional $\widehat{\mathbb{P}}$-distributions of the process $\widehat{Y}^{n_k}$ coincide with finite-dimensional $\mathbb{P}'$-distributions of Y^{n_k}, $k = 1, 2, \ldots$.)

$$\begin{aligned} &= \lim_{k\to\infty} \mathbb{E}'\left[\left(f\left(Y^{n_k}(t)\right) - f\left(Y^{n_k}(s)\right) - \int_s^t L_\tau^{(n_k)} f\left(Y^{n_k}(\tau)\right)\,d\tau\right)\right. \\ &\qquad \left.\times g\left(Y^{n_k}\left(s_1\right),\ldots,Y^{n_k}\left(s_n\right)\right)\right] \\ &= 0. \end{aligned} \tag{1.133}$$

The final step in (1.133) follows because the process $\left\{Y^\ell(t) : t \geqslant 0\right\}$ satisfies the stochastic differential equation in (1.128). From Theorem 1.3 we then infer that the process

$$t \mapsto f\left(Y^\ell(t)\right) - f\left(Y^\ell(0)\right) - \int_0^t L^\ell(\tau) f\left(Y^\ell(\tau)\right)\,d\tau, \quad t \geqslant s,$$

is a martingale after time s. Here the operators $L_\tau^{(\ell)}$, $\ell \in \mathbb{N}$, $\tau \geqslant 0$, are defined by

$$L^\ell(\tau) f(s) = \frac{1}{2}\sum_{i,j=1}^d a_{i,j}^\ell(s,x)\,\frac{\partial^2 f}{\partial x_j \partial x_k}(x) + \sum_{i=1}^d b_j^\ell(s,x)\,\frac{\partial f}{\partial x_i}(x). \tag{1.134}$$

As a consequence of the above observations we see that the processes

$$\left\{f\left(\widehat{Y}(t)\right) - f\left(\widehat{Y}(0)\right) - \int_0^t L(\tau) f\left(\widehat{Y}(\tau)\right)\,d\tau : t \geqslant 0\right\}, \quad f \in C_b^2\left(\mathbb{R}^d\right), \tag{1.135}$$

are local $\widehat{\mathbb{P}}$-martingales. Finally, we define the probability $\mathbb{P}$ on the space $W = C\left([0,\infty),\mathbb{R}^d\right)$ by the equality

$$\mathbb{P}\left[\left(X\left(t_1\right),\ldots,X\left(t_n\right)\right) \in B\right] = \widehat{\mathbb{P}}\left[\left(\widehat{Y}\left(t_1\right),\ldots,\widehat{Y}\left(t_n\right)\right) \in B\right], \tag{1.136}$$

where B is a Borel subset of $\left(\mathbb{R}^d\right)^n$, and where $\leqslant t_1 < t_2 < \cdots < t_n < \infty$. From the properties (1.135) and (1.136) it follows that the processes

$$\left\{f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L_(\tau) f\left(X(\tau)\right)\,d\tau : t \geqslant 0\right\}, \quad f \in C_b^2\left(\mathbb{R}^d\right), \tag{1.137}$$

are local $\mathbb{P}$-martingales, and the standard filtration on W. An application of item (iii) in Theorem 1.3 then yields the desired result.

This completes an outline of the proof of Theorem 1.12. □

1.2 Stochastic differential equations in the infinite-dimensional setting

In order to have a strong motivation for writing the present book we need the Hilbert space version of §1.1. In other words we need to prove the results of §1.1 for cylindrical Brownian motion. These stochastic differential equations are closely related to Partial Differential Equations (PSDE's): see e.g. [Cerrai (2001)], Seidler [Seidler (1997)], [Maslowski and Seidler (1998)], [Goldys and van Neerven (2003)], [Goldys and Maslowski (2001)], [Da Prato and Zabczyk (1992a, 1996)]. In this infinite-dimensional setting we put

$$Q(\tau,t)f(x) = \mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right] = \mathbb{E}\left[f\left(X^{\tau,x}(t)\right)\right] \tag{1.138}$$

where $X(t)$ is a unique weak solution to the equation (compare with (1.23))

$$X(t) = x + \int_\tau^t b\left(s, X(s)\right) ds + \int_\tau^t \sigma\left(s, X(s)\right) dW_H(s), \quad t \geqslant \tau. \tag{1.139}$$

If we have unique strong solutions, then we usually write $X^{\tau,x}(t)$, $t \geqslant \tau$, instead of $X(t)$. This means in case we have unique weak solutions the uniqueness is reflected in the measure $\mathbb{P}_{\tau,x}$, and if the paths are unique the uniqueness is reflected in the path, and the measure is directly related to cylindrical Brownian motion in the real Hilbert space H. In (1.139) the process $t \mapsto W_H(t)$ stands for cylindrical Brownian motion in a given Hilbert space $(H, \|\cdot\|_H)$, which is also called the Cameron-Martin Hilbert space. This Hilbert space is supposed to have a countable orthogonal basis.

Definition 1.4. Formally a cylindrical Brownian motion is a process of the form $W_H(t) = \sum_{j=1}^\infty W_{H,j}(t)e_j$, where the sequence $(e_j)_{j\in\mathbb{N}}$ is an orthonormal basis in H, and where each process $t \mapsto W_{H,j}(t)$ is a one-dimensional standard Brownian motion. The processes $t \mapsto W_{H,j_1}(t)$ and $t \mapsto W_{H,j_2}(t)$, $j_1 \neq j_2$, are $\mathbb{P}$-independent.

The following result contains a Hilbert space version of Theorem 1.2. As indicated above Theorem 1.2 is a d-dimensional version of Corollary 1.1, which is Lévy's theorem. The following theorem gives a characterization of cylindrical Brownian motion. Its proof follows that of Theorem 1.2. Let E be a Banach space and let E^* its topological dual. In the sequel E-valued process M will be called a (local) martingale if it is a (local) martingale in the weak sense, i.e. if for every $x^* \in E^*$ the process $t \mapsto \langle M(t), x^*\rangle$ is a (local) martingale.

Theorem 1.13. *Let H be an Hilbert space with a complete orthonormal system $(e_j)_{j\in\mathbb{N}}$, and let the process $t \mapsto M(t)$ be an E-valued local martingale with covariation process given by*

$$\langle \langle M(\cdot), x^* \rangle, \langle M(\cdot), y^* \rangle \rangle (t) = \int_0^t \langle \Phi(s)x^*, y^* \rangle_H \, ds, \tag{1.140}$$

where x^* and $y^* \in E^*$, and $s \mapsto \Phi(s)$ is an adapted process which attains its values in the cone of positive linear mappings from E^* to E. Let the operator valued adapted process $\chi(s) : E \to H$ be such that $\chi(s)\Phi(s)\chi(s)^* = I_H$. Put $W_H(t) = \int_0^t \chi(s)\, dM(s)$. This integral should be interpreted in Itô sense. Then the process $t \mapsto W_H(t)$ is cylindrical Brownian motion. Put $\Psi(t) = \Phi(t)\chi(t)^*$, and suppose that $\Psi(t)\chi(t) = I_E$. Then $M(t) - M(0) = \int_0^t \Psi(s)\, dW_H(s)$.

A mapping $\Phi : E^* \to E$ is called positive if $\langle \Phi x^*, x^* \rangle \geqslant 0$ for all $x^* \in E^*$.

Proof. First we calculate the covariation (process) of the processes

$$t \mapsto \langle W_H(t), e_j \rangle_H \quad \text{and} \quad t \mapsto \langle W_H(t), e_k \rangle_H \, .$$

This covariation process is given by

$$\begin{aligned} &\left\langle \left\langle \int_0^{\cdot} \chi(s) dM(s), e_j \right\rangle_H, \left\langle \int_0^{\cdot} \chi(s) dM(s), e_j \right\rangle_H \right\rangle (t) \\ &= \int_0^t \langle \Phi(s)\chi(s)^* e_j, \chi(s)^* e_k \rangle \, ds \\ &= \int_0^t \langle \chi(s)\Phi(s)\chi(s)^* e_j, e_k \rangle \, ds = t\delta_{j,k}. \end{aligned} \tag{1.141}$$

The finite-dimensional version of Lévy's theorem (see Theorem 1.2) then shows that the process $t \mapsto W_H(t)$ is cylindrical Brownian motion. In addition we have

$$\int_0^t \Psi(s)\, dW_H(s) = \int_0^t \Phi(s)\chi(s)^*\chi(s)\, dM(s) = \int_0^t dM(s) = M(t) - M(0). \tag{1.142}$$

This completes the proof of Theorem 1.13. □

The mapping $(s, x) \mapsto \sigma(s, x)$ is a mapping from the Hilbert space H to the real separable Banach space $(E, \|\cdot\|_E)$. The function $(s, x) \mapsto b(s, x)$ attains its values in E. Suppose that for every $t \in [0, T]$ the function $x \mapsto f(t, x)$ is twice continuously differentiable. Then we put

$$L(t)f(t, x) = \langle b(t, x), Df(t, x) \rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma(t, x)^* D^2 f(t, x)\sigma(t, x)\right). \tag{1.143}$$

Definition 1.5. Let $f : E \to \mathbb{C}$ be a function. The function f is called a C^1-function, if for every x, $y \in E$ the expression $\langle y, Df(x)\rangle$ which is given by

$$\langle y, Df(x)\rangle = \lim_{s\to 0} \frac{f(x+sy) - f(x)}{s}, \quad x \in E, \tag{1.144}$$

exists and if the mapping $(x, y) \mapsto \langle y, Df(x)\rangle$, $(x, y) \in E \times E$, is continuous. The derivative $Df(x)$ can be considered as an element of E^*. The function f is called a C^2-function if for every triple $(x, y_1, y_2) \in E \times E \times E$ the limit

$$\left\langle y_1, D^2 f(x) y_2\right\rangle = \lim_{s,t\to 0} \frac{f(x+sy_1+ty_2) - f(x+ty_2) - f(x+sy_1) + f(x)}{st}, \tag{1.145}$$

exists, and if the mapping $(x, y_1, y_2) \mapsto \left\langle y_1, D^2 f(x) y_2\right\rangle$, $(x, y_1, y_2) \in E \times E \times E$, is continuous. If $f : E \to \mathbb{C}$ be a C^2-function, then $D^2 f(x)$ can be interpreted as a mapping from E to E^*. More precisely, the equality in (1.145) defines such a mapping. For C^2-functions f it makes sense to write $\sigma(s,x)^* D^2 f(x) \sigma(s,x)$. For $(s,x) \in [0,T] \times E$ fixed, and f twice continuously differentiable (at x), this mapping is a linear operator from H to H.

A function $f : E \to \mathbb{C}$ is called a cylindrical function if there exists a finite number of elements $(x_1^*, \dots, x_n^*) \in (E^*)^n$ and a function $F : \mathbb{C}^n \to \mathbb{C}$ such that

$$f(x) = F\left(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle\right), \quad x \in E. \tag{1.146}$$

If everywhere $\mathbb{C}$ is replaced with $\mathbb{R}$, then f is called a real cylindrical function.

The derivatives in (1.144) and (1.145) are called Gâteaux derivatives of the function f, because the derivatives are taken in the weak sense. As notation we use $f \in C^1(E)$ for C^1-functions f defined on E, and $f \in C^2(E)$ for C^2-functions f defined on E.

The following proposition is left as an exercise for the reader.

Proposition 1.2. *Let f be a real cylindrical function as in (1.146). If the function F is a C^1-function defined on $\mathbb{R}^n$, then f is a C^1-function defined on E, and*

$$\langle y, Df(x)\rangle = \sum_{j=1}^{n} D_j F\left(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle\right) \left\langle y, x_j^*\right\rangle, \quad x,\ y \in E. \tag{1.147}$$

Here $D_j F(\xi) = \dfrac{\partial F(\xi)}{\partial \xi_j}$, $\xi \in \mathbb{R}^n$. *If the function* F *is a* C^2*-function defined on* $\mathbb{R}^n$, *then* f *given (1.146) is a* C^2*-function defined on* E, *and*

$$\left\langle y_1, D^2 f(x) y_2 \right\rangle = \sum_{k_1, k_2 = 1}^{n} D_{k_1} D_{k_2} F\left(\langle x, x_1^* \rangle, \dots, \langle x, x_n^* \rangle\right) \left\langle y_1, x_{k_1}^* \right\rangle \left\langle y_2, x_{k_2}^* \right\rangle \tag{1.148}$$

where $(x, y_1, y_2) \in E \times E \times E$.

In Theorem 1.14 below we assume that the mappings $\sigma(s,x) : H \to E$ are invertible in the sense that their null spaces are $\{0\}$ or that their ranges are dense in E. Moreover, we assume that there exists a family of operators $\chi(s,x) : E \to H$ and a complete orthonormal system $(e_j : j \in \mathbb{N})$ in H such that

$$\langle \sigma(s,x)^* \chi(s,x)^* e_{j_1}, \sigma(s,x)^* \chi(s,x)^* e_{j_2} \rangle_H = \delta_{j_1, j_2}, \quad j_1, j_2 \in \mathbb{N}. \tag{1.149}$$

In addition, suppose that $\chi(s,x)y = 0$, $y \in E$, implies $y = 0$. From (1.149) we see that

$$\chi(s,x)\sigma(s,x)\sigma(s,x)^* \chi(s,x)^* = I_H. \tag{1.150}$$

From (1.150) we get

$$\chi(s,x)\sigma(s,x)\sigma(s,x)^* \chi(s,x)^* \chi(s,x) = \chi(s,x), \tag{1.151}$$

and hence

$$\sigma(s,x)\sigma(s,x)^* \chi(s,x)^* \chi(s,x) = I_E. \tag{1.152}$$

From (1.152) we see that

$$\sigma(s,x)\sigma(s,x)^* \chi(s,x)^* \chi(s,x)\sigma(s,x) = \sigma(s,x). \tag{1.153}$$

Since the null space of $\sigma(s,x)$ is $\{0\}$ or its range is dense in E (1.153) implies:

$$\sigma(s,x)^* \chi(s,x)^* \chi(s,x)\sigma(s,x) = I_E. \tag{1.154}$$

Instead of the equalities in (1.149) through (1.154) the only property of the function $\chi(s,x)$ which is really required is the following:

$$\sigma(s,x)^* \chi(s,x)^* \chi(s,x)\sigma(s,x)\sigma(s,x)^* = \sigma(s,x)^*. \tag{1.155}$$

In fact on the range of $\sigma(s,x)$ we can construct the operator $\chi(s,x)$ as follows. Let $E_R(s,x)$ be the orthogonal projection on the closure of the range of the operator $\sigma(s,x)^*$ and define $\chi(s,x)\sigma(s,x)h = E_R(s,x)h$, $h \in H$. It is believed that this construction suffices to complete the proof of the implication (ii) $\Longrightarrow$ (iii) in Theorem 1.14. The following theorem is the infinite-dimensional analog of Theorem 1.3.

Theorem 1.14. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a right-continuous filtration $(\mathcal{F}_t)_{t \geqslant \tau}$. Let $\{X(t) : t \geqslant \tau\}$ be an E-valued continuous adapted process. Then the following assertions are equivalent:*

(i) For every cylindrical function $f \in C^2(E)$ the process

$$t \mapsto f(X(t)) - f(X(\tau)) - \int_\tau^t L(s) f(X(s))\, ds \tag{1.156}$$

is a local $\mathbb{P}$-martingale.

(ii) For every $x^* \in E^*$ the process

$$t \mapsto \langle M(t), x^* \rangle := \langle X(t), x^* \rangle - \int_\tau^t \langle b(s, X(s)), x^* \rangle\, ds, \quad t \geqslant \tau, \tag{1.157}$$

is local martingale with covariation processes

$$t \mapsto \langle \langle M, x^* \rangle, \langle M, y^* \rangle \rangle (t) = \int_\tau^t \left\langle \sigma(s, X(s))^* x^*, \sigma(s, X(s))^* y^* \right\rangle_H ds \tag{1.158}$$

where $t \geqslant \tau$, x^*, $y^* \in E^*$.

(iii) There exists a cylindrical Brownian motion $\{W_H(t) : t \geqslant \tau\}$ on some extension of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ starting at 0 at time τ such that

$$X(t) = X(\tau) + \int_\tau^t b(s, X(s))\, ds + \int_0^t \sigma(s, X(s))\, dW_H(s), \quad t \geqslant \tau. \tag{1.159}$$

Proof. We will give a proof for $\tau = 0$; the proof for general $\tau > 0$ is exactly the same.

(i) $\Longrightarrow$ (ii). Let $x^* \in E^*$, and put $f(x) = \langle x, x^* \rangle$, $x \in E$. Then the function f is linear, and so $D^2 f(x) = 0$. From (i) it follows that the process

$$t \mapsto \langle X(t), x^* \rangle - \int_0^t \langle b(s, X(s)), x^* \rangle\, ds = f(X(t)) - \int_0^t L(s) f(X(s))\, ds \tag{1.160}$$

is a local martingale. Let x^* and y^* belong to E^*. We will also show that the process

$$\left(\langle X(t), x^* \rangle - \int_0^t \langle b(s, X(s)), x^* \rangle\, ds \right) \left(\langle X(t), y^* \rangle - \int_0^t \langle b(s, X(s)), y^* \rangle\, ds \right)$$
$$- \int_0^t \left\langle \sigma(s, X(s))^* x^*, \sigma(s, X(s))^* y^* \right\rangle ds \tag{1.161}$$

is a local martingale. Once it is proved that the process in (1.161) is a local martingale, then assertion (ii) follows from semi-martingale theory.

So let us prove (1.161). Put $f(x) = \langle x, x^*\rangle \langle x, y^*\rangle$. Then we have

$$\begin{aligned} \langle y, Df(x)\rangle &= \langle y, x^*\rangle \langle x, y^*\rangle + \langle x, x^*\rangle \langle y, y^*\rangle , \\ \left\langle y, D^2 f(x) z\right\rangle &= \langle z, x^*\rangle \langle y, y^*\rangle + \langle y, x^*\rangle \langle z, y^*\rangle , \end{aligned}$$

and

$$\operatorname{Tr}\left(\sigma\left(s,x\right)^* D^2 f(x)\sigma\left(s,x\right)\right) = 2\left\langle \sigma\left(s,x\right)^* x^*, \sigma\left(s,x\right)^* y^*\right\rangle_H . \tag{1.162}$$

From the equalities in (1.162) we obtain the equality:

$$\begin{aligned} &\langle X(t), x^*\rangle \langle X(t), y^*\rangle - \int_0^t \left\langle \sigma\left(s, X(s)\right)^* x^*, \sigma\left(s, X(s)\right)^* y^*\right\rangle_H ds \\ &- \int_0^t \left(\langle b\left(s, X(s)\right), x^*\rangle \langle X(s), y^*\rangle + \langle X(s), x^*\rangle \langle b\left(s, X(s)\right), y^*\rangle\right) ds \\ &= f\left(X(t)\right) - \int_0^t L(s) f\left(X(s)\right) ds. \end{aligned} \tag{1.163}$$

From assertion (i) it follows that the process in the right-hand side of (1.163) is a local martingale. As a consequence the process in the left-hand side of (1.163) is a local martingale as well. For brevity we write

$$\begin{aligned} M_{x^*}(s) &= \left\langle X(s) - \int_0^s b\left(\tau, X(\tau)\right) d\tau, x^*\right\rangle , \\ M_{y^*}(s) &= \left\langle X(s) - \int_0^s b\left(\tau, X(\tau)\right) d\tau, y^*\right\rangle \\ &M_{x^*,y^*}(s) \\ &= \langle X(s), x^*\rangle \langle X(s), y^*\rangle \\ &\quad - \int_0^s \left(\langle b\left(\tau, X(\tau)\right), x^*\rangle \langle X(\tau), y^*\rangle + \langle X(\tau), x^*\rangle \langle b\left(\tau, X(\tau)\right), y^*\rangle\right) d\tau \\ &\quad + \int_0^s \left\langle \sigma\left(\tau, X(\tau)\right)^* x^*, \sigma\left(\tau, X(\tau)\right)^* y^*\right\rangle_H d\tau. \end{aligned} \tag{1.164}$$

Then the processes M_{x^*}, M_{y^*} and M_{x^*,y^*} are local martingales: see (1.160) and (1.163). Moreover, a calculation shows that:

$$\begin{aligned} &M_{x^*}(t) M_{y^*}(t) - \int_0^t \left\langle \sigma\left(s, X(s)\right)^* x^*, \sigma\left(s, X(s)\right)^* y^*\right\rangle_H ds \\ &= M_{x^*,y^*}(t) - \int_0^t \langle b\left(s, X(s)\right), x^*\rangle \left(M_{y^*}(t) - M_{y^*}(s)\right) ds \end{aligned}$$

$$- \int_0^t \left(M_{x^*}(t) - M_{x^*}(s)\right) \langle b(s, X(s)), y^* \rangle \, ds. \tag{1.165}$$

It is readily verified that the processes

$$\int_0^t \langle b(s, X(s)), x^* \rangle \left(M_{y^*}(t) - M_{y^*}(s)\right) ds \quad \text{and}$$
$$\int_0^t \left(M_{x^*}(t) - M_{x^*}(s)\right) \langle b(s, X(s)), y^* \rangle \, ds \tag{1.166}$$

are local martingales. It follows that the process in (1.161) is a local martingale. So that the covariation process $\langle M_{x^*}, M_{y^*} \rangle$ is given by

$$\langle M_{x^*}, M_{y^*} \rangle (t) = \int_0^t \left\langle \sigma(s, X(s))^* x^*, \sigma(s, X(s))^* y^* \right\rangle_H ds.$$

(ii) $\Longrightarrow$ (iii). Let the family of operators $\chi(s, X(s))$, $s \in [0, T]$, be such that

$$\left\langle \sigma(s, X(s))^* \chi(s, X(s))^* e_{j_1}, \sigma(s, X(s))^* \chi(s, X(s))^* e_{j_1} \right\rangle_H = \delta_{j_1, j_2}, \tag{1.167}$$

j_1, $j_2 \in \mathbb{N}$. Here $(e_j : j \in \mathbb{N})$ is a complete orthonormal system in H: compare with (1.149). Put

$$W_H(t) = \int_0^t \sigma(s, X(s))^* \chi(s, X(s))^* \chi(s, X(s)) \, dM(s) \tag{1.168}$$

where $M(s) = X(s) - \int_0^s b(\tau, X(\tau)) \, d\tau$. Then, employing (ii), the covariation process of the processes

$$t \mapsto \langle W_H(t), e_{j_1} \rangle_H = \int_0^t \left\langle dM(s), \chi(s, X(s))^* \chi(s, X(s)) \sigma(s, X(s)) e_{j_1} \right\rangle$$

and

$$t \mapsto \langle W_H(t), e_{j_2} \rangle_H = \int_0^t \left\langle dM(s), \chi(s, X(s))^* \chi(s, X(s)) \sigma(s, X(s)) e_{j_2} \right\rangle \tag{1.169}$$

is given by

$$\begin{aligned}
&\left\langle \langle W_H(\cdot), e_{j_1} \rangle_H, \langle W_H(\cdot), e_{j_2} \rangle_H \right\rangle (t) \\
&= \int_0^t \left\langle \sigma(s, X(s))^* \chi(s, X(s))^* \chi(s, X(s)) \sigma(s, X(s)) e_{j_1}, \right. \\
&\qquad \left. \sigma(s, X(s))^* \chi(s, X(s))^* \chi(s, X(s)) \sigma(s, X(s)) e_{j_2} \right\rangle_H ds \\
&= t\delta_{j_1, j_2}.
\end{aligned} \tag{1.170}$$

Here, as elsewhere, $\delta_{j_1,j_2} = 1$ when $j_1 = j_2$, and 0 otherwise. In the final equality in (1.170) we employed (1.154). From Lévy's theorem and (1.170) it follows that the process $t \mapsto W_H(t)$ is a cylindrical Brownian motion: see Theorem 1.13. In addition, from (1.168) in combination with (1.152) we infer:

$$\begin{aligned}
&\int_0^t \sigma(s, X(s))\, dW_H(s) \\
&= \int_0^t \sigma(s, X(s))\, \sigma(s, X(s))^* \chi(s, X(s))^* \chi(s, X(s))\; dM(s) \\
&= \int_0^t dM(s) = M(t) - M(0) = X(t) - X(0) - \int_0^t b(s, X(s))\; ds. \quad (1.171)
\end{aligned}$$

The equality in (1.171) completes the proof of the implication (ii) $\Longrightarrow$ (iii) in case the identities in (1.149) through (1.154) are assumed. If $\chi(s, x)$, $s \geqslant \tau$, only satisfies the equality in (1.155), then we proceed as follows. As in the proof of the implication (ii) $\Longrightarrow$ (iii) of Theorem 1.3 we take the standard extension $\left(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}\right)$ of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. On this extension we take a cylindrical Brownian motion $\{W_H'(t) : t \geqslant \tau\}$ which is $\widehat{\mathbb{P}}$-independent of the local martingale $M(t)$, $t \geqslant \tau$. Then instead of the definition of (1.168) we take

$$\begin{aligned}
W_H(t) = &\int_0^t \sigma(s, X(s))^* \chi(s, X(s))^* \chi(s, X(s))\; dM(s) \quad (1.172) \\
&+ \int_0^t \left(I - \sigma(s, X(s))^* \chi(s, X(s))^* \chi(s, X(s))\, \sigma(s, X(s))\right) dW_H'(s), \quad (1.173)
\end{aligned}$$

where $M(s) = X(s) - \int_0^s b(\tau, X(\tau))\; d\tau$. From (1.155) we also infer by taking adjoints that

$$\sigma(s, x)\sigma(s, x)^* \chi(s, x)^* \chi(s, x)\sigma(s, x) = \sigma(s, x). \quad (1.174)$$

We will show that the process $\{W_H(t) : t \geqslant 0\}$ is a cylindrical Brownian motion and that $M(t) = \int_0^t \sigma(s, X(s))\, dW_H(s)$. For brevity we write $\sigma(s) = \sigma(s, X(s))$, $\chi(s) = \chi(s, X(s))$, $E_R(s) = \sigma(s)^* \chi(s)^* \chi(s)\, \sigma(s)$, and $E_N(s) = I - E_R(s)$. Then $\sigma(s)E_N(s) = 0$, and $E_R(s)\sigma(s)^* = \sigma(s)^*$. We begin by proving that $\{W_H(t) : t \geqslant 0\}$ is a cylindrical Brownian motion. We will invoke Theorem 1.1 to establish this result. Let e_{j_1} and e_{j_2} be two orthogonal vectors in the Hilbert space H. Then we have

$$\left\langle \langle W_H(\cdot), e_{j_1}\rangle_H, \langle W_H(\cdot), e_{j_2}\rangle_H \right\rangle(t)$$

$$
\begin{aligned}
&= \left\langle \int_0^{\cdot} \left\langle dM(s), \chi(s)^*\chi(s)\sigma(s)e_{j_1}\right\rangle, \int_0^{\cdot} \left\langle dM(s), \chi(s)^*\chi(s)\sigma(s)e_{j_2}\right\rangle \right\rangle (t) \\
&\quad + \left\langle \int_0^{\cdot} \left\langle dM(s), \chi(s)^*\chi(s)\sigma(s)e_{j_1}\right\rangle, \right. \\
&\qquad \left. \int_0^{\cdot} \left\langle dW_H'(s), \left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) e_{j_2}\right\rangle \right\rangle (t) \\
&\quad + \left\langle \int_0^{\cdot} \left\langle dM(s), \chi(s)^*\chi(s)\sigma(s)e_{j_2}\right\rangle, \right. \\
&\qquad \left. \int_0^{\cdot} \left\langle dW_H'(s), \left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) e_{j_1}\right\rangle \right\rangle (t) \\
&\quad + \left\langle \int_0^{\cdot} \left\langle dW_H'(s), \left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) e_{j_1}\right\rangle, \right. \\
&\qquad \left. \int_0^{\cdot} \left\langle dW_H'(s), \left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) e_{j_2}\right\rangle \right\rangle (t)
\end{aligned}
$$

(employ the properties of M as set out in Assertion (ii); moreover, M and W_H are $\widehat{\mathbb{P}}$-independent)

$$
\begin{aligned}
&= \int_0^t \left\langle \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)e_{j_1}, \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)e_{j_2}\right\rangle_H \, ds \\
&\quad + \int_0^t \left\langle \left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) e_{j_1}, \left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) e_{j_2}\right\rangle_H \, ds \\
&= \int_0^t \left\langle E_R(s)e_{j_1}, E_R(s)e_{j_2}\right\rangle_H \, ds + \int_0^t \left\langle E_N(s)e_{j_1}, E_N(s)e_{j_2}\right\rangle_H \, ds
\end{aligned}
$$

(the operators $E_R(s)$ and $E_N(s)$ are orthogonal projections in the Hilbert space H)

$$
= \int_0^t \left\langle E_R(s)e_{j_1}, e_{j_2}\right\rangle_H \, ds + \int_0^t \left\langle E_N(s)e_{j_1}, e_{j_2}\right\rangle_H \, ds
$$

(the identity $E_R(s) + E_N(s) = I$ holds)

$$
= \int_0^t \left\langle e_{j_1}, e_{j_2}\right\rangle_H \, ds = t\delta_{j_1,j_2}. \tag{1.175}
$$

From Theorem 1.1 it follows that the process $\{W_H(t) : t \geqslant 0\}$ is cylindrical Brownian motion. In addition, we have

$$
M(t) - \int_0^t \sigma(s)\, dW_H(s)
$$

$$= M(t) - \int_0^t \sigma(s)\sigma(s)^*\chi(s)^*\chi(s)\, dM(s)$$

$$- \int_0^t \sigma(s)\left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right) dW_H'(s)$$

$$= \int_0^t \left(I - \sigma(s)\sigma(s)^*\chi(s)^*\chi(s)\right) dM(s). \tag{1.176}$$

In order to prove that the local martingale in (1.176) is zero we calculate its covariation process. Let x^* and y^* be members of E^*. Then we have

$$\left\langle\left\langle\int_0^\cdot \left(I - \sigma(s)\sigma(s)^*\chi(s)^*\chi(s)\right) dM(s), x^*\right\rangle,\right.$$

$$\left.\left\langle\int_0^\cdot \left(I - \sigma(s)\sigma(s)^*\chi(s)^*\chi(s)\right) dM(s), y^*\right\rangle\right\rangle(t)$$

$$= \left\langle\int_0^\cdot \langle dM(s), \left(I - \chi(s)^*\chi(s)\sigma(s)\sigma(s)^*\right) x^*\rangle,\right.$$

$$\left.\int_0^\cdot \langle dM(s), \left(I - \chi(s)^*\chi(s)\sigma(s)\sigma(s)^*\right) y^*\rangle\right\rangle(t)$$

$$= \int_0^t \langle \sigma(s)^*\left(I - \chi(s)^*\chi(s)\sigma(s)\sigma(s)^*\right) x^*,$$

$$\sigma(s)^*\left(I - \chi(s)^*\chi(s)\sigma(s)\sigma(s)^*\right) y^*\rangle_H\, ds$$

$$= \int_0^t \langle\left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right)\sigma(s)^*x^*,$$

$$\left(I - \sigma(s)^*\chi(s)^*\chi(s)\sigma(s)\right)\sigma(s)^*y^*\rangle_H\, ds$$

$$= \int_0^t \langle E_N(s)\sigma(s)^*x^*, E_N(s)\sigma(s)^*y^*\rangle_H\, ds = 0. \tag{1.177}$$

In the final equality of (1.177) we used the identity

$$E_N(s)\sigma(s)^* = \left(I - E_R(s)\right)\sigma(s)^* = 0.$$

From (1.177) we infer that the covariation of the local martingale $M(t) - \int_0^t \sigma(s)\, dW_H(s)$ vanishes. Consequently, $M(t) - \int_0^t \sigma(s)\, dW_H(s) = 0$. This shows the implication (ii) $\Longrightarrow$ (iii) of Theorem 1.14.

(iii) $\Longrightarrow$ (i). Let $f : E \to \mathbb{C}$ be a C^2-function. From Itô's formula (see equality (1.196) in Proposition 1.3 below with $C(t,\tau) = I$ and $A(t) = 0$), and assertion (iii), we get

$$f(X(t)) - f(X(0)) - \int_0^t L(s) f(X(s))\, ds$$

$$= \int_0^t \langle Df(X(s)), dX(s)\rangle + \frac{1}{2}\int_0^t \operatorname{Tr}\left(\sigma(s,X(s))^* D^2 f(X(s))\,\sigma(s,X(s))\right) ds$$
$$- \int_0^t L(s) f(X(s))\,ds$$
$$= \int_0^t \langle b(s,X(s)), Df(X(s))\rangle\, ds + \int_0^t \langle \sigma(s,X(s))\, dW_H(s), Df(X(s))\rangle$$
$$+ \frac{1}{2}\int_0^t \operatorname{Tr}\left(\sigma(s,X(s))^* D^2 f(X(s))\,\sigma(s,X(s))\right) ds - \int_0^t L(s) f(X(s))\, ds$$
$$= \int_0^t \langle \sigma(s,X(s))\, dW_H(s), Df(X(s))\rangle . \tag{1.178}$$

The stochastic integral in (1.178) represents a local martingale. This proves the implication (iii) $\Longrightarrow$ (i).

All this completes the proof of Theorem 1.14. □

The following definition is the infinite-dimensional analog of Definition 1.2.

Definition 1.6. The equation in (1.180) in Corollary 1.6 is said to have unique weak solutions, also called unique distributional solutions, provided that the finite-dimensional distributions of the process $X(t)$ which satisfy (1.180) do not depend on the particular cylindrical Brownian motion $W_H(t)$ which occurs in (1.180). This is the case if and only if for any pair of cylindrical Brownian motions

$$\{(W_H(t) : t \geqslant 0), (\Omega, \mathcal{F}, \mathbb{P})\} \quad \text{and} \quad \left\{(W_{H'}(t) : t \geqslant 0), \left(\Omega', \mathcal{F}', \mathbb{P}'\right)\right\}$$

and any pair of adapted processes $\{X(t) : t \geqslant 0\}$ and $\{X'(t) : t \geqslant 0\}$ for which

$$X(t) = x + \int_0^t \sigma(s,X(s))\, dW_H(s) + \int_0^t b(s,X(s))\, ds \quad \text{and}$$
$$X'(t) = x + \int_0^t \sigma\left(s,X'(s)\right) dW_{H'}(s) + \int_0^t b\left(s,X'(s)\right) ds$$

it follows that the finite-dimensional distributions of the process $\{X(t) : t \geqslant 0\}$ relative to $\mathbb{P}$ coincide with the finite-dimensional distributions of the process $\{X'(t) : t \geqslant 0\}$ relative to $\mathbb{P}'$.

The following corollary easily follows from Theorem 1.14. It is an infinite-dimensional analog of Corollary 1.2. In the infinite-dimensional setting it establishes a close relationship between unique weak solutions to stochastic differential equations and unique solutions to the martingale problem. The result serves as one of the main motivations to write the present book.

Corollary 1.6. *Let the notation and hypotheses be as in Theorem 1.14. In particular, suppose that (1.155) is satisfied. Put* $\Omega = C([0,\infty), E)$, *and* $X(t)(\omega) = \omega(t)$, $t \geqslant 0$, $\omega \in \Omega$. *Fix* $x \in E$. *Then the following assertions are equivalent:*

(i) *There exists a unique probability measure* $\mathbb{P}$ *on* $\mathcal{F}$ *such that* $\mathbb{P}[X(\tau) = x] = 1$, *and the process*

$$f(X(t)) - f(X(\tau)) - \int_\tau^t L(s) f(X(s))\, ds \tag{1.179}$$

is a local $\mathbb{P}$*-martingale for all cylindrical* C^2*-functions* f.

(ii) *The stochastic integral equation*

$$X(t) = x + \int_\tau^t \sigma(s, X(s))\, dW_H(s) + \int_\tau^t b(s, X(s))\, ds \tag{1.180}$$

has unique weak solutions.

In the infinite-dimensional setting we have the following version of Girsanov's theorem. Let $(E, \|\cdot\|)$ be a Banach space, and let $(H, \|\cdot\|_H)$ be a separable Hilbert space and let $W_H(t)$, $0 \leqslant t \leqslant T$, be a cylindrical Brownian motion in H. Let $(s,y) \mapsto b(s,y)$ be an E-valued weakly continuous function on $[0,T] \times E$, $(s,y) \mapsto c(s,y)$ be an H-valued weakly continuous function on $[0,T] \times E$, and let $(s,y) \mapsto \sigma(s,y)$ be an $L(H,E)$-valued function which is continuous for the weak operator topology, i.e. for every $z \in H$, and $x^* \in E^*$ the function $(s,y) \mapsto \langle \sigma(s,y) z, x^* \rangle$ is a continuous as a function from $[0,T] \times E$ to $\mathbb{R}$. The symbol $L(H,E)$ denotes the space of all continuous linear operators from H to E. The function $(s,y) \mapsto c_1(s,y)$ attains its values in E^*, it is such that $c(s,y) = \sigma(s,y)^* c(s,y)$, and such that the function $(s,y) \mapsto \langle x, c_1(s,y) \rangle$ is continuous for every $x \in E$.

Theorem 1.15. *Fix* $T > 0$, *and let the functions*

$$b(s,y), \quad \sigma(s,y), \quad c(s,y), \quad \text{and} \quad c_1(s,y), \quad 0 \leqslant s \leqslant T,$$

be weakly continuous vector or matrix functions such that

$$c(s,y) = \sigma(s,y)^* c_1(s,y), \quad 0 \leqslant s \leqslant T,\ y \in E.$$

Suppose that the equation

$$X(t) = x + \int_0^t \sigma(s, X(s))\, dW_H(s) + \int_0^t b(s, X(s))\, ds, \quad t \in [0,T], \tag{1.181}$$

possesses unique weak solutions on the interval $[0,T]$*: compare with (1.36).*

Uniqueness. Let weak solutions to the following stochastic differential equation exist (compare with (1.37)):

$$Y(t) = x + \int_0^t \sigma(s, X(s))\, dW_H(s) + \int_0^t \sigma(s, Y(s))\, c(s, Y(s))\, ds + \int_0^t b(s, X(s))\, ds, \tag{1.182}$$

$t \in [0,T]$. Then they are unique in the sense as explained next. In fact, let the couple $(Y(s), W_H(s))$, $0 \leqslant s \leqslant t$, be a solution to the equation in (1.182) with the property that the local martingale $M'(t)$ given by

$$M'(t) = \exp\left(-\int_0^t c(s, Y(s))\, dW_H(s) - \frac{1}{2}\int_0^t \|c(s, Y(s))\|_H^2\, ds\right). \tag{1.183}$$

satisfies $\mathbb{E}[M'(t)] = 1$. Then the finite-dimensional distributions of the process $Y(s)$, $0 \leqslant s \leqslant t$, are given by the Girsanov or Cameron-Martin transform:

$$\mathbb{E}[f(Y(t_1), \dots, Y(t_n))] = \mathbb{E}[M(t) f(X(t_1), \dots, X(t_n))], \tag{1.184}$$

$t \geqslant t_n > \cdots > t_1 \geqslant 0$, where $f : E^n \to \mathbb{R}$ is an arbitrary bounded Borel measurable function. The (local) martingale $M(s)$ is given by

$$M(s) = \exp\left(\int_0^s c(\tau, X(\tau))\, dW_H(s) - \frac{1}{2}\int_0^t \|c(\tau, X(\tau))\|_H^2\, d\tau\right). \tag{1.185}$$

If equation (1.182) has a solution such that $\mathbb{E}[M'(t)] = 1$, then necessarily $\mathbb{E}[M(t)] = 1$, and so $s \mapsto M(s)$ is a martingale on the interval $[0,t]$.

Existence. *Conversely, let the process $s \mapsto (X(s), W_H(s))$ be a solution to the equation in (1.181). Suppose that the local martingale $s \mapsto M(s)$, defined as in (1.185) is a martingale, i.e. suppose that $\mathbb{E}[M(t)] = 1$. Then there exists a couple $\left(\widetilde{Y}(s), \widetilde{W}_H(s)\right)$, $0 \leqslant s \leqslant t$, where $s \mapsto \widetilde{W}_H(s)$, $0 \leqslant s \leqslant t$, is a cylindrical Brownian motion on a probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ such that*

$$\widetilde{Y}(s) = x + \int_0^s \sigma\left(\tau, \widetilde{Y}(s)\right) d\widetilde{W}_H(s) + \int_0^s \sigma\left(\tau, \widetilde{Y}(s)\right) c\left(\tau, \widetilde{Y}(\tau)\right) d\tau + \int_0^s b\left(\tau, \widetilde{Y}(s)\right) d\tau, \tag{1.186}$$

and such that

$$\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s, \widetilde{Y}(s)\right) d\widetilde{W}_H(s) - \frac{1}{2}\int_0^t \left\|c\left(s, \widetilde{Y}(s)\right)\right\|_H^2 ds\right)\right] = 1. \tag{1.187}$$

The proof of Theorem 1.15 can be patterned after the proof of Theorem 1.4; it will be left as an exercise for the reader.

The following result should be compared with the equalities in (5.55). The definition in (1.188) is the same as the one in (1.138). The operator $\widetilde{L}(s)$ is the same as the one in (1.143).

Theorem 1.16. *Suppose that equation in (1.180) possesses unique weak solutions. Put*

$$Q(\tau, t)\varphi(x) = \mathbb{E}_{\tau,x}\left[\varphi\left(X(t)\right)\right], \quad \varphi \in C_b(E), \tag{1.188}$$

where the expectation $\mathbb{E}_{\tau,x}$ corresponds to the measure $\mathbb{P} = \mathbb{P}_{\tau,x}$ obtained in item (i) of Corollary 1.6. Define the operators $\widetilde{L}(s)$, $s \in [0,T]$, as the pointwise limits

$$\widetilde{L}(s)\varphi(x) = \lim_{t\downarrow s} \frac{Q(s,t)\varphi(x) - \varphi(x)}{t-s}, \quad s \in [0,T],\ x \in E, \tag{1.189}$$

and suppose that $\varphi \in C_b(E)$ is chosen in such a way that the function $(s,x) \mapsto \widetilde{L}(s)\varphi(x)$ is continuous. Then the following equalities hold:

$$\frac{\partial}{\partial s}Q(s,t)\varphi(x) = -\widetilde{L}(s)Q\left(s,t\right)\varphi(x), \text{ and } \frac{\partial}{\partial t}Q(s,t)\varphi(x) = Q\left(s,t\right)\widetilde{L}(t)\varphi(x). \tag{1.190}$$

In the first equality (1.190) it is assumed that the function φ is chosen in such a way that the pointwise derivative with respect to s exists. In the second equality it is assumed that the function $(\rho, y) \mapsto \widetilde{L}(\rho)\varphi(y)$ is bounded and continuous. In fact the operator $\widetilde{L}(s)$ is a linear extension of the operator $L(s)$ depicted in (1.143).

Proof. Suppose $0 < s < t < T$, and taking $h > 0$ small enough. Using the propagator property of the family $\{Q(s,t) :\ 0 \leqslant s \leqslant t \leqslant T\}$ yields the equalities

$$\begin{aligned} \frac{Q(s-h,t)\varphi - Q(s,t)\varphi}{-h} &= \frac{Q(s-h,s) - I}{-h}Q(s,t)\varphi \quad \text{and} \\ \frac{Q(s,t+h)\varphi - Q(s,t)\varphi}{h} &= Q(s,t)\frac{Q(t,t+h) - I}{h}\varphi. \end{aligned} \tag{1.191}$$

By the assumptions on the function φ the equalities in (1.190) follow. Next let φ be such that the process

$$\varphi\left(X(t)\right) - \varphi\left(X(s)\right) - \int_s^t L(s) f\left(X(\rho)\right)\, d\rho \tag{1.192}$$

is a $\mathbb{P}_{s,x}$-martingale. Then by taking $\mathbb{P}_{s,x}$-expectations in (1.192) we get

$$Q(s,t)\varphi(x) - \varphi(x) = \int_s^t Q(s,\rho)L(\rho)\varphi(x)\, d\rho. \tag{1.193}$$

From (1.193) we see that $\widetilde{L}(s)$ extends $L(s)$.

This completes the proof of Theorem 1.16. $\square$

Again we can discuss solutions to E-valued stochastic differential equations:

$$dX(t) = A(t)X(t)\,dt + \sigma\left(t, X(t)\right)\,dW_H(t) + b\left(t, X(t)\right)\,dt, \quad X(\tau) = x,\ t \geqslant \tau. \tag{1.194}$$

In (1.194) the family of operators $\{A(t) : 0 \leqslant t \leqslant T\}$ generates a forward propagator

$$\{C(t,\tau) : 0 \leqslant \tau \leqslant t \leqslant T\}$$

in the Banach space E. This means that $C(t,s)C(s,\tau) = C(t,\tau)$, $C(t,t) = I$, and

$$A(t)x = \lim_{h \downarrow 0} \frac{C(t+h,t)x - C(t,t)x}{h}, \quad x \in D(L(t)).$$

Then the integrated version of (1.194) reads as follows:

$$\begin{aligned} X^{\tau,x}(t) &= C(t,\tau)x + \int_\tau^t C(t,\rho)\,\sigma\left(\rho, X^{\tau,x}(\rho)\right)\,dW_H(\rho) \\ &\quad + \int_\tau^t C(t,\rho)b\left(\rho, X^{\tau,x}(\rho)\right)\,d\rho. \end{aligned} \tag{1.195}$$

Next we formulate and prove a version of Itô's formula in the infinite-dimensional setting. For related notions and results see e.g. [Krylov and Rozovskii (2007)].

Proposition 1.3. *Let the function f be such that for every $(s,x) \in [\tau,T] \times E$ the operator $\sigma(s,x)^*C(t,s)^*D^2f(x)C(t,s)\sigma(s,x)$ is a trace class operator for all $t \in [s,T]$. Let the process $X^{\tau,x}(t)$ be a solution to (1.195). Then the following equality holds $\mathbb{P}$-almost surely:*

$$\begin{aligned} &f\left(X^{\tau,x}(t)\right) - f(x) \\ &= \int_\tau^t \left\langle dX^{\tau,x}(\rho), Df\left(X^{\tau,x}(\rho)\right)\right\rangle \qquad (1.196) \\ &\quad + \frac{1}{2}\int_\tau^t \operatorname{Tr}\left(\sigma\left(\rho, X^{\tau,x}(\rho)\right)^* C(t,\rho)^* D^2 f\left(X^{\tau,x}(\rho)\right) C(t,\rho)\sigma\left(\rho, X^{\tau,x}(\rho)\right)\right)\,d\rho \\ &= \int_\tau^t \left\langle A(\rho)X^{\tau,x}(\rho) + b\left(\rho, X^{\tau,x}(\rho)\right), Df\left(X^{\tau,x}(\rho)\right)\right\rangle\,d\rho \\ &\quad + \int_\tau^t \left\langle \sigma\left(\rho, X^{\tau,x}(\rho)\right)\,dW_H(\rho), Df\left(X^{\tau,x}(\rho)\right)\right\rangle \\ &\quad + \frac{1}{2}\int_\tau^t \operatorname{Tr}\left(\sigma\left(\rho, X^{\tau,x}(\rho)\right)^* C(t,\rho)^* D^2 f\left(X^{\tau,x}(\rho)\right) C(t,\rho)\sigma\left(\rho, X^{\tau,x}(\rho)\right)\right)\,d\rho. \end{aligned}$$

Proof. By a general approximation argument it suffices to take $f : E \to \mathbb{R}$ of the form $f(x) = F(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle)$, $x \in E$, $x_1^*, \dots, x_n^* \in E^*$. Then we have

$$\langle y, Df(x)\rangle = \sum_{j=1}^{n} D_j F(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle)\left\langle y, x_j^*\right\rangle, \quad \text{and}$$

$$\left\langle y_1, D^2 f(x) y_2\right\rangle = \sum_{k_1,k_2=1}^{n} D_{k_1} D_{k_2} F(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle)\left\langle y_1, x_{k_1}^*\right\rangle\left\langle y_2, x_{k_2}^*\right\rangle \tag{1.197}$$

where y, y_1, $y_2 \in E$. Let $(e_j)_{j\in\mathbb{N}}$ be an orthonormal basis in H. From (1.197) we infer:

$$\begin{aligned}
&\operatorname{Tr}\left(\sigma(s,x)^* C(t,s)^* D^2 f(x) C(t,s)\sigma(s,x)\right) \\
&= \sum_{j=1}^{\infty}\left\langle C(t,s)\sigma(s,x)e_j, D^2 f(x) C(t,s)\sigma(s,x)e_j\right\rangle \\
&= \sum_{j=1}^{\infty}\sum_{k_1,k_2=1}^{\infty} D_{k_1} D_{k_2} F(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle) \\
&\quad \times \left\langle C(t,s)\sigma(s,x)e_j, x_{k_1}^*\right\rangle\left\langle C(t,s)\sigma(s,x)e_j, x_{k_2}^*\right\rangle \\
&= \sum_{k_1,k_2=1}^{\infty} D_{k_1} D_{k_2} F(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle) \\
&\quad \times \left\langle \sigma(s,x)^* C(t,s)^* x_{k_1}^*, \sigma(s,x)^* C(t,s)^* x_{k_2}^*\right\rangle_H .
\end{aligned} \tag{1.198}$$

From the finite-dimensional Itô formula we obtain:

$$\begin{aligned}
&f\left(X^{\tau,x}(t)\right) - f(x) \\
&= \int_\tau^t \sum_{j=1}^{n} d\left\langle X^{\tau,x}(\rho), x_j^*\right\rangle D_j F\left(\left\langle X^{\tau,x}(\rho), x_1^*\right\rangle, \dots, \left\langle X^{\tau,x}(\rho), x_n^*\right\rangle\right) \\
&\quad + \frac{1}{2}\sum_{k_1,k_2=1}^{n}\int_\tau^t D_{k_1} D_{k_2} F\left(\left\langle X^{\tau,x}(\rho), x_1^*\right\rangle, \dots, \left\langle X^{\tau,x}(\rho), x_n^*\right\rangle\right) \\
&\qquad d\left\langle\left\langle X^{\tau,x}(\cdot), x_{k_1}^*\right\rangle, \left\langle X^{\tau,x}(\cdot), x_{k_2}^*\right\rangle\right\rangle(\rho) \\
&= \int_\tau^t \sum_{j=1}^{n}\left\langle A(\rho)X^{\tau,x}(\rho), x_j^*\right\rangle D_j F\left(\left\langle X^{\tau,x}(\rho), x_1^*\right\rangle, \dots, \left\langle X^{\tau,x}(\rho), x_n^*\right\rangle\right) d\rho \\
&\quad + \int_\tau^t \sum_{j=1}^{n}\left\langle b\left(\rho, X^{\tau,x}(\rho)\right), x_j^*\right\rangle D_j F\left(\left\langle X^{\tau,x}(\rho), x_1^*\right\rangle, \dots, \left\langle X^{\tau,x}(\rho), x_n^*\right\rangle\right) d\rho
\end{aligned}$$

$$+\frac{1}{2}\sum_{k_1,k_2=1}^{n}\int_\tau^t D_{k_1}D_{k_2}F\left(\langle X^{\tau,x}(\rho),x_1^*\rangle,\dots,\langle X^{\tau,x}(\rho),x_n^*\rangle\right)$$
$$d\left\langle\left\langle\int_\tau^\cdot C(t,s)\sigma\left(s,X^{\tau,x}(s)\right)\,dW_H(s),x_{k_1}^*\right\rangle,\right.$$
$$\left.\left\langle\int_\tau^\cdot C(t,s)\sigma\left(s,X^{\tau,x}(s)\right)dW_H(s),x_{k_2}^*\right\rangle\right\rangle(\rho)$$
$$+\int_\tau^t\sum_{j=1}^{n}\left\langle\sigma\left(\rho,X^{\tau,x}(\rho)\right)dW_H(\rho),x_j^*\right\rangle$$
$$D_jF\left(\langle X^{\tau,x}(\rho),x_1^*\rangle,\dots,\langle X^{\tau,x}(\rho),x_n^*\rangle\right)$$

(employ Assertion (ii) in Theorem 1.14)

$$=\int_\tau^t\sum_{j=1}^{n}\left\langle A(\rho)X^{\tau,x}(\rho),x_j^*\right\rangle D_jF\left(\langle X^{\tau,x}(\rho),x_1^*\rangle,\dots,\langle X^{\tau,x}(\rho),x_n^*\rangle\right)\,d\rho$$
$$+\int_\tau^t\sum_{j=1}^{n}\left\langle b\left(\rho,X^{\tau,x}(\rho)\right),x_j^*\right\rangle D_jF\left(\langle X^{\tau,x}(\rho),x_1^*\rangle,\dots,\langle X^{\tau,x}(\rho),x_n^*\rangle\right)\,d\rho$$
$$+\frac{1}{2}\sum_{k_1,k_2=1}^{n}\int_\tau^t D_{k_1}D_{k_2}F\left(\langle X^{\tau,x}(\rho),x_1^*\rangle,\dots,\langle X^{\tau,x}(\rho),x_n^*\rangle\right)$$
$$\left\langle\sigma\left(\rho,X^{\tau,x}(\rho)\right)^*C(t,\rho)^*x_{k_1}^*,\sigma\left(\rho,X^{\tau,x}(\rho)\right)^*C(t,\rho)^*x_{k_2}^*\right\rangle_H\,d\rho$$
$$+\int_\tau^t\sum_{j=1}^{n}\left\langle\sigma\left(\rho,X^{\tau,x}(\rho)\right)dW_H(\rho),x_j^*\right\rangle$$
$$D_jF\left(\langle X^{\tau,x}(\rho),x_1^*\rangle,\dots,\langle X^{\tau,x}(\rho),x_n^*\rangle\right)$$

(apply the equalities in (1.197) and (1.198))

$$=\int_\tau^t\left\langle A(\rho)X^{\tau,x}(\rho),Df\left(X^{\tau,x}(\rho)\right)\right\rangle\,d\rho$$
$$+\int_\tau^t\left\langle b\left(\rho,X^{\tau,x}(\rho)\right),Df\left(X^{\tau,x}(\rho)\right)\right\rangle\,d\rho$$
$$+\frac{1}{2}\int_\tau^t\operatorname{Tr}\left(\sigma\left(\rho,X^{\tau,x}(\rho)\right)^*C(t,\rho)^*D^2f\left(X^{\tau,x}(\rho)\right)C(t,\rho)\sigma\left(\rho,X^{\tau,x}(\rho)\right)\right)\,d\rho$$
$$+\int_\tau^t\left\langle\sigma\left(\rho,X^{\tau,x}(\rho)\right)dW_H(\rho),Df\left(X^{\tau,x}(\rho)\right)\right\rangle. \tag{1.199}$$

The equality in (1.199) coincides with (1.196) in Proposition 1.3 in case $f(x)=F\left(\langle x,x_1^*\rangle,\dots,\langle x,x_n^*\rangle\right)$, $x\in E$, $x_1^*,\dots,x_n^*\in E^*$. Here F is a C^2-function defined on $\mathbb{R}^n$. An approximation argument then completes the proof of Proposition 1.3. □

Define the operators $L(t)$, $t \in [0,T]$, by

$$L(t)f(x) = \langle b(t,x), Df(x)\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma(t,x)^* D^2 f(x)\sigma(t,x)\right), \quad f \in C_b^2(E). \tag{1.200}$$

The following result is a consequence of Proposition 1.3.

Proposition 1.4. *Let the function $(t,x) \mapsto f(t,x)$ be such that $t \mapsto f(t,x)$ is once differentiable for all $x \in E$, and that $x \mapsto f(t,x)$ belongs to $C_b^2(E)$ for all $t \in [0,T]$. This time derivative is denoted by $D_1 f(t,x)$. Put*

$$u(t,x) = \mathbb{E}\left[f\left(t, X^{\tau,x}(t)\right)\right].$$

Then the following identity holds:

$$\begin{aligned} D_1 u(t,x) &= \mathbb{E}\left[D_1 f\left(t, X^{\tau,x}(t)\right)\right] + \mathbb{E}\left[L(t) f\left(t, X^{\tau,x}(t)\right)\right] \\ &\quad + \mathbb{E}\left[\langle A(t)X^{\tau,x}(t), Df\left(t, X^{\tau,x}(t)\right)\rangle\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[\mathrm{Tr}\left(\sigma\left(t, X^{\tau,x}(t)\right)^* A(t)^* D^2 f\left(t, X^{\tau,x}(t)\right)\sigma\left(t, X^{\tau,x}(t)\right)\right)\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[\mathrm{Tr}\left(\sigma\left(t, X^{\tau,x}(t)\right)^* D^2 f\left(t, X^{\tau,x}(t)\right) A(t)\sigma\left(t, X^{\tau,x}(t)\right)\right)\right]. \end{aligned} \tag{1.201}$$

In the following result we introduce a certain backward propagator starting from a propagator on the Banach space E. In the finite-dimensional case a statement like Proposition 1.5 can be found in Lemma 8.3: formula (8.116) in Subsection 8.3.1 is the finite-dimensional analog of (1.202) below.

Proposition 1.5. *Let $\{C(t,\tau) : 0 \leqslant \tau \leqslant t \leqslant T\}$ be a forward propagator on E. Let $S(t,\tau) : H \to E$, $0 \leqslant \tau < t \leqslant T$, be a family operators with the following property*

$$C(t,s)S(s,\tau)S(s,\tau)^*C(t,s)^* + S(t,s)S(t,s)^* = S(t,\tau)S(t,\tau)^*, \tag{1.202}$$

for all $0 \leqslant \tau \leqslant s \leqslant t \leqslant T$. Let $t \mapsto W_H(t)$ be cylindrical Brownian motion, and put for $0 \leqslant \tau \leqslant t \leqslant T$

$$Y(\tau,t) f(x) = \mathbb{E}\left[f\left(C(t,\tau)x + S(t,\tau)W_H(1)\right)\right], \quad f \in C_b(E). \tag{1.203}$$

Then $Y(\tau,s)\,Y(s,t) = Y(\tau,t)$ for all $0 \leqslant \tau \leqslant s \leqslant t \leqslant T$.

Let $\sigma(\rho)$, $0 \leqslant \rho \leqslant T$, be a family of operators from H to E, and let the family $\{S(t,\tau) : 0 \leqslant \tau \leqslant t \leqslant T\}$ be such that, for $0 \leqslant \tau \leqslant t \leqslant T$, and $x^* \in E^*$,

$$S(t,\tau)S(t,\tau)^*x^* = \int_\tau^t C(t,\rho)\sigma(\rho)\sigma(\rho)^*C(t,\rho)^*x^*\,d\rho. \tag{1.204}$$

Then the family $\{S(t,\tau) : 0 \leqslant \tau \leqslant t \leqslant T\}$ possesses property (1.202) and

$$Y(\tau,t)f(x) = \mathbb{E}\left[f\left(C(t,\tau)x + \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)\right)\right],$$

and so the process $t \mapsto C(t,\tau)x + \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ can be considered as an E-valued Ornstein-Uhlenbeck process.

Proof. [Proof of Proposition 1.5.] Let $f \in C_b(E)$, and $0 \leqslant \tau \leqslant s \leqslant t \leqslant T$. In addition let

$$\left\{\left(\Omega^1, \mathcal{F}_t^1, \mathbb{P}^1\right), W_H^1(t)\right\} \quad \text{and} \quad \left\{\left(\Omega^2, \mathcal{F}_t^2, \mathbb{P}^2\right), W_H^2(t)\right\}$$

be independent copies of cylindrical Brownian motion. Then we have

$$\begin{aligned}
&Y(\tau,s)\,Y(s,t)f(x)\\
&= \mathbb{E}^1\left[Y(s,t)f\left(C(s,\tau)x + S(s,\tau)W_H^1(1)\right)\right]\\
&= \mathbb{E}^1\left[\mathbb{E}^2\left[f\left(C(t,s)\left(C(s,\tau)x + S(s,\tau)W_H^1(1)\right) + S(t,s)W_H^2(1)\right)\right]\right]\\
&= \mathbb{E}^1\left[\mathbb{E}^2\left[f\left(C(t,s)C(s,\tau)x + C(t,s)S(s,\tau)W_H^1(1) + S(t,s)W_H^2(1)\right)\right]\right]\\
&= \mathbb{E}^1\left[\mathbb{E}^2\left[f\left(C(t,\tau)x + C(t,s)S(s,\tau)W_H^1(1) + S(t,s)W_H^2(1)\right)\right]\right]. \qquad (1.205)
\end{aligned}$$

By general arguments, like the use of (Fourier transforms of) cylindrical measures and the separability of the space E it suffices to prove the equality

$$Y(\tau,s)\,Y(s,t)f(x) = \mathbb{E}\left[f\left(C(t,\tau)x + S(t,\tau)W_H(1)\right)\right] \qquad (1.206)$$

for functions f of the form $f(x) = e^{-i\langle x,x^*\rangle}$, $x^* \in E$. For more details on cylindrical measures on topological vector spaces see e.g. [Schwartz (1973)] part II. For such a function f we have

$$\begin{aligned}
&Y(\tau,s)\,Y(s,t)f(x)\\
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\\
&\quad \mathbb{E}^1\left[\mathbb{E}^2\left[\exp\left(-i\left\langle C(t,s)S(s,\tau)W_H^1(1), x^*\right\rangle\right)\exp\left(-i\left\langle S(t,s)W_H^2(1), x^*\right\rangle\right)\right]\right]\\
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\\
&\quad \mathbb{E}^1\left[\exp\left(-i\left\langle C(t,s)S(s,\tau)W_H^1(1), x^*\right\rangle\right)\right]\mathbb{E}^2\left[\exp\left(-i\left\langle S(t,s)W_H^2(1), x^*\right\rangle\right)\right]\\
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\\
&\quad \exp\left(-\frac{1}{2}\left\|S(s,\tau)^*C(t,s)^*x^*\right\|_H\right)\exp\left(-\frac{1}{2}\left\|S(t,s)^*x^*\right\|_H\right)\\
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\\
&\quad \exp\left(-\frac{1}{2}\left\langle C(t,s)S(s,\tau)S(s,\tau)^*C(t,s)^*x^*, x^*\right\rangle_H\right)\\
&\quad \exp\left(-\frac{1}{2}\left\langle S(t,s)S(t,s)^*x^*, x^*\right\rangle_H\right)
\end{aligned}$$

(employ (1.202))

$$
\begin{aligned}
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\exp\left(-\frac{1}{2}\left\langle S(t,\tau)S(t,\tau)^*x^*, x^*\right\rangle_H\right)\\
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\exp\left(-\frac{1}{2}\left\|S(t,\tau)^*x^*\right\|_H^2\right)\\
&= \exp\left(-i\left\langle C(t,\tau)x, x^*\right\rangle\right)\mathbb{E}\left[\exp\left(-i\left\langle S(t,\tau)W_H(1), x^*\right\rangle\right)\right]\\
&= \mathbb{E}\left[f\left(C(t,\tau)x + S(t,\tau)W_H(1)\right)\right] = Y\left(\tau,t\right)f(x).
\end{aligned}
\tag{1.207}
$$

This completes the proof of Proposition 1.5. □

Next, let $(e_j)_{j\in\mathbb{N}}$ be an orthonormal basis for the Hilbert space H, and $\rho \mapsto \sigma(\rho)$, $\rho \in [\tau, t]$, be an $L(H,E)$-valued process such that for every $j \in \mathbb{N}$ the mapping $\rho \mapsto \sigma(\rho)e_j$ is strongly measurable and adapted to the filtration determined by the cylindrical Brownian motion $\rho \mapsto W_H(\rho)$, $\tau \leqslant \rho \leqslant t$. In Lemma 1.4 just below the variables $W_{H,j}(\rho)$ stand for independent one-dimensional Brownian motions; in fact $W_{H,j}(\rho) = \langle W_H(\rho), e_j\rangle_H$.

Lemma 1.4. *Fix $0 \leqslant \tau \leqslant t \leqslant T$, and suppose that for every $x^* \in E^*$ the inequality holds:*

$$
\mathbb{E}\left[\int_\tau^t \left\|\sigma(\rho)^*C(t,\rho)^*x^*\right\|_H^2\, d\rho\right] < \infty. \tag{1.208}
$$

Then for every $x^* \in E^*$ the L^2-limit

$$
\begin{aligned}
&L^2\text{-}\lim_{N\to\infty}\sum_{j=1}^{N}\int_\tau^t \left\langle C(t,\rho)\sigma(\rho)e_j, x^*\right\rangle\, dW_{H,j}(\rho)\\
&= L^2\text{-}\lim_{N\to\infty}\sum_{j=1}^{N}\int_\tau^t \left\langle e_j, \sigma(\rho)^*C(t,\rho)^*x^*\right\rangle_H\, dW_{H,j}(\rho)
\end{aligned}
\tag{1.209}
$$

defines an element in $L^2\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}\right)$, and $\mathbb{P}$-almost surely this limit defines an element in E^{**}. This limit is written as $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$. In other words

$$
\begin{aligned}
&\left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho)dW_h(\rho)\right\rangle\\
&= L^2\text{-}\lim_{N\to\infty}\sum_{j=1}^{N}\int_\tau^t \left\langle e_j, \sigma(\rho)^*C(t,\rho)^*x^*\right\rangle_H\, dW_{H,j}(\rho).
\end{aligned}
\tag{1.210}
$$

Moreover, the mapping

$$
x^* \mapsto \mathbb{E}\left[\int_\tau^t C(t,\rho)\sigma(\rho)\sigma(\rho)^*C(t,\rho)^*x^*d\rho\right], \quad x^* \in E^*, \tag{1.211}
$$

is a continuous linear mapping from E^* *to* E*, and the following equality holds for all* $x^*, y^* \in E^*$*:*

$$\mathbb{E}\left[\left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)\right\rangle\left\langle y^*, \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)\right\rangle\right]$$
$$= \mathbb{E}\left[\int_\tau^t \langle \sigma(\rho)^* C(t,\rho)^* x^*, \sigma(\rho)^* C(t,\rho)^* y^*\rangle_H\, d\rho\right]$$
$$= \left\langle \mathbb{E}\left[\int_\tau^t C(t,\rho)\sigma(\rho)\sigma(\rho)^* C(t,\rho)^* x^* d\rho\right], y^*\right\rangle. \tag{1.212}$$

Let $F \in L^2(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$*. Then the functional*

$$x^* \mapsto \mathbb{E}\left[F\left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)\right\rangle\right] \tag{1.213}$$

is sequentially continuous for the weak-topology. In other words the mapping*

$$A \mapsto \mathbb{E}\left[\mathbf{1}_A \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)\right], \quad A \in \mathcal{F}_t^\tau, \tag{1.214}$$

can be considered as an E*-valued vector measure which is absolutely continuous relative to the measure* $\mathbb{P}$*.*
Some conditions which guarantee that the stochastic integral

$$\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$$

belongs to E $\mathbb{P}$*-almost surely are inserted.*

(a) *If the Banach space* E *has the weak* L^2*-Radon-Nikodym property relative to the probability space* $(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$*, then the stochastic integral*

$$\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho),$$

which is the weak L^2*-Radon-Nikodym derivative of the vector measure in (1.214), belongs to* E $\mathbb{P}$*-almost surely.*

(b) *If for every* $x^{***} \in E^{***}$ *and every* $A \in \mathcal{F}_t^\tau$ *the equality*

$$\mathbb{E}\left[\mathbf{1}_A\left\langle \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho), x^{***}\right\rangle\right]$$
$$= \left\langle \mathbb{E}\left[\mathbf{1}_A \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)\right], x^{***}\right\rangle \tag{1.215}$$

holds, then the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ *belongs to* E $\mathbb{P}$*-almost surely.*

(c) If the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ is $\mathbb{P}$-almost surely contained in a $\|\cdot\|$-separable subspace of E^{**}, then it belongs E $\mathbb{P}$-almost surely.

If the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ belongs to E $\mathbb{P}$-almost surely, and if for every vector $h \in H$ and $x^* \in E^*$ the function $\rho \mapsto \langle C(t,\rho)\sigma(\rho)h, x^*\rangle$ is continuous, then it belongs to the $\|\cdot\|$-closure of the subset

$$\left\{ \frac{\left\| \mathbf{1}_A \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right\|}{\mathbb{P}[A]} \mathbf{1}_A : A \in \mathcal{F}_t^\tau,\, \mathbb{P}[A] \neq 0 \right\}. \tag{1.216}$$

In the final assertion the weak operator continuity condition on the operators $C(t,\rho)\sigma(\rho)$, $\tau \leqslant \rho \leqslant t$, can be relaxed. A somewhat more refined argument yields the following result. Suppose that the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ belongs to a separable subspace of E, and suppose that for every $h \in H$ and $x^* \in E^*$ the process $\rho \mapsto \langle C(t,\rho)\sigma(\rho)h, x^*\rangle$ is predictable. Then $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ belongs to the closure of the family in (1.216). This means that the variable $(s,\omega) \mapsto \langle C(t,s)\sigma(s)h, x^*\rangle$ is measurable relative to σ-field generated by the set $\{(a,b] \times A : \tau \leqslant a < b \leqslant t,\, A \in \mathcal{F}_a^\tau\}$.

Definition 1.7. A closed, bounded and convex subset C of E is said to have the WRNP (weak Radon-Nikodym property, or weak L^1-Radon-Nikodym property) with respect to $(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ if for every measure $G : \mathcal{F}_t^\tau \to E$ such that $G(A) \in \mathbb{P}(A) \cdot C$ for every $A \in \mathcal{F}_t^\tau$, there exists a Pettis integrable and $\mathcal{F}_t^\tau$-measurable function $g : \Omega \to C$ such that

$$\langle G(A), x^*\rangle = \mathbb{E}\left[\mathbf{1}_A \langle g, x^*\rangle\right] \tag{1.217}$$

for each $A \in \mathcal{F}_t^\tau$ and $x^* \in E^*$. We say that the set C has the WRNP if C has this property with respect to every probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Such a set C is called a weak Radon-Nikodym set. A Banach space E is said to have the WRNP (resp. WRNP with respect to $(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$) if the unit ball of E is a weak Radon-Nikodym set (resp. has the WRNP with respect to $(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$). The space E is said to have the weak L^p-Radon-Nikodym property, $1 \leqslant p < \infty$, with respect to $(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ if for every measure $G : \mathcal{F}_t^\tau \to E$ such that $\left\|\int_\Omega F\, dG\right\| \leqslant \|F\|_{L^p}$ for all $F \in L^p(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ there exists a Pettis integrable and $\mathcal{F}_t^\tau$-measurable function $g : \Omega \to E$ such that

$$\langle G(A), x^*\rangle = \mathbb{E}\left[\mathbf{1}_A \langle g, x^*\rangle\right] \tag{1.218}$$

for each $A \in \mathcal{F}_t^\tau$ and $x^* \in E^*$. Such a function g has the property that $\mathbb{E}\left[|\langle g, x^*\rangle|^q\right] \leqslant \|x^*\|^q$ for all $x^* \in E^*$. The $\mathbb{P}$-almost surely E-valued function g is called the weak, or Pettis, L^p-derivative of the measure G. Here q is the conjugate exponent of p: $q^{-1} + p^{-1} = 1$. If $p = 2$, then $q = 2$.

For the Radon-Nikodym theorem and related topics from a historical perspective see e.g. [Pietsch (2007)].

Let $g : \Omega \to C$ be a random variable such that for every $x^* \in E^*$ the variable $\langle g, x^*\rangle$ is $\mathcal{F}_t^\tau$-measurable. Then g is said to be Pettis-integrable if for every $A \in \mathcal{F}_t^\tau$ there exists an element $x_A \in E$ such that $\langle x_A, x^*\rangle = \mathbb{E}\left[\mathbf{1}_A \langle g, x^*\rangle\right]$ for all $x^* \in E^*$.

For more details on the weak Radon-Nikodym property see e.g. [Riddle (1984)], [Matsuda (1985)], or [Farmaki (1995)]. For more details on Pettis integrability see e.g. [Diestel and Uhl (1977)].

Proof. [Proof of Lemma 1.4.] First we calculate

$$\begin{aligned}
&\mathbb{E}\left[\left|\sum_{j=1}^N \int_\tau^t \langle C(t,\rho)\sigma(\rho)e_j, x^*\rangle_H \, dW_{H,j}(\rho)\right|^2\right] \\
&= \mathbb{E}\left[\left|\sum_{j=1}^N \int_\tau^t \langle e_j, \sigma(\rho)^* C(t,\rho)^* x^*\rangle_H \, dW_{H,j}(\rho)\right|^2\right] \\
&= \sum_{j=1}^N \mathbb{E}\left[\int_\tau^t \left|\langle e_j, \sigma(\rho)^* C(t,\rho)^* x^*\rangle_H\right|^2 d\rho\right] \\
&= \mathbb{E}\left[\int_\tau^t \sum_{j=1}^N \left|\langle e_j, \sigma(\rho)^* C(t,\rho)^* x^*\rangle_H\right|^2 d\rho\right] \\
&\leqslant \mathbb{E}\left[\int_\tau^t \sum_{j=1}^\infty \left|\langle e_j, \sigma(\rho)^* C(t,\rho)^* x^*\rangle_H\right|^2 d\rho\right] \\
&= \mathbb{E}\left[\int_\tau^t \|\sigma(\rho)^* C(t,\rho)^* x^*\|_H^2 \, d\rho\right].
\end{aligned} \tag{1.219}$$

Since the ultimate term in (1.212) is finite, it easily follows that the L^2-limit (1.209) exists. Next observe that $x_n^* \to x^*$ in $(E^*, \|\cdot\|)$, and

$$\left\langle x_n^*, \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right\rangle \to F \quad \text{in } L^2\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}\right)$$

implies $F = \left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right\rangle$. Hence, by the closed graph theorem we see that there exists a constant c such that

$$\begin{aligned} &\mathbb{E}\left[\left|\left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho) dW_H(\rho)\right\rangle\right|^2\right] \\ &= \mathbb{E}\left[\int_\tau^t \left\|\sigma(\rho)^* C(t,\rho)^* x^*\right\|_H^2 d\rho\right] \leqslant c^2 \left\|x^*\right\|^2. \end{aligned} \tag{1.220}$$

The proof of (1.212) can be patterned after the proof of equality (1.219). The fact that the mapping in (1.211) is E-valued can be proved by employing the Krein-Smulian theorem, or the Grothendieck completeness theorem. Fix $x^* \in E^*$. By separability of the subspace of E spanned by $C(t,\rho)\sigma(\rho)\sigma(\rho)^* C(t,\rho)^* x^*$, $\tau \leqslant \rho \leqslant t$, together with Grothendieck's completeness theorem it suffices to prove that

$$\begin{aligned} &\lim_{n\to\infty} \left\langle \mathbb{E}\left[\int_\tau^t C(t,\rho)\sigma(\rho)\sigma(\rho)^* C(t,\rho)^* x^* d\rho\right], x_n^*\right\rangle \\ &= \lim_{n\to\infty} \mathbb{E}\left[\int_{\tau,t} \left\langle C(t,\rho)\sigma(\rho)\sigma(\rho)^* C(t,\rho)^* x^*, x_n^*\right\rangle \, d\rho\right] \\ &= 0 \end{aligned} \tag{1.221}$$

whenever $(x_n^*)_{n\in\mathbb{N}}$ is a sequence in the unit ball of E^* which converges weak* to the zero-functional. The conclusion in (1.221) then follows from Lebesgue's dominated convergence theorem.

We continue by proving (1.213). First we do this for F of the form $F = Tx^* := \left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right\rangle$, $x^* \in E^*$. Let $(x_n^*)_{n\in\mathbb{N}}$ be a sequence in the unit ball of E^* which converges weak* to the zero-functional. Notice that by the Banach-Steinhaus theorem any sequence in E^* which converges weakly is bounded; without loss of generality we assume that such a sequence is contained in the dual unit ball. Then we have

$$\begin{aligned} &\mathbb{E}\left[F\left\langle x_n^*, \int_\tau^t C(t,\rho)\sigma(\rho) d W_H(\rho)\right\rangle\right] \\ &= \mathbb{E}\left[\left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho) d W_H(\rho)\right\rangle \left\langle x_n^*, \int_\tau^t C(t,\rho)\sigma(\rho) d W_H(\rho)\right\rangle\right] \\ &= \mathbb{E}\left[\int_\tau^t \left\langle C(t,\rho)\sigma(\rho)\sigma(\rho)^* C(t,\rho)^* x^*, x_n^*\right\rangle \, d\rho\right]. \end{aligned} \tag{1.222}$$

By dominated convergence the expression in (1.222) converges to zero when n tends to ∞. Let F belong to L_1, which by definition is the L^2-closure of the subspace

$$\left\{Tx^* = \left\langle x^*, \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right\rangle : \, x^* \in E^*\right\}.$$

Then an approximation argument shows the equality in (1.213) for such variables F. Finally, if $F \in L^2(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ we decompose $F = F_1 + F_2$, where $F_1 \in L_1$, and $F_2 \in L_1^\perp$. Then we see

$$\lim_{n\to\infty} \mathbb{E}[F T x_n^*] = \lim_{n\to\infty} \mathbb{E}[F_1 T x_n^*] = 0.$$

This proves (1.213). Next we show that mapping $G : \mathcal{F}_t^\tau \to E$, defined by $G(A) = \mathbb{E}\left[\mathbf{1}_A \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right]$ is an E-valued measure: see (1.214). To this end we take a sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{F}_t^\tau$ which decreases to the empty set. Then we have to prove that $\lim_{n\to\infty} \|G(A_n)\| = 0$. Then for $x^* \in E^*$, $\|x^*\| \leqslant 1$, we estimate

$$\begin{aligned}
&\left|\left\langle \mathbb{E}\left[\mathbf{1}_{A_n} \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\right], x^*\right\rangle\right|^2 \\
&= \left|\mathbb{E}\left[\mathbf{1}_{A_n} \left\langle \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho), x^*\right\rangle\right]\right|^2 \\
&\leqslant \mathbb{P}[A_n]\, \mathbb{E}\left[\left|\left\langle \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho), x^*\right\rangle\right|^2\right] \\
&= \mathbb{P}[A_n]\, \mathbb{E}\left[\int_\tau^t \|\sigma(\rho)^* C(t,\rho)^* x^*\|_H^2\, d\rho\right] \leqslant c^2 \mathbb{P}[A_n]\, \|x^*\|^2. \qquad (1.223)
\end{aligned}$$

In the final estimate of (1.223) we employed (1.220). By the Hahn-Banach theorem and (1.223) we see that $\lim_{n\to\infty} \|G(A_n)\| = 0$, which shows that the set function in (1.214) is an E-valued measure.

Next we prove the Assertions (a), (b) and (c).

(a). Since for every $x^* \in E^*$ the variable $\left\langle \int_\tau^t C(t,\rho)\sigma(\rho) dW_H(\rho), x^*\right\rangle$ is the Radon-Nikodym derivative of the measure $A \mapsto \langle G(A), x^*\rangle$ the weak L^2-Radon-Nikodym property implies that the stochastic integral

$$\int_\tau^t C(t,\rho)\sigma(\rho) dW_H(\rho)$$

belongs to E $\mathbb{P}$-almost surely.

(b). We already know that the set function in (1.214) is an E-valued measure. Pick $x^{***} \in E^{***}$. From the classical Radon-Nikodym theorem it follows that $\mathbb{P}$-almost surely the random variable $\left\langle \int_\tau^t C(t,\rho)\sigma(\rho) dW_H(\rho), x^{***}\right\rangle$ can be written as a limit of quotients of the form

$$\frac{\mathbb{E}\left[\mathbf{1}_A \left\langle \int_\tau^t C(t,\rho)\sigma(\rho) dW_H(\rho), x^{***}\right\rangle\right]}{\mathbb{P}[A]}$$

$$= \frac{\left\langle \mathbb{E}\left[\mathbf{1}_A \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)\right], x^{***}\right\rangle}{\mathbb{P}[A]}.$$

$A \in \mathcal{F}_t^\tau$, $P[A] \neq 0$. From this observation we see that the stochastic integral

$$\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho),$$

which $\mathbb{P}$-almost surely is a member of E^{**}, in fact belongs $\mathbb{P}$-almost surely to the weak-closure, i.e. the $\sigma\left(E^{**}, E^{***}\right)$-closure, of the collection of vectors of the form

$$\frac{\mathbb{E}\left[\mathbf{1}_A \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)\right]}{\mathbb{P}[A]}\mathbf{1}_A, \quad A \in \mathcal{F}_t^\tau. \tag{1.224}$$

By Mazur's theorem in functional analysis it follows that $\mathbb{P}$-almost surely the stochastic integral

$$\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$$

belongs to the $\|\cdot\|$-closed convex hull of vectors of the form (1.224). Since the vectors in (1.224) belong to E, it follows that the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ is a member of E $\mathbb{P}$-almost surely.

(c). Let $\left(\Omega', \mathcal{F}_t^{\prime,\tau}, \mathbb{P}'\right)$ be an independent copy of $(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$. Since the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ is $\mathbb{P}$-almost surely contained in $\|\cdot\|$-separable subspace of E^{**} we can find a double sequence of events $A_{m,k}$, m, $k \in \mathbb{N}$, with the following properties: $\mathbb{P}\left[A_{m,k}\right] \neq 0$, $\mathbb{P}\left[\bigcap_{m=1}^\infty \bigcup_{k=1}^\infty A_{m,k}\right] = 1$, and

$$\begin{aligned}\Bigg\| \mathbf{1}_{A_{m,k}}(\omega) \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)\,(\omega) \\ - \mathbf{1}_{A_{m,k}}\left(\omega'\right) \int_\tau^t C(t,\rho)\sigma(\rho)\sigma(\rho)\, dW_H(\rho)\left(\omega'\right)\Bigg\| \leqslant 2^{-m},\end{aligned} \tag{1.225}$$

$\mathbb{P} \times \mathbb{P}'$-almost surely. For brevity we temporarily employ the following notation:

$$M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho).$$

Then (1.225) reads as follows:

$$\left\|\mathbf{1}_{A_{m,k}}(\omega)\, M_\infty(\omega) - \mathbf{1}_{A_{m,k}}\left(\omega'\right) M_\infty\left(\omega'\right)\right\| \leqslant 2^{-m}, \quad \mathbb{P} \times \mathbb{P}'\text{-almost surely.} \tag{1.226}$$

Next we consider $\omega_0 \in \bigcap_{m=1}^{\infty} A_{m,k_m}$, and write

$$\begin{aligned}
&\frac{\mathbb{E}\left[\mathbf{1}_{A_{m,k_m}} M_\infty\right]}{\mathbb{P}\left[A_{m,k_m}\right]}\mathbf{1}_{A_{m,k_m}}\left(\omega'\right) - M_\infty\left(\omega_0\right) \\
&= \frac{\mathbb{E}\left[\mathbf{1}_{A_{m,k_m}}\left\{\mathbf{1}_{A_{m,k_m}} M_\infty - \mathbf{1}_{A_{m,k_m}}\left(\omega'\right) M_\infty\left(\omega'\right)\right\}\right]}{\mathbb{P}\left[A_{m,k_m}\right]}\mathbf{1}_{A_{m,k_m}}\left(\omega'\right) \\
&\quad + \mathbf{1}_{A_{m,k_m}}\left(\omega'\right) M_\infty\left(\omega'\right) - \mathbf{1}_{A_{m,k_m}}\left(\omega_0\right) M_\infty\left(\omega_0\right). \qquad (1.227)
\end{aligned}$$

From (1.226) and (1.227) we infer

$$\begin{aligned}
&\left\|\frac{\mathbb{E}\left[\mathbf{1}_{A_{m,k_m}} M_\infty\right]}{\mathbb{P}\left[A_{m,k_m}\right]}\mathbf{1}_{A_{m,k_m}}\left(\omega'\right) - M_\infty\left(\omega_0\right)\right\| \\
&\leqslant \frac{\mathbb{E}\left[\mathbf{1}_{A_{m,k_m}}\left\|\mathbf{1}_{A_{m,k_m}} M_\infty - \mathbf{1}_{A_{m,k_m}}\left(\omega'\right) M_\infty\left(\omega'\right)\right\|\right]}{\mathbb{P}\left[A_{m,k_m}\right]}\mathbf{1}_{A_{m,k_m}}\left(\omega'\right) \\
&\quad + \left\|\mathbf{1}_{A_{m,k_m}}\left(\omega'\right) M_\infty\left(\omega'\right) - \mathbf{1}_{A_{m,k_m}}\left(\omega_0\right) M_\infty\left(\omega_0\right)\right\| \\
&\leqslant 2^{-m}\mathbf{1}_{A_{m,k_m}}\left(\omega'\right) + 2^{-m} \leqslant 2^{-m+1}. \qquad (1.228)
\end{aligned}$$

Consequently, $M_\infty \in E$ $\mathbb{P}$-almost surely.

Finally, suppose that $M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)$ belongs to E $\mathbb{P}$-almost surely. Let $(x_k)_{k\in\mathbb{N}}$ be an enumeration of random vectors of the form

$$\sum_{n,N=1}^{\infty} \alpha_{n,N} \sum_{\ell=1}^{2^n}\sum_{j=1}^{N} C\left(t,\rho_{\ell,n}\right)\sigma\left(\rho_{\ell,n}\right) e_j \left(W_{H,j}\left(\rho_{\ell+1,n}\right) - W_{H,j}\left(\rho_{\ell,n}\right)\right), \qquad (1.229)$$

where $\rho_{\ell,n} = \tau + \ell 2^{-n}(t-\tau)$, and the $\alpha_{n,N}$'s are non-negative rational numbers such that $\sum_{n,N=1}^{\infty}\alpha_{n,N} = 1$ and such that only finitely many of them are non-zero. From the definition of the stochastic integral $M_\infty := \int_\tau^t C(t,\rho)\sigma(\rho)dW_H(\rho)$ it follows that M_∞ belongs to the weak closure (i.e. $\sigma(E,E^*)$-closure) of the family in (1.229). But then M_∞ belongs to the $\|\cdot\|$-closure of the sequence $(x_k)_{k\in\mathbb{N}}$. Define the events $A_{m,k}$, m, $k \in \mathbb{N}$, by $A_{m,k} = \left\{\|M_\infty - x_k\| < 2^{-m-1}\right\}$. Then $\mathbb{P}\left[\bigcap_{m=1}^{\infty}\bigcup_{k=1}^{\infty} A_{m,k}\right] = 1$. Define the E-valued martingale M_∞^m, $m \in \mathbb{N}$, by $M_\infty^m = \mathbb{E}\left[M_\infty \,\middle|\, \Pi_m\right]$ where Π_m is the σ-field generated by $\{A_{m,k} : k \in \mathbb{N}\}$. As in the proof of assertion (c) of this Lemma it follows that $M_\infty = \|\cdot\|\text{-}\lim_{m\to\infty} M_\infty^m$.

All this together completes the proof of Lemma 1.4. □

Lemma 1.5. *Put $M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)$ Let y^* be the weak* limit of a necessarily bounded sequence $(y_n^*)_{n\in\mathbb{N}}$. Then the following equality holds:*

$$\mathbb{P}\left[\langle M_\infty, y^*\rangle \leqslant \limsup_{n\to\infty}\langle M_\infty, y_n^*\rangle < \infty\right] = 1. \qquad (1.230)$$

Proof. From (1.213) we see that the sequence $\{\langle M_\infty, y_n^*\rangle : n \in \mathbb{N}\}$ converges in the L^2-weak sense to $\langle M_\infty, y^*\rangle$. From Mazur's theorem it follows that for an appropriate sequence of convex combinations $x_n^* = \sum_{j=n}^\infty \alpha_{j,n} y_j^*$ we have $\langle M_\infty, y^*\rangle = L^2\text{-}\lim_{n\to\infty} \langle M_\infty, x_n^*\rangle$. By passing to a subsequence we have that $\langle M_\infty, y^*\rangle = \lim_{k\to\infty} \langle M_\infty, x_{n_k}^*\rangle$, $\mathbb{P}$-almost surely. Since $\alpha_{j,n} \geqslant 0$ and sum to 1 we see $\langle M_\infty, y^*\rangle \leqslant \limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle$ $\mathbb{P}$-almost surely. This completes the proof of Lemma 1.5. □

Theorem 1.17. *The stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)$ attains its values in E $\mathbb{P}$-almost surely if and only if there exists an $\mathbb{P}$-almost sure event Ω' such that the following inequalities*

$$0 \leqslant \limsup_{n\to\infty} \left\langle \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho), y_n^* \right\rangle < \infty \tag{1.231}$$

hold on Ω' for all sequences $(y_n^*)_{n\in\mathbb{N}}$ in E^* which converge in weak-sense to* 0.

The point here is that the event Ω' does not depend on the weak*-convergent sequence $(y_n^*)_{n\in\mathbb{N}}$: compare with Lemma 1.5. Observe that the arguments of the proof of Theorem 1.17, which is given below, also occur on page 268 (Chapter 5 of Part II) in [Schwartz (1973)]. In addition, notice that the following construction gives the corresponding cylindrical measure

$$\mu = \left\{\mu_{E/F} : F \subset E, \operatorname{codim}(F) < \infty\right\} \quad \text{on } E.$$

For a closed linear subspace F with $\operatorname{codim}(F) = n$ choose an independent subset consisting of n elements in E^* such that $F = \bigcap_{j=1}^n \left\{x \in E : \left\langle x, x_j^*\right\rangle = 0\right\}$, and define the mapping $\widetilde{\pi}_{x_1^*,\dots,x_n^*} : E/F \to \mathbb{R}^n$ by

$$\widetilde{\pi}_{x_1^*,\dots,x_n^*}(x+F) = \left(\langle x, x_1^*\rangle, \dots, \langle x, x_n^*\rangle\right), \quad x \in E.$$

The measure $\mu_{E/F}$ on the Borel field of E/F is determined by

$$\mu_{E/F}(B+F) = \mathbb{P}\left[\left(\langle M_\infty, x_1^*\rangle, \dots, \langle M_\infty, x_1^*\rangle\right) \in \widetilde{\pi}_{x_1^*,\dots,x_n^*}(B+F)\right]. \tag{1.232}$$

Here $B+F$ is a Borel subset of E/F. Then it can be checked that μ is a cylindrical measure indeed. The variable M_∞ is given by e.g.

$$M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho).$$

Of course the variable M_∞ could be replaced with any other variable $Y \in L^2_{\text{weak}}(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$. For the definition of this space the reader is referred to Definition 1.8 below.

Notice that the stochastic integral attains its values in E^{**} $\mathbb{P}$-almost surely if and only if there exists an $\mathbb{P}$-almost sure event Ω' such that the following inequality

$$\limsup_{n\to\infty} \left\langle \int_{\tau}^{t} C(t,\rho)\sigma(\rho)\, dW_H(\rho), y_n^* \right\rangle < \infty \tag{1.233}$$

holds on Ω' for all $y^* \in E^*$ and all sequences $(y_n^*)_{n\in\mathbb{N}}$ in E^* which converge in weak*-sense to y^*. This result is a consequence of the Banach-Alaoglu theorem, and the separability of E.

The following theorem in functional analysis can be proved along the same lines as Theorem 1.17. The theorem of Krein-Smulian (see Theorem 6.4 Corollary in [Schaefer (1971)]), or Grothendieck (see Corollary 2 to Theorem 6.2 in [Schaefer (1971)]) plays a dominant role in the proof of Theorem 1.18. By definition a sequence $(x_n^*)_{n\in\mathbb{N}} \subset E^*$ belongs to $c_0(\mathbb{N}, E^*)$ if $\lim_{n\to\infty} \langle x, x_n^* \rangle = 0$ for every $x \in E$.

Theorem 1.18. *Let E be a separable Banach space, and let $f : E^* \to \mathbb{R}$ be a linear functional. Then the following assertions are equivalent:*

(a) There exists $x \in E$ such that $f(x^*) = \langle x, x^* \rangle$ for all $x^* \in E^*$;
(b) For every sequence $(x_n^*)_{n\in\mathbb{N}} \in c_0(\mathbb{N}, E^*)$ the following inequalities hold:

$$0 \leqslant \sup_{n\in\mathbb{N}} f(x_n^*) < \infty.$$

(c) For every sequence $(x_n^*)_{n\in\mathbb{N}} \in c_0(\mathbb{N}, E^*)$ the following inequalities hold:

$$0 \leqslant \limsup_{n\to\infty} f(x_n^*) < \infty.$$

Proof. [Proof of Theorem 1.18.] (a) $\Longrightarrow$ (b). A sequence in $c_0(\mathbb{N}, E^*)$ is norm-bounded in E^*; this is a consequence of e.g. the Banach-Steinhaus theorem. It is also a consequence of a Baire-category argument applied to the dual unit ball. Hence assertion (b) follows from (a).

(b) $\Longrightarrow$ (c). Let $(x_n^*)_{n\in\mathbb{N}}$ be any sequence in $c_0(\mathbb{N}, E^*)$. Then $(x_k^*)_{k\in\mathbb{N}, k\geqslant n}$ is a sequence in $c_0(\mathbb{N}, E^*)$, and so, by (b), $0 \leqslant \sup_{k\geqslant n} f(x_k^*) < \infty$, from which assertion (c) readily follows.

(c) $\Longrightarrow$ (a). In this implication we will employ the Krein-Smulian theorem, or Grothendieck's completeness result. So suppose that (c) holds, and let $(y_n^*)_{n\in\mathbb{N}}$ be any sequence in E^* which converges to in weak*-sense to $y^* \in E^*$. By (c) we see $0 \leqslant \limsup_{n\to\infty} f(y_n^* - y^*) < \infty$, and hence

$$f(y^*) \leqslant \limsup_{n\to\infty} f(y_n^*) < \infty. \tag{1.234}$$

From (1.234) it follows that for every $M \in \mathbb{N}$ and every $\alpha \in \mathbb{R}$ the subset

$$\{x^* \in E^* : \|x^*\| \leqslant M,\, f(x^*) \leqslant \alpha\} \tag{1.235}$$

is sequentially weak*-closed. Since E is separable, and the set in (1.235) is equi-continuous, it follows that sets of the form (1.235) are weak*-closed, not just sequentially weak*-closed. From Krein-Smulian's theorem it follows that for every $\alpha \in \mathbb{R}$ the half-space $\{x^* \in E^* : f(x^*) \leqslant \alpha\}$ is weak*-closed. It then follows that the hyper-plane $\{x^* \in E^* : f(x^*) = 0\}$ is weak*-closed. Consequently, there exists a vector $x \in E$ such that $f(x^*) = \langle x, x^*\rangle$, $x^* \in E^*$.

We can also use Grothendieck's theorem. Then we proceed as follows. Instead of considering a set of the form (1.235) we look at the subset $H_{M,\alpha}$ defined by

$$H_{M,\alpha} = \{x^* \in E^* : \|x^*\| \leqslant M,\, f(x^*) = \alpha\}. \tag{1.236}$$

Then the set in (1.236) is sequentially weak*-closed. Let $(x_n^*)_{n\in\mathbb{N}}$ be a sequence in $H_{M,\alpha}$ which converges to $x^* \in E^*$ in weak*-sense. Then, by (c),

$$f(x^*) \leqslant \limsup_{n\to\infty} f(x_n^*) = \limsup_{n\to\infty} \alpha = \alpha. \tag{1.237}$$

Applying the same argument to the sequence $(-x_n^*)_{n\in\mathbb{N}}$ which converges in weak*-sense to $-x^*$ shows $f(-x^*) \leqslant -\alpha$. This in combination with (1.237) yields $f(x^*) = \alpha$, and consequently the subset $H_{M,\alpha}$ is sequentially weak*-closed. Since the space is separable and the set $H_{M,\alpha}$ is equi-continuous it follows that $H_{M,\alpha}$ is weak*-closed. Grothendieck's theorem then implies that the hyper-plane $\{x^* \in E^* : f(x^*) = \alpha\}$ is weak*-closed. Again it follows that there exists $x \in E$ such that $f(x^*) = \langle x, x^*\rangle$, $x^* \in E^*$.

This completes the proof of Theorem 1.18. □

Proof. [Proof of Theorem 1.17.] Put $M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)$. If $M_\infty \in E$ $\mathbb{P}$-almost surely, then we have equality in (1.231) on the event $\Omega' = \{M_\infty \in E\}$.

Next we prove that the stochastic integral $\int_\tau^s C(s,\rho)\sigma(\rho)\, dW_H(\rho)$ attains its values in E $\mathbb{P}$-almost surely provided that (1.231) is satisfied. The scalar L^2-space $L^2(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ is separable, and so is its subspace L_1 which by definition is the L^2-closure of $\{Tx^* : x^* \in E^*\}$. So there exists a countable family $\left(x_j^*\right)_{j\in\mathbb{N}}$ in the closed unit ball of E^* such that the linear span of the countable family $\left(Tx_j^*\right)_{j\in\mathbb{N}}$ is dense in L_1. Since E is separable, and T is sequentially continuous relative to the weak*-topology on E^* and the weak topology in $L^2(\Omega, \mathcal{F}_s^\tau, \mathbb{P})$, we may assume additionally that the sequence $\left(x_j^*\right)_{j\in\mathbb{N}}$ is dense in the dual unit ball. Then by the theorems of Krein-Smulian (see Theorem 6.4 Corollary in [Schaefer (1971)]) and Grothendieck (see Corollary 2 to Theorem 6.2 in [Schaefer (1971)]) we obtain the following equality of events:

$$\begin{aligned}
&\{M_\infty \in E\} \\
&= \{\{x^* \in E^*, \|x^*\| \leqslant M, \langle M_\infty, x^*\rangle \leqslant \alpha\} \\
&\qquad \text{is sequentially weak*-closed for all } M \geqslant 0 \text{ and } \alpha \in \mathbb{R}\} \qquad (1.238) \\
&= \left\{\langle M_\infty, y^*\rangle \leqslant \limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle < \infty \ \text{ whenever } y_n^* \to y^* \text{ weak*}\right\} \qquad (1.239) \\
&= \left\{0 \leqslant \limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle < \infty \ \text{ whenever } y_n^* \to 0 \text{ weak*}\right\} \qquad (1.240) \\
&= \{\{x^* \in E^*, \|x^*\| \leqslant M, \langle M_\infty, x^*\rangle = \alpha\} \\
&\qquad \text{is sequentially weak*-closed for all } M \geqslant 0 \text{ and } \alpha \in \mathbb{R}\} \qquad (1.241) \\
&= \{\{x^* \in E^*, \|x^*\| \leqslant M, \langle M_\infty, x^*\rangle = \alpha\} \\
&\qquad \text{is weak*-closed for all } M \geqslant 0 \text{ and } \alpha \in \mathbb{R}\}. \qquad (1.242)
\end{aligned}$$

The equality of the event $\{M_\infty \in E\}$ and the one in (1.238) is a consequence of Krein-Smulian's theorem, and in proving the equality of the events in (1.238) and (1.239) the fact is used that weak*-bounded subsets are norm-bounded, and that the space E is separable. A consequence of the latter is that the dual unit ball is a compact metric space. Therefore the inclusion of the event in (1.238) in the one in (1.239) can be seen as follows. Let $(y_n^*)_{n\in\mathbb{N}}$ be a sequence in E^* which converges in weak*-sense to $y^* \in E^*$. Fix $\alpha \in \mathbb{R}$, and consider the event

$$\left\{\limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle \leqslant \alpha\right\}.$$

Then $\sup_{n\in\mathbb{N}} \|y_n^*\| \leqslant M < \infty$, and on the event in (1.238) we have

$$\left\{\limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle \leqslant \alpha\right\} \supset \{\langle M_\infty, y^*\rangle \leqslant \alpha\}. \qquad (1.243)$$

Since $\alpha \in \mathbb{R}$ is arbitrary, from (1.243) we see that

$$\langle M_\infty, y^*\rangle \leqslant \limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle < \infty,$$

and hence the event in (1.238) is contained in the one (1.239). Again let $(y_n^*)_{n\in\mathbb{N}}$ be a, necessarily bounded, sequence in E^* which converges in weak*-sense to $y^* \in E^*$. Then consider the event

$$\bigcap_{n\in\mathbb{N}} \{\langle M_\infty, y_n^*\rangle = \alpha\} = \bigcap_{n\in\mathbb{N}} \{\langle M_\infty, -y_n^*\rangle = -\alpha\}.$$

On the intersection of this event and the one in (1.239) we have

$$\begin{aligned} \langle M_\infty, y^*\rangle &\leqslant \limsup_{n\to\infty} \langle M_\infty, y_n^*\rangle = \alpha, \quad \text{and also} \\ \langle M_\infty, -y^*\rangle &\leqslant \limsup_{n\to\infty} \langle M_\infty, -y_n^*\rangle = -\alpha, \end{aligned} \tag{1.244}$$

and hence $\langle M_\infty, y^*\rangle = \alpha$. From (1.244) we then easily infer that the event in (1.239) is contained in the fifth one, i.e. the event in (1.241). The equality of the events in (1.239) and in (1.240) follows by taking $y_n^* - y^*$ instead of y_n^*. The equalities of the events (1.241) and (1.242) is a consequence of the separability of the space E. That the event in (1.242) is contained in the event $\{M_\infty \in E\}$ is a consequence of Grothendieck's theorem. By hypothesis we know that the event in (1.240) contains the $\mathbb{P}$-almost sure event Ω'. By the equalities of the event $\{M_\infty < \infty\}$ and the one in (1.240) this shows that the event $\{M_\infty \in E\}$ is $\mathbb{P}$-almost sure.

This concludes the proof of Theorem 1.17. □

In the following theorem we give some alternative formulations for conditions which guarantee that $M_\infty \in E^{**}$ $\mathbb{P}$-almost surely.

Theorem 1.19. *Let $M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)$ be a stochastic integral: see Theorem 1.17. Suppose that the dual space endowed with the norm topology is separable. The following assertions are equivalent:*

(i) M_∞ belongs to E^{**} $\mathbb{P}$-almost surely.
(ii) For every sequence $(y_n^*)_{n\in\mathbb{N}}$ in E^* for which $\|y_n^*\| = 1$ the following inequality holds:

$$\sup_{n\in\mathbb{N}} |\langle M_\infty, y_n^*\rangle| < \infty, \quad \mathbb{P}\text{-almost surely.} \tag{1.245}$$

Moreover, if (i) or (ii) is satisfied, then there exists a sequence $(y_n^*)_{n\in\mathbb{N}}$ in E^* for which $\|y_n^*\| = 1$ such that $\|M_\infty\| = \sup\{|\langle M_\infty, y_n^*\rangle| : n \in \mathbb{N}\}$ $\mathbb{P}$-almost surely.

Note that in assertion (ii) the exceptional set may depend on the sequence $(y_n^*)_{n\in\mathbb{N}}$.

Proof. The implication (i) $\Longrightarrow$ (ii) being trivial we only need to prove (ii) $\Longrightarrow$ (i). Therefore, let B^{**} be the closed unit ball of E^{**}, and $F \subset$

E^{**} a linear subspace of E^{**} which is $\sigma(E^{**}, E^{*})$-closed and of finite co-dimension. The latter means that F is of the form

$$F = \left\{x^{**} \in E^{**} : \left\langle x_j^*, x^{**} \right\rangle = 0,\, 1 \leqslant j \leqslant n\right\}.$$

For $R > 0$ we consider the probability of events of the form $\{M_\infty \in RB + F\}$. Choose a sequence $\left(x_j^*\right)_{j\in\mathbb{N}}$ such that

$$\bigcap_{j=1}^{\infty} \left\{x^{**} \in E^{**} : \left\langle x_j^*, x^{**} \right\rangle = 0\right\} = \{0\},$$

and put $F_n = \bigcap_{j=1}^{n} \left\{x^{**} \in E^{**} : \left\langle x_j^*, x^{**} \right\rangle = 0\right\}$. Then we have

$$\bigcap_{n\in\mathbb{N}} \{M_\infty \in RB^{**} + F_n\} = \{M_\infty \in RB^{**}\}, \tag{1.246}$$

and hence

$$\begin{aligned} \mathbb{P}\left[M_\infty \in RB^{**}\right] &= \inf_{n\in\mathbb{N}} \mathbb{P}\left[M_\infty \in RB^{**} + F_n\right] \\ &= \inf_{n\in\mathbb{N}} \mathbb{P}\left[\sup\left\{\frac{\left\langle M_\infty, \sum_{j=1}^n \alpha_j x_j^*\right\rangle}{\left\|\sum_{j=1}^n \alpha_j x_j^*\right\|},\, \alpha_j \in \mathbb{Q},\, 1 \leqslant j \leqslant n\right\} \leqslant R\right]. \end{aligned} \tag{1.247}$$

Let $(y_n^*)_{n\in\mathbb{N}}$ be an enumeration of the countable family:

$$\left\{\frac{\sum_{j=1}^n \alpha_j x_j^*}{\left\|\sum_{j=1}^n \alpha_j x_j^*\right\|} : \alpha_j \in \mathbb{Q},\, 1 \leqslant j \leqslant n,\, n \in \mathbb{N}\right\}. \tag{1.248}$$

Then from (1.247) and (1.248) we infer:

$$\mathbb{P}\left[M_\infty \in RB^{**}\right] = \mathbb{P}\left[\sup_{n\in\mathbb{N}} \langle M_\infty, y_n^*\rangle \leqslant R\right]. \tag{1.249}$$

Notice that the sequence $(y_n^*)_{n\in\mathbb{N}}$ does not depend on the choice of $R > 0$. As a consequence we see that $\|M_\infty\| = \sup\{|\langle M_\infty, y_n^*\rangle| : n \in \mathbb{N}\}$ $\mathbb{P}$-almost surely.

This completes the proof of Theorem 1.19. □

Theorem 1.20. *Suppose that the stochastic integral*

$$M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)$$

belongs to E^{**} $\mathbb{P}$-almost surely: see Theorems 1.17 and 1.19. Suppose that E^* is separable for the norm-topology. The following assertions are equivalent:

(i) M_∞ *belongs to* E $\mathbb{P}$*-almost surely.*
(ii) For every $\varepsilon > 0$ *the following equality holds:*

$$\sup_{n\in\mathbb{N}} \inf_{\substack{\Lambda \subset \left\{y^* \in E^*,\, \|y^*\| \leqslant 1,\, d(y^*,0) < n^{-1}\right\} \\ \Lambda \text{ finite}}} \mathbb{P}\left[\max_{y^*\in\Lambda} |\langle M_\infty, y^*\rangle| < \varepsilon\right] = 1. \tag{1.250}$$

Here d denotes a metric on E^* which turns the dual unit ball B^* into a compact metric space. An appropriate metric d is given by

$$d(x^*, y^*) = \sum_{k=1}^{\infty} 2^{-k} |\langle x_k, x^* - y^*\rangle|$$

where the linear span of the sequence $(x_k)_{k\in\mathbb{N}}$ is dense in E, and where $\|x_k\| = 1$, $k \in \mathbb{N}$.

Proof. [Proof of Theorem 1.20.] Since the dual space is separable for the norm-topology, and $M_\infty \in E^{**}$ $\mathbb{P}$-almost surely, the equality in (1.250) can be rewritten as:

$$\sup_{n\in\mathbb{N}} \mathbb{P}\left[\sup_{\substack{\Lambda \subset \left\{y^* \in E^*,\, \|y^*\| \leqslant 1,\, d(y^*,0) < n^{-1}\right\} \\ \Lambda \text{ finite}}} \max_{y^*\in\Lambda} |\langle M_\infty, y^*\rangle| < \varepsilon\right]$$
$$= \sup_{n\in\mathbb{N}} \mathbb{P}\left[\sup_{\{y^*\in E^*,\, \|y^*\|\leqslant 1,\, d(y^*,0)<n^{-1}\}} |\langle M_\infty, y^*\rangle| < \varepsilon\right] = 1. \tag{1.251}$$

From (1.251) we see that assertion (ii) is equivalent to the equality:

$$\mathbb{P}\left[\inf_{n\in\mathbb{N}} \sup_{\{y^*\in E^*,\, \|y^*\|\leqslant 1,\, d(y^*,0)<n^{-1}\}} |\langle M_\infty, y^*\rangle| = 0\right] = 1. \tag{1.252}$$

If assertion (i) holds, then $M_\infty \in E$ $\mathbb{P}$-almost surely. Then the equality in (1.252) holds automatically, because a sequence $(y_n^*)_{n\in\mathbb{N}}$ converges for the weak*-topology to 0 if and only if $\lim_{n\to\infty} d(y_n^*, 0) = 0$. For the converse implication we invoke Theorem 1.17. If $(y_n^*)_{n\in\mathbb{N}}$ is a sequence in the dual unit ball which converges to 0 for the weak*-topology, then $\lim_{n\to\infty} d(y_n^*, 0) = 0$. By equality (1.252) it follows that $\lim_{n\to\infty} \langle M_\infty, y_n^*\rangle = 0$ $\mathbb{P}$-almost surely, where the exceptional set does not depend on the specific sequence $(y_n^*)_{n\in\mathbb{N}}$. An appeal to Theorem 1.17 then guarantees that $M_\infty \in E$ $\mathbb{P}$-almost surely.

This completes the proof of Theorem 1.20. □

For a concise formulation of some of the following results we introduce the space $L^2_{\text{weak}}(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ in the following definition. In such a space solutions to stochastic differential equations of the form (1.272) ought to be found.

Definition 1.8. For $s \in [\tau, t]$ the space $L^2_{\text{weak}}(\Omega, \mathcal{F}_s^\tau, \mathbb{P})$ is defined as follows. An element $Y(s)$ belongs to the vector space $L^2_{\text{weak}}(\Omega, \mathcal{F}_s^\tau, \mathbb{P})$ if it has the following properties.

(i) $Y(s) : E^* \to \mathbb{R}$ is $\mathbb{P}$-almost surely linear; we write $\langle Y(s), x^*\rangle$ for this action.
(ii) For every $x^* \in E^*$ the variable $\langle Y(s), x^*\rangle$ is $\mathcal{F}_s^\tau$-measurable, and it belongs to $L^2(\Omega, \mathcal{F}_s^\tau, \mathbb{P})$.
(iii) The supremum

$$\|Y(s)\|^2_{L^2_{\text{weak}}} := \sup\left\{\mathbb{E}\left[|\langle Y(s), x^*\rangle|^2\right] : x^* \in E^*, \|x^*\| \leqslant 1\right\}$$

is finite; equipped with the norm $\|Y(s)\|_{L^2_{\text{weak}}}$ the space $L^2_{\text{weak}}(\Omega, \mathcal{F}_s^\tau, \mathbb{P})$ is a Banach space.

Proposition 1.6. *Let $\mathcal{F}_s^{\tau,n}$, $\tau \leqslant s \leqslant t$, be the σ-field generated by*

$$(W_{H,j}(\rho) : \tau \leqslant \rho \leqslant s, 1 \leqslant j \leqslant n).$$

Let the sequence of $L^2_{\text{weak}}(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$-valued random variables $(M_n)_{n\in\mathbb{N}}$ be defined by the requirement that for all $x^* \in E^*$ the following equality holds $\mathbb{P}$-almost surely:

$$\langle M_n, x^*\rangle = \sum_{j=1}^n \int_\tau^t \mathbb{E}\left[\langle C(t,\rho)\sigma(\rho)e_j, x^*\rangle \mid \mathcal{F}_\rho^{\tau,n}\right] dW_{H,j}(\rho). \tag{1.253}$$

Suppose that for all $j \in \mathbb{N}$ and $x^* \in E^*$ the process $s \mapsto \langle C(t,s)\sigma(s)e_j, x^*\rangle$, $\tau \leqslant s \leqslant t$, is adapted to the filtration $\mathcal{F}_s^\tau = \sigma(W_{H,j}(\rho) : \tau \leqslant \rho \leqslant s, j \in \mathbb{N})$. Then with $M_0 = 0$ the following equality holds for all $x^* \in E^*$:

$$\sum_{n=1}^\infty \mathbb{E}\left[|\langle M_n - M_{n-1}, x^*\rangle|^2\right] = \mathbb{E}\left[\int_\tau^t \|\sigma(\rho)^* C(t,\rho)^* x^*\|_H^2 \, d\rho\right]. \tag{1.254}$$

From (1.254) it follows that for every $x^* \in E^*$ the L^2-limit $\langle M_\infty, x^*\rangle = L^2\text{-}\lim_{n\to\infty}\langle M_n, x^*\rangle$ exists, and that for every $F \in L^2(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ the vector $\mathbb{E}[FM_\infty]$ can be considered as E-valued. The vector $M_\infty \in L^2_{\text{weak}}(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ can be written as the stochastic integral:

$$M_\infty = \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho) = \sum_{j=1}^\infty \int_\tau^t C(t,\rho)\sigma(\rho)e_j \, dW_{H,j}(\rho). \tag{1.255}$$

Moreover, if

$$\sup_{x^*\in E^*,\,\|x^*\|\leqslant 1}\sum_{n=1}^{\infty}\left|\langle M_n - M_{n-1}, x^*\rangle\right|^2 < \infty, \quad \mathbb{P}\text{-almost surely,} \tag{1.256}$$

then the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)$ *belongs to* E^{**} $\mathbb{P}$*-almost surely. If, in addition, the conditional stochastic integrals*

$$\mathbb{E}\left[\int_\tau^t C(t,\rho)\sigma(\rho)e_j\,dW_{H,j}(\rho)\;\Big|\;\mathcal{F}_t^{\tau,n}\right],\quad 1\leqslant j\leqslant n,\ n\in\mathbb{N}, \tag{1.257}$$

belong to E*, then so does the stochastic integral* $\int_\tau^t C(t,\rho)\sigma(\rho)\,dW_H(\rho)$*. Similar results are true if the* $L^2_{\text{weak}}(\Omega,\mathcal{F}_t^\tau,\mathbb{P})$*-valued martingale* $n\mapsto M_n$ *is replaced by* $L^2_{\text{weak}}(\Omega,\mathcal{F}_t^\tau,\mathbb{P})$*-valued process* $n\mapsto\widetilde{M}_n$ *where*

$$\widetilde{M}_n = \sum_{j=1}^{n}\int_\tau^t C(t,\rho)\sigma(\rho)e_j\,dW_{H,j}(\rho).$$

In particular, with $\widetilde{M}_0 = 0$ *the following equality holds for all* $x^*\in E^*$*:*

$$\sum_{n=1}^{\infty}\mathbb{E}\left[\left|\left\langle \widetilde{M}_n - \widetilde{M}_{n-1}, x^*\right\rangle\right|^2\right] = \mathbb{E}\left[\int_\tau^t \|\sigma(\rho)^*C(t,\rho)^*x^*\|_H^2\,d\rho\right]. \tag{1.258}$$

The following lemma gives a sufficient condition in order that $M_\infty\in E$ $\mathbb{P}$-almost surely.

Lemma 1.6. *Put* $\varphi_N(\rho) = \tau + \dfrac{t-\tau}{2^N}\left\lfloor\dfrac{(\rho-\tau)2^N}{t-\tau}\right\rfloor$*, and suppose that*

$$\lim_{N\to\infty}\mathbb{E}\left[\left\|M_\infty - \sum_{j=1}^{N}\int_\tau^t C\left(t,\varphi_N(\rho)\right)\sigma\left(\varphi_N(\rho)\right)\,dW_{H,j}(\rho)\right\|\right] = 0. \tag{1.259}$$

Then the vector M_∞ *belongs to* E $\mathbb{P}$*-almost surely.*

Proof. By a standard result from integration theory there exists a subsequence such that

$$\lim_{k\to\infty}\left\|M_\infty - \sum_{j=1}^{N_k}\int_\tau^t C\left(t,\varphi_{N_k}(\rho)\right)\sigma\left(\varphi_{N_k}(\rho)\right)\,dW_{H,j}(\rho)\right\| = 0,\ \mathbb{P}\text{-almost surely.} \tag{1.260}$$

By definition of the functions $\varphi_N(\rho)$ and the definition of stochastic integral, we see that the integrals $\int_\tau^t C\left(t,\varphi_N(\rho)\right)\sigma\left(\varphi_N(\rho)\right)\,dW_{H,j}(\rho)$ belong to E. The equality in (1.260) shows that $M_\infty\in E$ $\mathbb{P}$-almost surely. □

In fact, we conjecture that if the functional $p: E^* \to [0, \infty)$ defined by

$$p(x^*) = \mathbb{E}\left[\int_\tau^t \|\sigma(\rho)^* C(t,\rho)^* x^*\|_H^2 \, d\rho\right] \tag{1.261}$$

is sequentially continuous when E^* is endowed with the weak*-topology, then the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)$ belongs to E $\mathbb{P}$-almost surely.

Proof. [Proof of Proposition 1.6.] The equality in (1.253) is a consequence of the following string of equalities, which are self-explanatory:

$$\begin{aligned}
\langle M_n, x^* \rangle &= \sum_{j=1}^n \int_\tau^t \mathbb{E}\left[\langle C(t,\rho)\sigma(\rho)e_j, x^*\rangle \mid \mathcal{F}_\rho^{\tau,n}\right] dW_{H,j}(\rho) \\
&= \sum_{j=1}^n \int_\tau^t \mathbb{E}\left[\langle C(t,\rho)\sigma(\rho)e_j, x^*\rangle \mid \mathcal{F}_t^{\tau,n}\right] dW_{H,j}(\rho) \\
&= \sum_{j=1}^n \int_\tau^t \mathbb{E}\left[\langle C(t,\rho)\sigma(\rho)e_j, x^*\rangle \, dW_{H,j}(\rho) \mid \mathcal{F}_t^{\tau,n}\right] \\
&= \sum_{j=1}^\infty \int_\tau^t \mathbb{E}\left[\langle C(t,\rho)\sigma(\rho)e_j, x^*\rangle \, dW_{H,j}(\rho) \mid \mathcal{F}_t^{\tau,n}\right] \\
&= \mathbb{E}\left[\left\langle \int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho), x^* \right\rangle \mid \mathcal{F}_t^{\tau,n}\right].
\end{aligned} \tag{1.262}$$

By employing martingale convergence in (1.262) the equality in (1.253) follows. The claim that for every $F \in L^2(\Omega, \mathcal{F}_t^\tau, \mathbb{P})$ the vector $\mathbb{E}[FM_\infty]$ can be considered as being E-valued in Proposition 1.6 follows from Lemma 1.4 formula (1.213).

If (1.256) holds, then, uniformly on the dual unit ball, we have $\langle M_\infty, x^* \rangle = \lim_{n\to\infty} \langle M_n, x^* \rangle$. From (1.256) we also see that $M_n \in E^{**}$ $\mathbb{P}$-almost surely for all $n \in \mathbb{N}$. Consequently, $M_\infty \in E^{**}$ $\mathbb{P}$-almost surely. If (1.257) is satisfied, then these arguments show that $M_\infty \in E$ $\mathbb{P}$-almost surely.

The assertions concerning the process $n \mapsto \widetilde{M}_n$ can be proved in more or less the same manner; instead of a martingale argument one employs a Hilbert space argument to prove equality (1.258).

This completes the proof of Proposition 1.6. □

In Definition 1.9 and Theorem 1.22 we assume that for every $s \in [\tau, t]$ the mapping $y \mapsto \sigma(s, y)$, $y \in E$, is defined on the space E, and attains its values in $L(H, E)$, i.e. the space of all bounded linear operators from the

Hilbert space H to the Banach space E. Similarly, for every $s \in [\tau, t]$ the mapping $y \mapsto b(s,y)$ is defined on E, and attains its values in E.

Definition 1.9. A solution to (1.195) in E is a process $s \mapsto X^{\tau,x}(s)$, $s \in [\tau, t]$, such that for every $s \in [\tau, t]$ the stochastic vector $X^{\tau,x}(s)$ belongs to the space E, and the following identity holds $\mathbb{P}$-almost surely for all $s \in [\tau, t]$:

$$X^{\tau,x}(s) = C(s,\tau)x + \int_\tau^s C(s,\rho)\sigma\left(\rho, X^{\tau,x}(\rho)\right) dW_H(\rho) \\ + \int_\tau^s C(s,\rho) b\left(\rho, X^{\tau,x}(\rho)\right) d\rho, \tag{1.263}$$

It is not so easy to work in the space E or $L^2_{\mathrm{weak}}\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}\right)$ directly. In the latter space we use a supremum-norm over the dual unit ball. Instead of taking the supremum-norm we can also take a Borel measure μ on E^* and look at the following subspace of E:

$$\left\{x \in E : \|x\|_\mu^2 = \int \left|\langle x, x^*\rangle\right|^2 d\mu(x^*) < \infty\right\} \tag{1.264}$$

Denote by E_μ the completion of the space in (1.264) with respect to $\|\cdot\|_\mu$. Denote by $L^2_\mu\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}\right)$ the space of stochastic $\mathbb{E}_\mu$-valued vectors X such that for μ-almost all $x^* \in E^*$ the variable $\langle X, x^*\rangle$ is $\mathcal{F}_t^\tau$-measurable, and such that $\|X\|^2_{L^2_\mu} = \mathbb{E}\left[\int \left|\langle X, x^*\rangle\right|^2 d\mu(x^*)\right] < \infty$.

Proposition 1.7. *Suppose that*

$$\mathbb{E}\left[\int\int_\tau^t \left\|\sigma(\rho)^* C(t,\rho)^* x^*\right\|_H^2 d\rho\, d\mu(x^*)\right] < \infty.$$

Then the stochastic integral $\int_\tau^t C(t,\rho)\sigma(\rho)\, dW_H(\rho)$ *belongs to* $L^2_\mu\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}\right)$.

Proof. From (1.258) we obtain

$$\sum_{n=1}^\infty \int \mathbb{E}\left[\left|\left\langle \widetilde{M}_n - \widetilde{M}_{n-1}, x^*\right\rangle\right|^2\right] d\mu(x^*) \\ = \mathbb{E}\left[\int\int_\tau^t \left\|\sigma(\rho)^* C(t,\rho)^* x^*\right\|_H^2 d\rho\, d\mu(x^*)\right] < \infty \tag{1.265}$$

From (1.265) we deduce that the sequence $\left(\widetilde{M}_n\right)_{n\in\mathbb{N}}$ converges to M_∞ in the space $L^2_\mu\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}\right)$. □

In Definition 1.10 and Theorem 1.21 below we assume that for every $s \in [\tau, t]$ the mapping $Y(s) \mapsto \sigma(s, Y(s))$ is defined on the space $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$, and attains its values in $L(H, E)$, i.e. the space of all bounded linear operators from the Hilbert space H to the Banach space E. Similarly, for every $s \in [\tau, t]$ the mapping $Y(s) \mapsto b(s, Y(s))$ is defined on $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$, and attains its values in E.

Definition 1.10. A solution to (1.195) is a process $s \mapsto X^{\tau,x}(s)$, $s \in [\tau, t]$, such that for every $s \in [\tau, t]$ the stochastic vector $X^{\tau,x}(s)$ belongs to the space $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$, and the following identity holds $\mathbb{P}$-almost surely for all $s \in [\tau, t]$:

$$\begin{aligned} X^{\tau,x}(s) = C(s,\tau)x + \int_\tau^s C(s,\rho)\sigma\left(\rho, X^{\tau,x}(\rho)\right) dW_H(\rho) \\ + \int_\tau^s C(s,\rho) b\left(\rho, X^{\tau,x}(\rho)\right) d\rho. \end{aligned} \tag{1.266}$$

Next we give an existence and uniqueness theorem for solutions to stochastic differential equations with values in $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$.

Theorem 1.21. *Assume that the coefficients $\sigma(s, Y(s))$, and $b(s, Y(s))$ satisfy the following Lipschitz conditions. There exist functions $c_1(s)$ and $c_2(s)$, $s \in [\tau, t]$, such that for all $y^* \in E^*$, and all $Y_1(s)$, $Y_2(s) \in L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$ the following inequalities hold:*

$$\mathbb{E}\left[\left\|\left(\sigma(s, Y_2(s))^* - \sigma(s, Y_1(s))^*\right) y^*\right\|_H^2\right] \leqslant c_1(s)^2 \left\|Y_2(s) - Y_1(s)\right\|^2_{L^2_{\text{weak}}} \left\|y^*\right\|^2, \tag{1.267}$$

and

$$\mathbb{E}\left[\left|\left\langle b(s, Y_2(s)) - b(s, Y_1(s)), y^*\right\rangle\right|\right] \leqslant c_2(s) \left\|Y_2(s) - Y_1(s)\right\|_{L^2_{\text{weak}}} \left\|y^*\right\|. \tag{1.268}$$

Fix $x \in E$ and suppose that

$$\int_\tau^t \sup_{s \in [\rho, t]} \left\|C(s,\rho)\right\|^2 c_1(\rho)^2 d\rho < \infty, \quad \textit{and} \quad \int_\tau^t \sup_{s \in [\tau, t]} \left\|C(s,\rho)\right\|^2 c_2(\rho)^2 \, d\rho < \infty.$$

In addition, suppose that for $\tau \leqslant s \leqslant t$ the stochastic integral

$$\int_\tau^s C(s,\rho)\sigma(\rho, x)\, dW_H(\rho)$$

belongs to $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$. Then the equation in (1.263) possesses a unique solution $s \mapsto X^{\tau,x}(s)$ and $X^{\tau,x}(s)$ belongs to $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$ for all $s \in$

$[\tau, t]$. *If the process* $s \mapsto X^{\tau,x}(s)$ *can be realized in such a way that for all sequences* $(y_n^*)_{n\in\mathbb{N}} \in c_0(\mathbb{N}, E^*)$ *the inequalities*

$$0 \leqslant \sup_{n\in\mathbb{N}} \langle X^{\tau,x}(s), y_n^*\rangle < \infty \tag{1.269}$$

holds on an almost sure event Ω' *which does not depend on the choice of the sequence* $(y_n^*)_{n\in\mathbb{N}} \in c_0(\mathbb{N}, E^*)$. *Then the solution* $s \mapsto X^{\tau,x}(s)$ *belongs to* E $\mathbb{P}$*-almost surely. Here a sequence* $(y_n^*)_{n\in\mathbb{N}} \subset E^*$ *belongs to* $c_0(\mathbb{N}, E^*)$ *if* $\lim_{n\to\infty} \langle y, y_n^*\rangle = 0$ *for every* $y \in E$.

Proof. We will construct a solution to the equation in (1.266). To this end we introduce the functions $\varphi_n : [\tau, t] \to [\tau, t]$ by

$$\varphi_n(\rho) = \tau + \frac{t-\tau}{2^n}\left\lfloor \frac{(\rho-\tau)2^n}{t-\tau}\right\rfloor. \tag{1.270}$$

Notice that $\varphi_0(\rho) = \tau$, $\tau \leqslant \rho < t$, and

$$\rho - \frac{t-\tau}{2^n} \leqslant \varphi_n(\rho) \leqslant \rho.$$

By induction we define the sequence of E-valued stochastic processes

$$\{X_n^{\tau,x}(s) : \tau \leqslant s \leqslant t,\ n \in \mathbb{N}\} \tag{1.271}$$

as follows. For $\tau \leqslant s \leqslant t$ we write $X_0^{\tau,x}(s) = C(s,\tau)x$, and the process $X_1^{\tau,x}(s)$ is defined by

$$\begin{aligned} X_1^{\tau,x}(s) &= C(s,\tau)x + \int_\tau^s C(s,\rho)\sigma(\rho,x)\, dW_H(\rho) + \int_\tau^s C(s,\rho)b(\rho,x)\, d\rho \\ &= C(s,\tau)x + \int_\tau^s C(s,\rho)\sigma\left(\rho, X_0^{\tau,x}(\varphi_0(\rho))\right) dW_H(\rho) \\ &\quad + \int_\tau^s C(s,\rho) b\left(\rho, X_0^{\tau,x}(\varphi_0(\rho))\right) d\rho. \end{aligned} \tag{1.272}$$

Then we define the process $s \mapsto X_{n+1}^{\tau,x}(s)$ in terms of $\rho \mapsto X_n^{\tau,x}(\varphi_n(\rho))$ as follows:

$$\begin{aligned} X_{n+1}^{\tau,x}(s) &= C(s,\tau)x + \int_\tau^s C(s,\rho)\sigma\left(\rho, X_n^{\tau,x}(\varphi_n(\rho))\right) dW_H(\rho) \\ &\quad + \int_\tau^s C(s,\rho)b\left(\rho, X_n^{\tau,x}(\varphi_n(\rho))\right) d\rho. \end{aligned} \tag{1.273}$$

Since for $\tau \leqslant s_1 \leqslant s_2 \leqslant t$ and $y \in E$ we have

$$\begin{aligned} &\int_{s_1}^{s_2} C(s_2,\rho)\,\sigma(\rho,y)\,dW_H(\rho) \\ &= \int_\tau^{s_2} C(s_2,\rho)\,\sigma(\rho,y)\,dW_H(\rho) - \int_\tau^{s_1} C(s_2,\rho)\,\sigma(\rho,y)\,dW_H(\rho) \\ &= \int_\tau^{s_2} C(s_2,\rho)\,\sigma(\rho,y)\,dW_H(\rho) - C(s_2,s_1)\int_\tau^{s_1} C(s_1,\rho)\,\sigma(\rho,y)\,dW_H(\rho) \end{aligned} \tag{1.274}$$

it follows by induction that the processes $s \mapsto X_n^{\tau,x}(s)$, $s \in [\tau, t]$, take their values in E $\mathbb{P}$-almost surely. Let $x^* \in E^*$. Then from (1.273) we get:

$$\begin{aligned}
&\left\langle X_{n+1}^{\tau,x}(s) - X_n^{\tau,x}(s), x^* \right\rangle \qquad (1.275)\\
&= \left\langle \int_\tau^s C(s,\rho) \left(\sigma\left(\rho, X_n^{\tau,x}(\varphi_n(\rho))\right) - \sigma\left(\rho, X_{n-1}^{\tau,x}(\varphi_{n-1}(\rho))\right)\right) dW_H(\rho), x^* \right\rangle\\
&\quad + \left\langle \int_\tau^s C(s,\rho) \left(b\left(\rho, X_n^{\tau,x}(\varphi_n(\rho))\right) - b\left(\rho, X_{n-1}^{\tau,x}(\varphi_{n-1}(\rho))\right)\right) d\rho, x^* \right\rangle,
\end{aligned}$$

and hence

$$\begin{aligned}
&\left|\left\langle X_{n+1}^{\tau,x}(s) - X_n^{\tau,x}(s), x^* \right\rangle\right|^2 \qquad (1.276)\\
&\leqslant 2\left|\left\langle \int_\tau^s C(s,\rho) \left(\sigma\left(\rho, X_n^{\tau,x}(\varphi_n(\rho))\right) - \sigma\left(\rho, X_{n-1}^{\tau,x}(\varphi_{n-1}(\rho))\right)\right) dW_H(\rho), x^* \right\rangle\right|^2\\
&\quad + 2\left|\left\langle \int_\tau^s C(s,\rho) \left(b\left(\rho, X_n^{\tau,x}(\varphi_n(\rho))\right) - b\left(\rho, X_{n-1}^{\tau,x}(\varphi_{n-1}(\rho))\right)\right) d\rho, x^* \right\rangle\right|^2.
\end{aligned}$$

For brevity we write $Y_n^{\tau,x}(\rho) = X_n^{\tau,x}(\varphi_n(\rho))$. From (1.276) and (1.212) in Lemma 1.4 we deduce

$$\begin{aligned}
&\mathbb{E}\left[\left|\left\langle X_{n+1}^{\tau,x}(s) - X_n^{\tau,x}(s), x^* \right\rangle\right|^2\right]\\
&\leqslant 2\mathbb{E}\left[\left|\left\langle \int_\tau^s C(s,\rho) \left(\sigma\left(\rho, Y_n^{\tau,x}(\rho)\right) - \sigma\left(\rho, Y_{n-1}^{\tau,x}(\rho)\right)\right) dW_H(\rho), x^* \right\rangle\right|^2\right]\\
&\quad + 2\mathbb{E}\left[\left(\int_\tau^s \left|\left\langle C(s,\rho) \left(b\left(\rho, Y_n^{\tau,x}(\rho)\right) - b\left(\rho, Y_{n-1}^{\tau,x}(\rho)\right)\right), x^* \right\rangle\right| d\rho\right)^2\right]\\
&= 2\mathbb{E}\left[\int_\tau^s \left\|\left(\sigma\left(\rho, Y_n^{\tau,x}(\rho)\right)^* - \sigma\left(\rho, Y_{n-1}^{\tau,x}(\rho)\right)^*\right) C(s,\rho)^* x^*\right\|_H^2 d\rho\right]\\
&\quad + 2\mathbb{E}\left[\left(\int_\tau^s \left|\left\langle C(s,\rho) \left(b\left(\rho, Y_n^{\tau,x}(\rho)\right) - b\left(\rho, Y_{n-1}^{\tau,x}(\rho)\right)\right), x^* \right\rangle\right| d\rho\right)^2\right].
\end{aligned} \qquad (1.277)$$

Inserting the inequalities (1.267) and (1.268) into (1.277) yields:

$$\begin{aligned}
&\mathbb{E}\left[\left|\left\langle X_{n+1}^{\tau,x}(s) - X_n^{\tau,x}(s), x^* \right\rangle\right|^2\right]\\
&\leqslant 2\int_\tau^s c_1(\rho)^2 \left\|X_n^{\tau,x}(\varphi_n(\rho)) - X_{n-1}^{\tau,x}(\varphi_{n-1}(\rho))\right\|_{L^2_{\text{weak}}}^2 \left\|C(s,\rho)^* x^*\right\|^2 d\rho\\
&\quad + 2\left(\int_\tau^s c_2(\rho) \left\|X_n^{\tau,x}(\varphi_n(\rho)) - X_{n-1}^{\tau,x}(\varphi_{n-1}(\rho))\right\|_{L^2_{\text{weak}}} \left\|C(s,\rho)^* x^*\right\| d\rho\right)^2\\
&\leqslant 2\int_\tau^s \left(c_1(\rho)^2 + (s-\tau)c_2(\rho)^2\right)
\end{aligned}$$

$$\left\|X_n^{\tau,x}\left(\varphi_n(\rho)\right)-X_{n-1}^{\tau,x}\left(\varphi_{n-1}(\rho)\right)\right\|_{L^2_{\rm weak}}^2\left\|C(s,\rho)^*x^*\right\|^2 d\rho \tag{1.278}$$

$$\leqslant 2\int_\tau^s \left(c_1(\rho)^2+(s-\tau)c_2(\rho)^2\right)\left\|C(s,\rho)\right\|^2\left\|x^*\right\|^2$$

$$\left\|X_n^{\tau,x}\left(\varphi_n(\rho)\right)-X_{n-1}^{\tau,x}\left(\varphi_{n-1}(\rho)\right)\right\|_{L^2_{\rm weak}}^2 d\rho. \tag{1.279}$$

Put

$$\chi(\rho)=2\left(c_1(\rho)^2+(t-\tau)c_2(\rho)^2\right)\sup_{\rho\leqslant s\leqslant t}\left\|C(s,\rho)\right\|^2 \quad\text{and}$$

$$\Phi_n(\rho)=\left\|X_n^{\tau,x}\left(\varphi_n(\rho)\right)-X_{n-1}^{\tau,x}\left(\varphi_{n-1}(\rho)\right)\right\|_{L^2_{\rm weak}}^2. \tag{1.280}$$

Then from (1.279) we see:

$$\Phi_{n+1}(s)\leqslant\int_\tau^s\chi(\rho)\Phi_n(\rho)\,d\rho,\quad \tau\leqslant s\leqslant t. \tag{1.281}$$

By induction with respect to k, $0\leqslant k\leqslant n-2$, from (1.281) we infer

$$\Phi_n(s)\leqslant\frac{1}{k!}\int_\tau^s\chi\left(\rho_{k+1}\right)\left(\int_{\rho_{k+1}}^s\chi(\rho)\,d\rho\right)^k\Phi_{n-k-1}\left(\rho_{k+1}\right)\,d\rho_{k+1}. \tag{1.282}$$

With $k=n-2$ we get

$$\Phi_n(s)\leqslant\frac{1}{(n-2)!}\int_\tau^s\chi\left(\rho'\right)\left(\int_{\rho'}^s\chi(\rho)\,d\rho\right)^{n-2}\Phi_1\left(\rho'\right)\,d\rho'. \tag{1.283}$$

From (1.283) it follows that $\sum_{n=2}^{\infty}\Phi_n(s)\infty$. Then from (1.280) it follows that the limit

$$X^{\tau,x}(s)=L^2_{\rm weak}\text{-}\lim_{n\to\infty}X_n^{\tau,x}\left(\varphi_n(\rho)\right) \tag{1.284}$$

exists. The equality in (1.273) then implies that the process $s\mapsto X^{\tau,x}(s)$ satisfies the stochastic differential equation in (1.266).

Let $X_1^{\tau,x}(s)$ and $X_2^{\tau,x}(s)$ be two solutions to the equation in (1.266). Then the above arguments applied to the equality

$$X_2^{\tau,x}(s)-X_1^{\tau,x}(s)=\int_\tau^s C(s,\rho)\left(\sigma\left(\rho,X_2^{\tau,x}(\rho)\right)-\sigma\left(\rho,X_1^{\tau,x}(\rho)\right)\right)\,dW_H(\rho)$$
$$+\int_\tau^s C(s,\rho)\left(b\left(\rho,X_2^{\tau,x}(\rho)\right)-b\left(\rho,X_1^{\tau,x}(\rho)\right)\right)\,dW_H(\rho)$$

shows that the $\mathbb{P}$-almost sure equality $X_2^{\tau,x}(s)=X_1^{\tau,x}(s)$.

If (1.269) is satisfied for all sequences $(y_n^*)_{n\in\mathbb{N}}\in c_0\left(\mathbb{N},E^*\right)$, then Theorem 1.18 entails the inclusion $X^{\tau,x}(s)\in E$ $\mathbb{P}$-almost surely.

This concludes the proof of Theorem 1.21. □

In the following theorem we prove a result similar to the one in Theorem 1.21, but in the space $L^2_\mu(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$ instead of $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$.

Theorem 1.22. *Assume that the coefficients $\sigma(s, y)$, and $b(s, y)$ satisfy the following Lipschitz conditions. There exist functions $c_1(s)$ and $c_2(s)$, $s \in [\tau, t]$, such that for all $y^* \in E$, and all y_1, $y_2 \in E$ the following inequalities hold:*

$$\left\| \left(\sigma(s, y_2)^* - \sigma(s, y_1)^*\right) y^* \right\|_H \leqslant c_1(s) \left\| y_2 - y_1 \right\|_\mu \left\| y^* \right\|, \tag{1.285}$$

and

$$\left| \left\langle b(s, y_2) - b(s, y_1), y^* \right\rangle \right| \leqslant c_2(s) \left\| y_2 - y_1 \right\|_\mu \left\| y^* \right\|. \tag{1.286}$$

Suppose that

$$\limsup_{\delta \downarrow 0} \sup_{s_1 \in [\tau, t]} \sup_{s_2 \in [s_1, t],\, s_2 - s_1 < \delta} \int_{s_1}^{s_2} \int \left\| C(s_2, \rho)^* x^* \right\|^2 d\mu(x^*)\, c_1(\rho)^2 d\rho < \frac{1}{2},$$

and

$$\limsup_{\delta \downarrow 0} \sup_{s_1 \in [\tau, t]} \sup_{s_2 \in [s_1, t],\, s_2 - s_1 < \delta} \int_{s_1}^{s_2} \int \left\| C(s_2, \rho)^* x^* \right\|^2 d\mu(x^*)\, c_2(\rho)^2\, d\rho < \infty. \tag{1.287}$$

Then the equation in (1.263) possesses a unique solution $s \mapsto X^{\tau,x}(s)$ and $X^{\tau,x}(s)$ belongs to $L^2_{\text{weak}}(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$ for all $s \in [\tau, t]$, provided that for every $(s, x) \in [\tau, t] \times E$ the stochastic vector $\int_\tau^s C(s, \rho)\sigma(\rho, x)\, dW_H(\rho)$ belongs to $L^2_\mu(\Omega, \mathcal{F}^\tau_s, \mathbb{P})$.

The latter means that

$$\mathbb{E}\left[\int \int_\tau^s \left\| \sigma(\rho, x)^* C(s, \rho)^* x^* \right\|_H^2 d\rho\, d\mu(x^*)\right] < \infty.$$

Proof. The proof of Theorem 1.22 follows the same pattern as that of Theorem 1.21. Again we construct the sequence $\{X^{\tau,x}_n(s) : \tau \leqslant s \leqslant t,\, n \in \mathbb{N}\}$ in (1.271) satisfying (1.272) and (1.273). Inserting the inequalities (1.285) and (1.286) into (1.277) yields the following inequality:

$$\begin{aligned}
&\mathbb{E}\left[\left|\left\langle X^{\tau,x}_{n+1}(s) - X^{\tau,x}_n(s), x^* \right\rangle\right|^2\right] \\
&\leqslant 2 \int_\tau^s c_1(\rho)^2 \left\| X^{\tau,x}_n(\varphi_n(\rho)) - X^{\tau,x}_{n-1}(\varphi_{n-1}(\rho)) \right\|^2_{L^2_\mu} \left\| C(s, \rho)^* x^* \right\|^2 d\rho \\
&\quad + 2 \left(\int_\tau^s c_2(\rho) \left\| X^{\tau,x}_n(\varphi_n(\rho)) - X^{\tau,x}_{n-1}(\varphi_{n-1}(\rho)) \right\|_{L^2_\mu} \left\| C(s, \rho)^* x^* \right\| d\rho\right)^2
\end{aligned}$$

$$\leqslant 2\int_\tau^s \left(c_1(\rho)^2+(s-\tau)c_2(\rho)^2\right)$$
$$\left\|X_n^{\tau,x}\left(\varphi_n(\rho)\right)-X_{n-1}^{\tau,x}\left(\varphi_{n-1}(\rho)\right)\right\|_{L^2_\mu}^2\left\|C(s,\rho)^*x^*\right\|^2 d\rho. \tag{1.288}$$

The inequality in (1.288) is the same as the one in (1.278) except that here we write $\|\cdot\|_{L^2_\mu}$ instead of $\|\cdot\|_{L^2_{\text{weak}}}$. Instead of $\chi(\rho)$ and $\Phi_n(\rho)$ as in (1.280) we know introduce the functions:

$$\chi_\mu(\rho,s)=2\left(c_1(\rho)^2+(s-\tau)c_2(\rho)^2\right)\int\left\|C(s,\rho)^*x^*\right\|^2 d\mu\left(x^*\right)\quad\text{and}$$
$$\Phi_{\mu,n}(\rho)=\left\|X_n^{\tau,x}\left(\varphi_n(\rho)\right)-X_{n-1}^{\tau,x}\left(\varphi_{n-1}(\rho)\right)\right\|_{L^2_\mu}^2. \tag{1.289}$$

Then we choose $\delta>0$ so small that

$$\sup_{\rho\in[s_1,s_2]}\int_\rho^{s_2}\chi_\mu(\rho,s)\,ds\leqslant 1-\eta<1, \tag{1.290}$$

for some $\eta>0$ and for all s_1, $s_2\in[\tau,t]$ such that $0\leqslant s_2-s_1\leqslant\delta$. By the assumptions in (1.287) such a choice is possible. Integrating (1.288) relative to $d\mu\left(x^*\right)$ yields:

$$\Phi_{\mu,n+1}(s)\leqslant\int_\tau^s\chi_\mu(\rho,s)\Phi_{\mu,n}(\rho)\,d\rho. \tag{1.291}$$

Next we define the sequence of functions $\chi_{\mu,n}(\rho,s)$, $\tau\leqslant\rho\leqslant s\leqslant t$, $n\in\mathbb{N}$, as follows:

$$\chi_{\mu,1}(\rho,s)=\chi_\mu(\rho,s),\quad \chi_{\mu,2}(\rho,s)=\int_\rho^s\chi_\mu\left(\rho,\rho_1\right)\chi_\mu\left(\rho_1,s\right)\,d\rho_1,\quad\text{and}$$
$$\chi_{\mu,n}(\rho,s) \tag{1.292}$$
$$=\int\cdots\int_{\rho<\rho_{n-1}<\cdots<\rho_1<s} d\rho_{n-1}\dots d\rho_1\chi_\mu\left(\rho,\rho_{n-1}\right)\prod_{j=2}^{n-1}\chi_\mu\left(\rho_j,\rho_{j-1}\right)\chi_\mu\left(\rho_1,s\right)$$

for $n\geqslant 3$. The function $\chi_{\mu,n}(\rho,s)$ is kind of a generalized n-fold convolution product of $\chi_\mu(\rho,s)$ with itself. From the choice of $\delta>0$ and $\eta>0$ we see that

$$\int_{s_1}^{s_2}\chi_{\mu,n}\left(s_1,s\right)\,ds\leqslant(1-\eta)^n,\quad\text{for } 0\leqslant s_2-s_1\leqslant\delta,\ n\in\mathbb{N}. \tag{1.293}$$

Moreover, it is not difficult to show that

$$\Phi_{\mu,n+1}(s)\leqslant\int_\tau^s\chi_{\mu,n}(\rho,s)\Phi_{\mu,1}(\rho)\,d\rho,\quad n\geqslant 1. \tag{1.294}$$

From (1.293) and (1.293) we get:

$$\int_\tau^{\tau+\delta}\Phi_{\mu,n+1}(s)\,ds\leqslant(1-\eta)^n\int_\tau^{\tau+\delta}\Phi_{\mu,1}(s)\,ds. \tag{1.295}$$

From (1.295) we infer that

$$\sum_{n=1}^{\infty}\int_{\tau}^{\tau+\delta}\Phi_{\mu,n}(s)\,ds<\infty. \tag{1.296}$$

From (1.296) it follows that $\sum_{n=1}^{\infty}\Phi_{\mu,n}(s)<\infty$ for almost all $s\in[\tau,\tau+\delta]$. This means that for almost all $s\in[\tau,\tau+\delta]$ the process $X^{\tau,x}(s)=L^2_{\mu}\text{-}\lim_{n\to\infty}X_n^{\tau,x}(\varphi_n(s))$ exists, and that for such s the equality in (1.266) holds; i.e.

$$\begin{aligned}X^{\tau,x}(s)=C(s,\tau)x&+\int_{\tau}^{s}C(s,\rho)\sigma\left(\rho,X^{\tau,x}(\rho)\right)\,dW_H(\rho)\\&+\int_{\tau}^{s}C(s,\rho)b\left(\rho,X^{\tau,x}(\rho)\right)\,d\rho.\end{aligned} \tag{1.297}$$

Then we use continuity in $s\in[\tau,\tau+\delta]$ to prove that (1.297) holds for all $s\in[\tau,\tau+\delta]$. Once this is done we repeat the previous argument on the interval $[\tau+\delta,\tau+2\delta]$ with initial value $X^{\tau,x}(\tau+\delta)$ instead of x. In finite many steps we construct a (unique) solution on the interval $[\tau,t]$.

All this completes the proof of Theorem 1.22. □

Remark 1.6. Let the operator families $C_1(\rho)$, and $C_2(\rho)$, $\rho\in[\tau,t]$, consist of operators from the Hilbert space H to the Banach space E with appropriate measurability properties. Instead of the Lipschitz conditions (1.285) and (1.286) we could have taken conditions of the form:

$$\left\|\left(\sigma\left(s,y_2\right)^*-\sigma\left(s,y_1\right)^*\right)y^*\right\|_H\leqslant\left\|y_2-y_1\right\|_{\mu}\left\|C_1(s)^*y^*\right\|, \tag{1.298}$$

and

$$\left|\left\langle b\left(s,y_2\right)-b\left(s,y_1\right),y^*\right\rangle\right|\leqslant\left\|y_2-y_1\right\|_{\mu}\left\|C_2(s)^*y^*\right\|. \tag{1.299}$$

In order to get conclusions like in Theorem 1.22 the conditions in (1.287) have to pe replaced with

$$\limsup_{\delta\downarrow 0}\sup_{s_1\in[\tau,t]}\sup_{s_2\in[s_1,t],\,s_2-s_1<\delta}\int_{s_1}^{s_2}\int\left\|C_1(\rho)^*C\left(s_2,\rho\right)^*x^*\right\|_H^2\,d\mu\left(x^*\right)\,d\rho<\frac{1}{2},$$

and

$$\limsup_{\delta\downarrow 0}\sup_{s_1\in[\tau,t]}\sup_{s_2\in[s_1,t],\,s_2-s_1<\delta}\int_{s_1}^{s_2}\int\left\|C_2(\rho)^*C\left(s_2,\rho\right)^*x^*\right\|_H^2\,d\mu\left(x^*\right)\,d\rho<\infty. \tag{1.300}$$

We conclude this introduction by collecting some well-known and not so well-known results about martingales and stopping times for time-homogeneous Markov processes.

1.3 Martingales

In this section we recall some interesting facts about martingales. This material is taken from [Van Casteren (2002)].

(1) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(\mathcal{F}_t : t \geqslant 0)$ be a filtration on Ω; i.e. $s < t$ implies $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$. Suppose that $\mathcal{F}$ is the σ-field generated by $\mathcal{F}_t$, $t \geqslant 0$. Moreover, let Y belong to $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Put $M(t) = \mathbb{E}\left[Y \mid \mathcal{F}_t\right]$. Then the process is the standard example of a closed martingale. This martingale is closed, because $Y = L^1\text{-}\lim_{t\to\infty} M(t)$. This limit is also an $\mathbb{P}$-almost sure limit.

(2) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $W(t) : \Omega \to \mathbb{R}^d$ be Brownian motion starting at zero. Then the process $W(t)$, $t \geqslant 0$, is a martingale. The same is true for the process $t \mapsto |W(t)|^2 - dt$.

(3) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $W(t) : \Omega \to \mathbb{R}^d$ be Brownian motion. Let $\{H(t) : t \geqslant 0\}$ be a predictable process. This means that $H(t)$ is $\mathcal{F}_t$-measurable for each $t \geqslant 0$, and that the mapping $(t, \omega) \mapsto H(t, \omega)$ is measurable with respect to the σ-field generated by

$$\left\{\mathbf{1}_{(s,t]} \otimes 1_A : A\ \mathcal{F}_s\text{-measurable},\ s < t\right\}.$$

Suppose that $\mathbb{E}\left[\int_0^t |H(s)|^2\, ds\right] < \infty$ for all $t > 0$. Then the process $t \mapsto \int_0^t H(s)\, dW(s)$ is a martingale in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. If we only assume that the expression $\int_0^t |H(s)|^2\, ds$ are finite $\mathbb{P}$-almost surely for all $t > 0$, then this process is a local martingale. A process $t \mapsto M(t)$ is called a *local martingale*, if there exists a sequence of stopping times T_n, $n \in \mathbb{N}$, which increases to ∞, such that every process $t \mapsto M(t \wedge T_n)$ is a genuine martingale. A similar notion is available for local sub-martingales, local super-martingales, and processes which are locally of bounded variation. A process $X(t) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ with the property that $\mathbb{E}\left[X(t) \mid \mathcal{F}_s\right] \geqslant X(s)$, $\mathbb{P}$-almost surely for $t > s$, is called a *sub-martingale*, and a process with $\mathbb{E}\left[X(t) \mid \mathcal{F}_s\right] \leqslant X(s)$, $\mathbb{P}$-almost surely for $t > s$, is called a *super-martingale*. A process $X(t) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ is of bounded variation on the interval $[0, T]$ if

$$\sup\left\{\sum_{j=0}^{N-1} |X(X(t_{j+1})) - X(X(t_j))| : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T\right\}$$

is finite. Doob-Meyer's decomposition theorem says that every local sub-martingale $X(t)$ of class DL (locally) can be decomposed as a sum

$X(t) = M(t) + A(t)$, where $M(t)$ is a local martingale and $A(t)$ is an increasing process. By definition the process $t \mapsto X(t \wedge T)$ is of *class DL* if the collection $\{X(\tau) : \tau \leqslant T,\ \tau \text{ stopping time}\}$ is uniformly integrable.

(4) Let $M_1(t)$ and $M_2(t)$ be two martingales in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists a process of bounded variation $\langle M_1(\cdot), M_2(\cdot)\rangle(t)$, the co-variation process of $M_1(t)$ and $M_2(t)$ such that the process $t \mapsto M_1(t)M_2(t) - \langle M_1(\cdot), M_2(\cdot)\rangle(t)$ is an L^1-martingale. A similar result is true for local martingales. If $M_j(t) = \int_0^t H_j(s)dW(s)$, $j = 0, 1$, where $H_1(t)$ and $H_2(t)$ are predictable processes for which $\int_0^t |H_j(s)|^2 ds$ is $\mathbb{P}$-almost surely finite for all $t \geqslant 0$i, and for $j = 1, 2$. Then $\langle M_1(\cdot), M_2(\cdot)\rangle(t) = \int_0^t H_1(s)H_2(s)ds$. Instead of $\langle M_1(\cdot), M_2(\cdot)\rangle(t)$ we often write $\langle M_1, M_2\rangle(t)$; if $M_1(t) = M_2(t) = M(t)$ we also write $\langle M\rangle(t) = \langle M, M\rangle(t)$.

(5) Exponential martingales. Suppose that $M(t)$ and $N(t)$ are martingales. Then the process

$$t \mapsto \mathcal{E}(-N)(t) := \exp\left(-N(t) - \frac{1}{2}\langle N, N\rangle(t)\right)$$

is a martingale, provided Novikov's condition, i.e.

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\langle N, N\rangle(t)\right)\right] < \infty$$

is satisfied for all $t \geqslant 0$. In addition, the process

$$t \mapsto \exp\left(-N(t) - \frac{1}{2}\langle N, N\rangle(t)\right)(M(t) + \langle N, M\rangle(t)) \tag{1.301}$$

is a martingale. If $M(t) = N(t)$, for all $t \geqslant 0$, then the martingale in (1.301) is the same as the second one in

$$t \mapsto \exp\left(-M(t) - \frac{1}{2}\langle M, M\rangle(t)\right) \quad \text{and} \tag{1.302}$$

$$t \mapsto \exp\left(-M(t) - \frac{1}{2}\langle M, M\rangle(t)\right)(M(t) + \langle M, M\rangle(t)). \tag{1.303}$$

The factor $\mathcal{E}(-N)$ can be considered as an risk adjustment factor, $M(t)$ can be interpreted as the volatility (fluctuation, diffusion part), and $\langle N, M\rangle(t)$ is the drift or trend of the process. Define the exponential measure $\mathbb{Q}_N$ by $\mathbb{Q}_N[A] = \mathbb{E}[\mathcal{E}_N(T)1_A]$, $A \in \mathcal{F}_T$. Let $\mathbb{E}_N$ denote the corresponding expectation. The process $M + \langle N, M\rangle$ is then a local martingale with respect to the measure $\mathbb{Q}_N$. This follows from Itô calculus in the following manner. First notice that $d\mathcal{E}_N(t) = -\mathcal{E}_N(t)dN(t)$, and hence, for $0 \leqslant t_1 < t_2 \leqslant T$ we have

$$\begin{aligned}
&\mathcal{E}_N(t_2)\left(M(t_2) + \langle N, M\rangle(t_2)\right) - \mathcal{E}_N(t_1)\left(M(t_1) + \langle N, M\rangle(t_1)\right) \\
&= -\int_{t_1}^{t_2} \mathcal{E}_N(s)\left(M(s) + \langle N, M\rangle(s)\right) dN(s) \\
&\quad + \int_{t_1}^{t_2} \mathcal{E}_N(s)\left(dM(s) + d\langle N, M\rangle(s)\right) - \int_{t_1}^{t_2} \mathcal{E}_N(s)\, d\langle N, M\rangle(s) \\
&= -\int_{t_1}^{t_2} \mathcal{E}_N(s)\left(M(s) + \langle N, M\rangle(s)\right) dN(s) + \int_{t_1}^{t_2} \mathcal{E}_N(s)\, dM(s).
\end{aligned} \tag{1.304}$$

As a consequence of (1.304) we see that the process

$$t \mapsto \mathcal{E}_N(t)\left(M(t) + \langle N, M\rangle(t)\right) \tag{1.305}$$

is a (local) $\mathbb{P}$-martingale. Here we use the fact that stochastic integrals with respect to martingales are (local) martingales. If the expectations $\mathbb{E}\left[e^{\frac{1}{2}\langle N,N\rangle(T)}\right]$, $\mathbb{E}\left[e^{\frac{1}{2}\langle N,N\rangle(T)}\langle N, N\rangle(T)\right]$, and $\mathbb{E}\left[e^{\frac{1}{2}\langle N,N\rangle(T)}\langle M, M\rangle(T)\right]$ are finite, then the stochastic integrals in (1.304) are genuine martingales. This follows from the equalities:

$$\begin{aligned}
&\mathbb{E}_N\left[M(t_2) + \langle N, M\rangle(t_2) \,\middle|\, \mathcal{F}_{t_1}\right] - \left(M(t_1) + \langle N, M\rangle(t_1)\right) \\
&= \mathbb{E}_N\left[M(t_2) + \langle N, M\rangle(t_2) - \left(M(t_1) + \langle N, M\rangle(t_1)\right) \,\middle|\, \mathcal{F}_{t_1}\right] \\
&= \mathbb{E}\left[\mathcal{E}_N(T)\left(M(t_2) + \langle N, M\rangle(t_2)\right) - \mathcal{E}_N(T)\left(M(t_1) + \langle N, M\rangle(t_1)\right) \,\middle|\, \mathcal{F}_{t_1}\right]
\end{aligned}$$

(the process $\mathcal{E}_N(t)$ is a $\mathbb{P}$-martingale)

$$\begin{aligned}
&= \mathbb{E}\left[\mathcal{E}_N(t_2)\left(M(t_2) + \langle N, M\rangle(t_2)\right) - \mathcal{E}_N(t_1)\left(M(t_1) + \langle N, M\rangle(t_1)\right) \,\middle|\, \mathcal{F}_{t_1}\right] \\
&= 0,
\end{aligned}$$

where in the final step we used the martingale property of the process in (1.305).

Corollary 1.7. *Let $N(t)$ be a martingale for which Novikov's condition is satisfied. Put (Radon-Nikodym derivative)*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-N(T) - \frac{1}{2}\langle N, N\rangle(T)\right).$$

Suppose that $W(t) = M(t)$ is a Brownian motion with respect to $\mathbb{P}$. Then $W(t) + \langle N, W\rangle(t)$ is a Brownian motion with respect to $\mathbb{Q}$. In particular, if $N(t) = \int_0^t b(s)dW(s)$, then $W(t) + \int_0^t b(s)ds$, $0 \leqslant t \leqslant T$, is a Brownian motion with respect to $\mathbb{Q}$.

As a consequence we get the following result. Suppose that the process $t \mapsto \exp\left(-\int_0^t b(s)dW(s) - \frac{1}{2}\int_0^t |b(s)|^2\, ds\right)$ is a martingale, and let Φ, $\Psi : C([0,T],\mathbb{R}^n) \to \mathbb{C}$ be bounded continuous functions. Then it follows that

$$\mathbb{E}\left[\Phi\left(t \mapsto W(t) + \int_0^t b(s)ds\right) \exp\left(-\int_0^T b(s)dW(s) - \frac{1}{2}\int_0^T |b(s)|^2\, ds\right)\right]$$
$$= \mathbb{E}\left[\Phi\left(t \mapsto W(t)\right)\right], \quad \text{and} \tag{1.306}$$
$$\mathbb{E}\left[\Psi\left(t \mapsto W(t) + \int_0^t b(s)dW(s)\right)\right]$$
$$= \mathbb{E}\left[\Psi\left(t \mapsto W(t)\right) \exp\left(\int_0^T b(s)dW(s) - \frac{1}{2}\int_0^T |b(s)|^2\, ds\right)\right]. \tag{1.307}$$

Let $\Psi : C([0,T],\mathbb{R}^n) \to \mathbb{C}$ be a bounded continuous function, and put $\widetilde{W}(t) = W(t) + \int_0^t b(s)\, ds$. By applying (1.306) to the function Φ defined by

$$\Phi\left(t \mapsto \widetilde{W}(t)\right) = \Psi\left(t \mapsto \widetilde{W}(t)\right) \exp\left(\int_0^T b(s)d\widetilde{W}(s) - \frac{1}{2}\int_0^T |b(s)|^2\, ds\right)$$

we see that (1.307) is a consequence of (1.306).

(6) Let $H(t)$ be a predictable process, and let $M(t)$ be a martingale. Suppose

$$\mathbb{E}\left[\int_0^t |H(s)|^2\, d\langle M,M\rangle(s)\right] < \infty, \quad t > 0. \tag{1.308}$$

Then the stochastic integral $t \mapsto \int_0^t H(s)dM(s)$ is well defined (as an Itô integral). Moreover, it is a martingale and the equality

$$\mathbb{E}\left[\left|\int_0^t H(s)dM(s)\right|^2\right] = \mathbb{E}\left[\int_0^t |H(s)|^2\, d\langle M,M\rangle(s)\right]$$

is valid. If $H_1(t)$ and $H_2(t)$ are predictable processes which satisfy (1.308), then

$$\mathbb{E}\left[\int_0^t H_1(s)dM(s) \cdot \int_0^t H_2(s)dM(s)\right]$$
$$= \mathbb{E}\left[\int_0^t H_1(s) \cdot H_2(s)d\langle M,M\rangle(s)\right]. \tag{1.309}$$

For $H_j(t) = 1_{A_j} \otimes 1_{(u_j,\infty)}(t)$, $A_j \in \mathcal{F}_{u_j}$, $j = 1, 2$, the equality in (1.309) is readily established, for linear combinations of such indicator functions (i.e. for simple processes) the result also follows easily. A density argument will do the rest.

(7) Suppose that L generates a Feller semigroup with corresponding Markov process

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t) : t \geqslant 0), (\vartheta_t : t \geqslant 0), (E, \mathcal{E})\}.$$

Let f be a function in $D(L)$. Then the process

$$t \mapsto M_f(t) := f(X(t)) - f(X(0)) - \int_0^t Lf(X(s))ds$$

is a martingale.

(8) Let L generate the semigroup e^{tL}, $t \geqslant 0$. Suppose that the marginals of the corresponding Markov process have a density:

$$e^{tL} f(x) = \mathbb{E}_x [f(X(t))] = \int p_0(t, x, y) dm(y)$$

for some reference measure m. Then the process $s \mapsto p_0(t-s, X(s), y)$ is a $\mathbb{P}_x$-martingale on the half open interval $[0, t)$.

(9) Let L be the second order differential operator

$$Lf = b \cdot \nabla f + \frac{1}{2} \sum_{j,k=1}^{d} a_{jk} \frac{\partial^2 f}{\partial x_j \partial x_k}.$$

Then for C^2-functions f_1, f_2 we have

$$\langle M_{f_1}, M_{f_2} \rangle (t) = \int_0^t \Gamma_1 (f_1, f_2) (X(s)) ds$$

where

$$\Gamma_1 (f_1, f_2)(x) = \sum_{j,k=1}^{d} a_{jk}(x) \frac{\partial f_1(x)}{\partial x_j} \frac{\partial f_2(x)}{\partial x_k}.$$

The operator $(f_1, f_2) \mapsto \Gamma_1 (f_1, f_2)$ is called the squared gradient operator, or in French, the *opérateur carré du champ*. The process $\langle M_{f_1}, M_{f_2} \rangle (t)$ is called the (quadratic) covariation process of the local martingales M_{f_1} and M_{f_2}.

(10) Itô's formula. Let

$$X(t) = M(t) + A(t) = (M_1(t), \ldots, M_d(t)) + (A_1(t), \ldots, A_d(t))$$

be a continuous semi-martingale, where $M_j(t)$, $1 \leqslant j \leqslant d$, are local martingales, and where the process A_j, $1 \leqslant j \leqslant d$, are locally of bounded variation. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a C^2-function. Then

$$\begin{aligned} f(X(t)) =& f(X(0)) + \int_0^t \nabla f(X(s)) dM(s) + \int_0^t \nabla f(X(s)) dA(s) \\ &+ \frac{1}{2} \sum_{j,k=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_j \partial x_k} (X(s))\, d \langle M_j, M_k \rangle (s). \end{aligned}$$

We notice that $\langle X_j, X_k \rangle (t) = \langle M_j, M_k \rangle (t)$. Itô's formula says that, under the action of C^2-functions local semi-martingales are preserved. In other words, if $X(t) = M(t) + A(t)$ is a local semi-martingale (i.e. a sum of a local martingale and a process which is locally of bounded variation), and if $f : \mathbb{R}^d \to \mathbb{R}$ is a C^2-function, then the process $t \mapsto f(X(t))$ is a again a local semi-martingale.

1.4 Operator-valued Brownian motion and the Heston volatility model

We insert a definition of an operator-valued Brownian motion. For matrix-valued Brownian motion see e.g. [Biane (2009)] and the references given therein. Perhaps this section can be phrased in terms of quantum probability: see e.g. [Franz and Schott (1999)], [Meyer (1993)], [Biane (1995)], [Hudson and Lindsay (1998)], [Rebolledo *et al.* (2004)]. A main motivation to include it in this book is that the results in Theorem 1.23 put the Heston volatility model in an operator framework, so that, in principle the material could also be used for stochastic volatility matrices.

Definition 1.11. Let $(H, \langle \cdot, \cdot \rangle_H)$ be a Hilbert space, and let $L(H)$ denote the C^*-algebra of all bounded linear operators defined on H with values in H. Let $(\Omega, \mathcal{F}_T, \mathbb{E})$ be a probability space. An $L(H)$-valued process $(B(\tau))_{\tau \in [0,T]}$ is called a Brownian motion if for every every pair $(f, g) \in H \times H$ the process $\tau \mapsto \langle B(\tau)f, g \rangle_H$ is a $\mathbb{P}$-martingale relative to the filtration determined by the variables $B(\tau)$, $\tau \in [0, T]$, and if for every quadruple $(f_1, f_2, g_1, g_2) \in H \times H \times H \times H$ the equality

$$\langle \langle B(\cdot)f_1, f_2 \rangle_H , \langle g_1, B(\cdot)g_2 \rangle_H \rangle (\tau) = \tau \langle f_1, f_2 \rangle_H \langle g_1, g_2 \rangle_H \tag{1.310}$$

holds.

In other words, for every $f \in H$ with $\|f\|_H = 1$ the process $\tau \mapsto \langle B(\tau)f, f \rangle_H$ is a Brownian motion, and if f and g are vectors in H such that $\langle f, g \rangle_H = 0$, and $\|g\|_H = \|f\|_H = 1$, then the Brownian motions $\tau \mapsto \langle B(\tau)f, f \rangle_H$ and $\tau \mapsto \langle B(\tau)g, g \rangle_H$ are $\mathbb{P}$-independent. It follows from Lévy's theorem that such processes $\tau \mapsto \langle B(\tau)f, f \rangle_H$, $\|f\|_H = 1$, are classical Brownian motions. Again we can introduce $L(H)$-valued stochastic integrals using the weak-operator topology. Let $\tau \mapsto \Phi_1(\tau)$ and $\tau \mapsto \Phi_2(\tau)$ be adapted process with property that $\mathbb{E}\left[\int_0^T \|\Phi_j(\rho)f\|_H^2 \, d\rho\right] < \infty$ for $j = 1, 2$. Let $(e_\ell)_{\ell \in \mathbb{N}}$ be an orthonormal basis in H, and let f, $g \in H$. Then the stochastic integrals

$\int_0^\tau \Phi_j(\rho)\, dB(\rho) f$, $j = 1, 2$, have the following basic properties:

$$\begin{aligned}
&\left\langle \int_0^\tau \Phi_1(\rho)\, dB(\rho) f, g \right\rangle_H = \int_0^\tau \int_0^\tau d\left\langle B(\cdot) f, \Phi_1(\rho)^* g\right\rangle_H (\rho) \\
&= \int_0^\tau \int_0^\tau d\left\langle f, (dB(\rho)^*)\, \Phi_1(\rho)^* g\right\rangle_H (\rho) \\
&= \sum_{\ell\in\mathbb{N}} \int_0^\tau d\left\langle B(\rho) f, e_\ell\right\rangle_H \left\langle \Phi_1(\rho) e_\ell, g\right\rangle_H, \quad \text{and} \\
&\mathbb{E}\left[\left\langle \int_0^\tau \Phi_1(\rho)\, dB(\rho) f, \int_0^\tau \Phi_1(\rho)\, dB(\rho) g\right\rangle_H\right] \\
&= \mathbb{E}\left[\int_0^\tau \left\langle \int_0^\tau \Phi_1(\rho) f, \Phi_1(\rho) g\right\rangle_H d\rho\right].
\end{aligned} \tag{1.311}$$

It follows that the stochastic integrals $\tau \mapsto \int_0^\tau \Phi_j(\rho)\, dB(\rho) f$, $j = 1, 2$, belong to that subspace of $L^2_{\text{weak}}(\Omega, \mathcal{F}_T, \mathbb{P})$ which are H-valued L^2-martingales relative to the filtration determined by the stochastic variables $\tau \mapsto \langle B(\tau) h_1, h_2\rangle_H$, $\tau \in [0, T]$, h_1, $h_2 \in H$. For the definition of this space the reader is referred to Definition 1.8. Itô's formula is available in some restricted sense. Then the following equality holds for all f, $g \in H$:

$$\begin{aligned}
&\left\langle \int_0^\tau \Phi_1(\rho)\, dB(\rho) f, \int_0^\tau \Phi_1(\rho)\, dB(\rho) g\right\rangle_H \\
&= \int_0^\tau \left\langle \Phi_1(\rho)\, dB(\rho) f, \int_0^\rho \Phi_2\left(\rho'\right)\, dB\left(\rho'\right)\right\rangle_H \\
&\quad + \int_0^\tau \left\langle \int_0^\rho \Phi_1\left(\rho'\right)\, dB\left(\rho'\right) f, \Phi_2\left(\rho\right)\, dB\left(\rho\right)\right\rangle_H \\
&\quad + \int_0^\tau \left\langle \Phi_1(\rho) f, \Phi_2(\rho)\right\rangle_H d\rho.
\end{aligned} \tag{1.312}$$

A proof of the equality in (1.312) can be based on the following arguments. By density and bilinearity it suffices to prove (1.312) for Φ_j of the form $\Phi_j(\tau) = T_j \mathbf{1}_{(t_j,\infty)}(\tau)$ where, for every f, $g \in H$ the random variable $\langle T_j f, g\rangle_H$ is $\mathcal{F}_{t_j}$-measurable, $j = 1, 2$. Here $\mathcal{F}_t$ is the σ-field generated by the variables $\tau \mapsto \langle B(\tau) f, g\rangle_H$, $0 \leqslant \tau \leqslant t$, f, $g \in H$. Let the $\tau \mapsto A_1(\tau) + M_1(\tau)$ and $\tau \mapsto A_2(\tau) + M_2(\tau)$ be $L(H)$-valued local semi-martingales, i.e. the processes $\tau \mapsto \langle A_j(\tau) f, g\rangle_H$. $j = 1, 2$, are locally of bounded variation $\mathbb{P}$-almost surely, and the processes $\tau \mapsto \langle M_j(\tau) f, g\rangle_H$, $j = 1, 2$, are local martingales for all elements f, $g \in H$. Then for the co-variation of the processes of the processes $A_1(\tau) + M_1(\tau)$ and $A_2(\tau) + M_2(\tau)$ we write

$$\left\langle \left(A_1(\cdot) + M_1(\cdot)\right) f, \left(A_2(\cdot) + M_2(\cdot)\right) g\right\rangle (\tau)$$

$$= \sum_{\ell=1}^{\infty} \langle \langle (A_1(\cdot) + M_1(\cdot)) f, e_\ell \rangle_H, \langle e_\ell, (A_2(\cdot) + M_2(\cdot)) g \rangle_H \rangle (\tau)$$
$$= \sum_{\ell=1}^{\infty} \langle \langle M_1(\cdot) f, e_\ell \rangle_H, \langle e_\ell, M_2(\cdot) g \rangle_H \rangle (\tau). \tag{1.313}$$

The second equality in (1.313) follows because the classical covariation of real or complex semi-martingales only depends depends on the martingale parts of these process. Since, by the definition of operator-valued Brownian motion,

$$\langle \langle B(\cdot) f, e_m \rangle_H, \langle e_n, B(\cdot) g \rangle_H \rangle (\tau) = \tau \langle f, e_m \rangle_H \langle e_n, g \rangle_H, \quad \tau \in [0,T], \ f, \ g \in H,$$

we obtain the following equalities:

$$\left\langle \left(A_1(\cdot) + \int_0^{(\cdot)} \Phi_1(\rho)\, dB(\rho) \right) f, \left(A_2(\cdot) + \int_0^{(\cdot)} \Phi_2(\rho)\, dB(\rho) \right) g \right\rangle (\tau)$$
$$= \int_0^{\tau} \langle \Phi_1(\rho) f, \Phi_2(\rho) g \rangle_H \, d\rho, \tag{1.314}$$

whenever the operator-valued processes $\Phi_1(\tau)$ and $\Phi_2(\tau)$ are predictable and satisfy:

$$\mathbb{E}\left[\int_0^T \langle \Phi_j(\rho) f, \Phi_j(\rho) f \rangle_H \, d\rho \right] < \infty, \ f \in H, \quad j = 1, 2.$$

We also observe that the process $\tau \mapsto B(\tau)^*$ is a Brownian motion. In the following (proof of) Theorem 1.23 the equalities in (8.217), (8.81), (1.312), (1.313) and (1.314) will be freely used. Throughout the present section it is assumed that the stochastic processes are adapted to the filtration determined by the stochastic variables $\{\langle B(\tau) f, g \rangle_H : \tau \in [0,T], \ f, \ g \in H\}$. Let $\mathcal{F}_\tau$ be the σ-field generated by $\{\langle B(\rho) f, g \rangle_H : 0 \leqslant \rho \leqslant \tau \in [0,T], \ f, \ g \in H\}$. Moreover, it is assumed that, unless stated otherwise, all operator-valued processes are adapted and continuous for the weak operator topology, and that therefore they are automatically predictable. This in the sense that the mappings $(\tau, \omega) \mapsto \Phi_j(\tau, \omega)$, $j = 1, 2$, are measurable with respect to the σ-field generated by $\left\{ \mathbf{1}_{(a,b]} \otimes \mathbf{1}_A : 0 \leqslant a < b \leqslant T, \ A \in \mathcal{F}_a \right\}$; compare with item (3) in §1.3. Let $\tau \mapsto M_j(\tau)$, $\tau \in [0,T]$, $j = 1, 2$, be two semi-martingales with values in the Hilbert space H. Notice that one has to distinguish between $\langle M_1(\cdot), M_2(\cdot) \rangle (\tau)$, which the covariation process between $\tau \mapsto M_1(\tau)$ and $\tau \mapsto M_2(\tau)$ and $\langle M_1(\tau), M_2(\tau) \rangle_H$, which denotes an inner-product.

Theorem 1.23. *The following assertions are equivalent.*

(1) *There exists an adapted $L(H)$-valued process $\tau \mapsto A(\tau)$, $\tau \in [0,T]$, which satisfies the integral equation:*

$$A(\tau) - A(0) \tag{1.315}$$
$$= \frac{1}{2}\int_0^\tau \left(\left(\frac{4\kappa\eta}{\lambda^2}-1\right)\left(A(\rho)^* + B(\rho)^*\right)^{-1} - \kappa\left(A(\rho)+B(\rho)\right)\right) d\rho.$$

(2) *There exists an adapted $L(H)$-valued process $\tau \mapsto A(\tau)$, $\tau \in [0,T]$, which satisfies the integral equation*

$$2\int_0^\tau \left(A(\rho)^* + B(\rho)^*\right) dA(\rho) \tag{1.316}$$
$$+ \kappa\int_0^\tau \left(A(\rho)+B(\rho)\right)^*\left(A(\rho)+B(\rho)\right) d\rho = \left(\frac{4\kappa\eta}{\lambda^2}-1\right)\tau I.$$

(3) *There exists a pair of adapted $L(H)$-valued processes $\tau \mapsto (V(\tau), U_\lambda(\tau))$, $\tau \in [0,T]$, with $V(\tau) = U_\lambda(\tau)^* U_\lambda(\tau)$, which has the following properties.*
   (a) *The following stochastic differential equation holds:*
$$dV(\tau) = \kappa\left(\eta I - V(\tau)\right) d\tau + \frac{\lambda}{2}\left(U_\lambda(\tau)^* \, dB(\tau) + dB(\tau)^* U_\lambda(\tau)\right). \tag{1.317}$$
   (b) *For all $f,\ g \in H$ the following equality holds: $\langle U_\lambda(\cdot)f, U_\lambda(\cdot)g\rangle(\tau) = \frac{\lambda^2}{4}\tau\langle f,g\rangle$, $\tau \in [0,T]$.*
   (c) *For all $\tau \in [0,T]$ the operators $\int_0^\tau U_\lambda(\rho)^* \, d\left(U_\lambda - \frac{\lambda}{2}B\right)(\rho)$ are self-adjoint $\mathbb{P}$-almost surely.*

(4) *There exists an adapted $L(H)$-valued semi-martingale $\tau \mapsto U(\tau)$, $\tau \in [0,T]$, with the following properties.*
   (a) *The following stochastic differential equation holds:*
$$d\left(U(\tau)^* U(\tau)\right) = \kappa\left(\frac{4\eta}{\lambda^2}I - U(\tau)^* U(\tau)\right) d\tau$$
$$+ \left(dB(\tau)\right)^* U(\tau) + U(\tau)^* \, dB(\tau). \tag{1.318}$$
   (b) *For all $f,\ g \in H$ the following equality holds: $\langle U(\cdot)f, U(\cdot)g\rangle(\tau) = \tau\langle f,g\rangle$, $\tau \in [0,T]$.*
   (c) *For every $\tau \in [0,T]$ the operator $\int_0^\tau U(\rho)^* d\left(U-B\right)(\rho)$ is self-adjoint $\mathbb{P}$-almost surely.*

(5) *There exists an adapted $L(H)$-valued semi-martingale $\tau \mapsto U(\tau)$, $\tau \in [0,T]$, which satisfies the following stochastic differential equation:*
$$dU(\tau) = \frac{1}{2}\left(\left(\frac{4\kappa\eta}{\lambda^2}-1\right)\left(U(\tau)^*\right)^{-1} - \kappa U(\tau)\right) d\tau + dB(\tau). \tag{1.319}$$

(6) *There exists an $L(H)$-valued predictable process $\tau \mapsto A(\tau)$ with the following properties:*

(a) *The operator-valued martingales $\tau \mapsto M_j(\tau)$, $\tau \in [0,T]$, $j = 1, 2$, defined by*

$$M_1(\tau) = M_1(0) + \int_0^\tau (A(\rho) + B(\rho))\, dB(\rho), \quad \text{and}$$
$$M_2(\tau) = M_2(0) + \int_0^\tau (A(\rho)^* + B(\rho)^*)\, dB(\rho), \tag{1.320}$$

satisfy the following integral equality:

$$\begin{aligned}&\frac{d}{d\tau} \langle M_1(\cdot)f, M_1(\cdot)g\rangle (\tau) + \kappa \langle M_1(\cdot)f, M_1(\cdot)g\rangle (\tau) \qquad (1.321)\\ &= \langle A(0)f, A(0)g\rangle_H + \frac{4\kappa\eta\tau}{\lambda^2} \langle f, g\rangle_H \\ &\quad + \langle (M_2(\tau) - M_2(0))\, f, g\rangle_H + \langle f, (M_2(\tau) - M_2(0))\, g\rangle_H ,\end{aligned}$$

for all f, $g \in H$.

(b) *The predictable process $\tau \mapsto \langle A(\tau)f, g\rangle_H$ is locally of bounded variation for all f, $g \in H$,*

(c) *The operators $\int_0^\tau (A(\rho)^* + B(\rho)^*)\, dA(\rho)$ are almost surely self-adjoint.*

(7) *There exists an $L(H)$-valued predictable process $\tau \mapsto A(\tau)$, which satisfies the conditions in (b) and (c) of item (6), possesses the following additional property. The operator-valued martingales $\tau \mapsto M_j(\tau)$, $\tau \in [0,T]$, $j = 1, 2$ defined as in (1.320) satisfy the following integral equality:*

$$\begin{aligned}&\langle M_1(\cdot)f, M_1(\cdot)g\rangle (\tau)\\ &= \frac{4\eta}{\kappa\lambda^2} \left(e^{-\kappa\tau} - 1 + \tau\kappa\right) \langle f, g\rangle_H + \frac{1 - e^{-\kappa\tau}}{\kappa} \langle A(0)f, A(0)g\rangle_H \\ &\quad + \int_0^\tau e^{-\kappa(\tau-\rho)} (\langle (M_2(\rho) - M_2(0))f, g\rangle_H + \langle f, (M_2(\rho) - M_2(0))g\rangle_H) d\rho.\end{aligned} \tag{1.322}$$

Proof. The equivalence of (1) and (2) follows by differentiating the expressions in (1.315) and (1.316) respectively.

(2) $\Longrightarrow$ (3). Let the process $A(\tau)$ be as in (2). Put $U_\lambda(\tau) = \frac{1}{2}\lambda\, (A(\tau) + B(\tau))$. From (1.316) we get:

$$\int_0^\tau (A(\rho)^* + B(\rho)^*)\, dA(\rho) + \int_0^\tau (dA(\rho))^* \, (A(\rho) + B(\rho))$$

$$+\kappa\int_0^\tau (A(\rho)+B(\rho))^*(A(\rho)+B(\rho))\,d\rho = \left(\frac{4\kappa\eta}{\lambda^2}-1\right)\tau I. \tag{1.323}$$

From Itô calculus and (1.323) we obtain:

$$\begin{aligned} dV(\tau) &= \frac{\lambda^2}{4} d\left((A(\rho)^*+B(\rho)^*)(A(\rho)+B(\rho))\right) \\ &= \frac{\lambda^2}{4}\left((A(\tau)^*+B(\tau)^*)\,dA(\tau)+dA(\tau)^*(A(\tau)+B(\tau))\right) \\ &\quad + \frac{\lambda^2}{4}\left(dB(\tau)^*(A(\tau)+B(\tau))+(A(\tau)+B(\tau)^*)\right)+\frac{\lambda^2}{4}I\,d\tau \\ &= \kappa(\eta I - V(\tau))\,d\tau + \frac{\lambda}{2}\left(dB(\tau)^*U_\lambda(\tau)+U_\lambda(\tau)^*\,dB(\tau)\right). \end{aligned} \tag{1.324}$$

The equality in (1.317) follows from (1.324). This shows (a) of Assertion (3). The statement in (b) follows from the equality in (1.316). Finally, (c) is a consequence of the representation $U_\lambda(\tau)=\frac{1}{2}\lambda(A(\tau)+B(\tau))$, where $A(\tau)$ is locally of differentiable $\mathbb{P}$-almost surely. It follows that

$$\langle U_\lambda(\cdot)f, U_\lambda(\cdot)g\rangle(\tau) = \frac{\lambda^2}{4}\langle B(\cdot)f, B(\cdot)g\rangle(\tau) = \frac{\lambda^2}{4}\langle f,g\rangle_H\,\tau. \tag{1.325}$$

Then (1.325) implies (c).

The equivalence of (3) and (4) follows by the relationship $\lambda U(\tau) = 2U_\lambda(\tau)$ where $U_\lambda(\cdot)$ and $U(\cdot)$ are as in (3) and (4) respectively.

(4) $\Longrightarrow$ (5). Let the process $\tau\mapsto U(\tau)$ be as in (4). In particular it satisfies the equation in (1.318). This together with (b) and (c) in (4) and Itô calculus shows:

$$\begin{aligned} &2U(\tau)^*\,d(U-B)(\tau) \\ &= U(\tau)^*\,d(U(\tau)-B(\tau))+d(U(\tau)^*-B(\tau)^*)\,U(\tau) \\ &= d(U(\tau)^*U(\tau)) - I\,d\tau - U(\tau)^*\,dB(\tau) - d(B(\tau)^*)\,U(\tau) \\ &= \kappa\left(\frac{4\eta}{\lambda^2}-U(\tau)^*U(\tau)\right)d\tau - I\,d\tau. \end{aligned} \tag{1.326}$$

The equation in (1.319) follows from (1.326). Consequently, (5) follows from (4). In fact the Assertions (2) and (4) are also easy consequences of (5), as we shall see next.

(5) $\Longrightarrow$ (2). Let the $L(H)$-valued process $\tau\mapsto U(\tau)$ have the properties in (5). Put

$$A(\tau) = U(\tau)-B(\tau) = U(0)+\int_0^\tau\left(\left(\frac{4\kappa\eta}{\lambda^2}-1\right)(U(\rho)^*)^{-1}-\kappa U(\rho)\right)d\rho. \tag{1.327}$$

Then from (1.319) we get

$$
\begin{aligned}
&\left(A(\tau)^* + B(\tau)^*\right) dA(\tau) + dA(\tau)^* \left(A(\tau) + B(\tau)\right) \\
&\quad + \kappa \left(A(\tau)^* + B(\tau)^*\right)\left(A(\tau) + B(\tau)\right) d\tau \\
&= U(\tau)^* \, d\left(U - B\right)(\tau) + \left(U(\tau)^* \, d\left(U - B\right)(\tau)\right)^* + \kappa U(\tau)^* U(\tau)\, d\tau
\end{aligned}
$$

(apply (1.319))

$$
= \left(\frac{4\kappa\eta}{\lambda^2} - 1\right) I \, d\tau. \tag{1.328}
$$

Since, by (c) of (5) operators of the form

$$
\int_0^\tau \left(A(\rho)^* + B(\rho)^*\right) dA(\rho) = \int_0^\tau U(\rho)^* \, d\left(U - B\right)(\rho)
$$

are self-adjoint the equality in (1.316) follows. This completes the proof of the implication (5) $\Longrightarrow$ (2).

(2) $\Longrightarrow$ (6). Let the process $\tau \mapsto A(\tau)$ be as in (1.316), and define the $L(H)$-valued martingales $M_j(\tau)$, $j = 1, 2$, as in (1.320). Then by Itô calculus and by the proof of the implication (2) $\Longrightarrow$ (3) we obtain

$$
\begin{aligned}
&\left(A(\tau)^* + B(\tau)^*\right)\left(A(\tau) + B(\tau)\right) \\
&= A(0)^* A(0) + \int_0^\tau \left(A(\rho)^* + B(\rho)^*\right) dA(\rho) + \int_0^\tau dA(\rho)^* \left(A(\rho) + B(\rho)\right) \\
&\quad + \int_0^\tau \left(A(\rho)^* + B(\rho)^*\right) dB(\rho) + \int_0^\tau dB(\rho)^* \left(A(\rho) + B(\rho)\right) + \tau I
\end{aligned}
$$

(apply (1.323) in the proof of the implication (2) $\Longrightarrow$ (3))

$$
\begin{aligned}
&= A(0)^* A(0) + \left(\frac{4\kappa\eta}{\lambda^2} - 1\right) \tau I - \kappa \int_0^\tau \left(A(\rho)^* + B(\rho)^*\right)\left(A(\rho) + B(\rho)\right) d\rho I \\
&\quad + \tau I + M_2(\tau) + M_2(\tau)^* - M_2(0) - M_2(0)^* \\
&= A(0)^* A(0) + \frac{4\kappa\eta}{\lambda^2} \tau I - \kappa \int_0^\tau \left(A(\rho)^* + B(\rho)^*\right)\left(A(\rho) + B(\rho)\right) d\rho \\
&\quad + M_2(\tau) + M_2(\tau)^* - M_2(0) - M_2(0)^*.
\end{aligned} \tag{1.329}
$$

Let $f, g \in H$. Then by (8.81) and (1.329) we see

$$
\begin{aligned}
&\frac{d}{d\tau} \left\langle M_1(\cdot) f, M_1(\cdot) g \right\rangle (\tau) + \kappa \left\langle M_1(\cdot) f, M_1(\cdot) g \right\rangle (\tau) \\
&= \frac{d}{d\tau} \int_0^\tau \left\langle \left(A(\rho) + B(\rho)\right) f, \left(A(\rho) + B(\rho)\right) g \right\rangle_H d\rho
\end{aligned}
$$

$$
\begin{aligned}
&+\kappa\int_0^\tau \left\langle \left(A(\rho)+B(\rho)\right)f, \left(A(\rho)+B(\rho)\right)g\right\rangle_H \, d\rho \\
&= \frac{d}{d\tau}\left\langle M_1(\cdot)f, M_1(\cdot)g\right\rangle(\tau) + \kappa\left\langle M_1(\cdot)f, M_1(\cdot)g\right\rangle(\tau) \\
&= \left\langle \left(A(\tau)^* + B(\tau)^*\right)\left(A(\tau)+B(\tau)\right)f, g\right\rangle_H \\
&\quad + \kappa\int_0^\tau \left\langle \left(A(\rho)^* + B(\rho)^*\right)\left(A(\rho)+B(\rho)\right)f, g\right\rangle_H \, d\rho \\
&= \left\langle \left(A(0)^*A(0) + \frac{4\kappa\eta}{\lambda^2}\tau + M_2(\tau) + M_2(\tau)^* - M_2(0) - M_2(0)^*\right) f, g\right\rangle_H \\
&= \left\langle A(0)f, A(0)g\right\rangle_H + \frac{4\kappa\eta}{\lambda^2}\tau\left\langle f, g\right\rangle_H \\
&\quad + \left\langle \left(M_2(\tau) - M_2(0)\right)f, g\right\rangle_H + \left\langle f, \left(M_2(\tau) - M_2(0)\right)g\right\rangle_H . \qquad (1.330)
\end{aligned}
$$

The equality in (1.195) completes the proof of the implication (2) $\Longrightarrow$ (6).

(6) $\Longleftrightarrow$ (7). The fact that the equalities in (1.321) and (1.322) are equivalent is a simple exercise in ordinary differential equations. It follows that the assertions (6) and (7) are equivalent.

(6) $\Longrightarrow$ (2). The arguments in the proof of the implication (2) $\Longrightarrow$ (6) can be reversed. More precisely, if (6) is true, then the equality in (1.321) holds. It follows that the equalities in (1.330) hold. An application of Itô calculus then yields the equality:

$$
\begin{aligned}
&\left(A(\tau)^* + B(\tau)^*\right) dA(\tau) + dA(\tau)^*\left(A(\tau)+B(\tau)\right) + I\, d\tau \\
&\quad + \kappa\left(A(\tau)^* + B(\tau)^*\right)\left(A(\tau)+B(\tau)\right) d\tau \\
&= \frac{4\kappa\eta}{\lambda^2} I\, d\tau + \left(A(\tau)^* + B(\tau)^*\right) dB(\tau) + dB(\tau)^*\left(A(\tau)+B(\tau)\right). \qquad (1.331)
\end{aligned}
$$

(See (1.328) in the proof of the implication (5) $\Longrightarrow$ (2).) From property (c) in (6) we get $dA(\tau)^*\left(A(\tau)+B(\tau)\right) = \left(A(\tau)^* + B(\tau)^*\right) dA(\tau)$, and hence (1.331) implies

$$
\begin{aligned}
&2\left(A(\tau)^* + B(\tau)^*\right) dA(\tau) + \kappa\left(A(\tau)^* + B(\tau)^*\right)\left(A(\tau)+B(\tau)\right) d\tau \\
&= \left(\frac{4\kappa\eta}{\lambda^2} - 1\right) I\, d\tau,
\end{aligned}
$$

and so (1.316) in Assertion (2) follows.

This completes the proof of Theorem 1.23. □

As a corollary to Theorem 1.23 we have the following results.

Corollary 1.8. *Theorem 1.23 is also true if throughout we assume that the operators* $A(\tau) + B(\tau)$, *respectively,* $U(\tau)$, $\tau \in [0, T]$, *are predictable* *and self-adjoint. In this case the martingales $M_1(\tau)$ and $M_2(\tau)$ in Assertions (6) and (7) may be taken equal. If moreover, it is assumed that throughout Theorem 1.23 the operators $A(\tau) + B(\tau)$, respectively, $U(\tau)$, $\tau \in [0,T]$, are predictable and positive, then not only the martingales $M_1(\tau)$ and $M_2(\tau)$ in Assertions (6) and (7) may be taken equal, but the operator process $\tau \mapsto V(\tau)$ satisfies an operator version of the volatility Heston model. This means that the equation in (1.317) can be written in the form:*

$$dV(\tau) = \kappa\left(\eta I - V(\tau)\right)\, d\tau + \frac{\lambda}{2}\left(\sqrt{V(\tau)}\, dB(\tau) + dB(\tau)^*\sqrt{V(\tau)}\right). \quad (1.332)$$

In the following theorem we specialize the result in Theorem 1.23 to the case where $H = \mathbb{R}$. If in (1.335) the process $U_\lambda(\tau)$ is non-negative, then this equation corresponds to the classical Heston model for the volatility. For more details on the Heston volatility model see e.g. [Feng *et al.* (2010); In 't Hout and Foulon (2010)] and many others. It was Heston [Heston (1993)] who first used this stochastic volatility model. For a related stochastic interest rate model the reader is referred to [Cox *et al.* (1985)].

Theorem 1.24. *The following assertions are equivalent.*

(1) There exists an adapted $\mathbb{R}$-valued process $\tau \mapsto A(\tau)$, $\tau \in [0,T]$, which satisfies the integral equation:

$$A(\tau) = A(0) + \frac{1}{2}\int_0^\tau \left(\left(\frac{4\kappa\eta}{\lambda^2} - 1\right)\frac{1}{A(\rho) + B(\rho)} - \kappa\left(A(\rho) + B(\rho)\right)\right) d\rho. \quad (1.333)$$

(2) There exists an adapted $\mathbb{R}$-valued process $\tau \mapsto A(\tau)$, $\tau \in [0,T]$, which satisfies the integral equation

$$2\int_0^\tau \left(A(\rho) + B(\rho)\right) dA(\rho) + \kappa\int_0^\tau \left(A(\rho) + B(\rho)\right)^2 d\rho = \left(\frac{4\kappa\eta}{\lambda^2} - 1\right)\tau. \quad (1.334)$$

(3) There exists a pair of adapted $\mathbb{R}$-valued processes $\tau \mapsto (V(\tau), U_\lambda(\tau))$, $\tau \in [0,T]$, with $V(\tau) = U_\lambda(\tau)^2$, such that the following stochastic differential equation holds:

$$dV(\tau) = \kappa\left(\eta I - V(\tau)\right)\, d\tau + \lambda U_\lambda(\tau)\, dB(\tau). \quad (1.335)$$

(4) There exists an adapted $\mathbb{R}$-valued semi-martingale $\tau \mapsto U(\tau)$, $\tau \in [0,T]$, such that the following stochastic differential equation holds:

$$d\left(U(\tau)^2\right) = \kappa\left(\frac{4\eta}{\lambda^2} I - U(\tau)^2\right) d\tau + 2U(\tau)\, dB(\tau). \quad (1.336)$$

(5) *There exists an adapted* $\mathbb{R}$*-valued semi-martingale* $\tau \mapsto U(\tau)$, $\tau \in [0,T]$, *which satisfies the following stochastic differential equation:*

$$dU(\tau) = \frac{1}{2}\left(\left(\frac{4\kappa\eta}{\lambda^2} - 1\right) U(\tau)^{-1} - \kappa U(\tau)\right) d\tau + dB(\tau). \tag{1.337}$$

(6) *There exists an* $\mathbb{R}$*-valued martingale* $\tau \mapsto M(\tau)$ *for which the following integral equality is satisfied:*

$$\begin{aligned}&\frac{d}{d\tau}\left\langle M(\cdot), M(\cdot)\right\rangle(\tau) + \kappa\left\langle M(\cdot), M(\cdot)\right\rangle(\tau)\\ &= M(0)^2 + \frac{4\kappa\eta\tau}{\lambda^2} + 2\left(M(\tau) - M(0)\right).\end{aligned}$$

(7) *There exists a* $\mathbb{R}$*-valued martingale* $\tau \mapsto M(\tau)$, *which satisfies the following integral equality:*

$$\begin{aligned}&\left\langle M(\cdot), M(\cdot)\right\rangle(\tau) - \frac{4\eta}{\kappa\lambda^2}\left(e^{-\kappa\tau} - 1 + \tau\kappa\right)\\ &= \frac{1 - e^{-\kappa\tau}}{\kappa} M(0)^2 + 2\int_0^\tau e^{-\kappa(\tau-\rho)}\left(M(\rho) - M(0)\right).\end{aligned}$$

The proof follows the same pattern as the proof of Theorem 1.23 with the following extra arguments. From the stochastic differential equation in (1.335) it follows that $\left\langle U_\lambda(\cdot), U_\lambda(\cdot)\right\rangle(\tau) = \frac{1}{4}\lambda^2\tau$. From the stochastic differential equation in (1.336) it follows that $\left\langle U(\cdot), U(\cdot)\right\rangle(\tau) = \tau$. By the martingale representation theorem a martingale $M(\tau)$ which is adapted to a Brownian $\tau \mapsto B(\tau)$ can be written in the form $M(\tau) = M(0) + \int_0^\tau U(\rho)\, dB(\rho)$. If such a martingale also satisfies the integral equation in Assertion (6), then by Itô calculus it follows that $\left\langle U(\cdot), U(\cdot)\right\rangle(\tau) = \tau$. As a consequence it follows that $U(\tau) = A(\tau) + B(\tau)$ where the process $A(\tau)$ is predictable and locally of bounded variation. The proof of the implication (6) $\Longrightarrow$ (2) then proceeds along the same lines as the proof of the corresponding implication in Theorem 1.23.

1.5 Stopping times and time-homogeneous Markov processes

Next we explain the strong Markov property. Let E be a Polish space (i.e. a complete metrtizable separable Hausdorff space) with Borel field $\mathcal{E}$, and let

$$\left\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t) : t \geqslant 0), (\vartheta_t : t \geqslant 0), (E, \mathcal{E})\right\}$$

be a family of probability spaces with state variables $X(t) : \Omega \to E$ and time translation operators $\vartheta_s : \Omega \to \Omega$ such that $X(t) \circ \vartheta_s = X(t+s)$, $\mathbb{P}_x$-almost surely for all $s, t \geqslant 0$, $x \in E$. Moreover, $\vartheta_{s+t} = \vartheta_s \circ \vartheta_t$, $s, t \geqslant 0$. Assume that the process $t \mapsto X(t)$ is $\mathbb{P}_x$-almost surely right-continuous for all $x \in E$. If E is locally compact and not compact, then $E^{\triangle}$ is the one-point-compactification of E; otherwise $\triangle$ is an isolated point of the topological space $E^{\triangle}$. Sometimes E is augmented with an extra absorption state $\triangle$: $E^{\triangle} = E \bigcup \triangle$. The σ-field $\mathcal{F}_t$ is generated by the state variables $X(s)$, $0 \leqslant s \leqslant t$, and $\mathcal{F}$ is generated by the process $t \mapsto X(t)$. Since the sample paths $t \mapsto X(t)$, $t \geqslant 0$ are right continuous $\mathbb{P}_x$-almost surely our Markov process is a strong Markov process. Let $S : \Omega \to [0, \infty]$ be a *stopping* meaning that for every $t \geqslant 0$ the event $\{S \leqslant t\}$ belongs to $\mathcal{F}_t$. This is the same as saying that the process $t \mapsto \mathbf{1}_{[S \leqslant t]}$ is adapted. Let $\mathcal{F}_S$ be the natural σ-field associated with the stopping time S, i.e.

$$\mathcal{F}_S = \bigcap_{t \geqslant 0} \left\{ A \in \mathcal{F} : A \bigcap \{S \leqslant t\} \in \mathcal{F}_t \right\}.$$

Define $\vartheta_S(\omega)$ by $\vartheta_S(\omega) = \vartheta_{S(\omega)}(\omega)$. Consider $\mathcal{F}_S$ as the information from the past, $\sigma\big(X(S)\big)$ as information from the present, and

$$\sigma \{X(t) \circ \vartheta_S : t \geqslant 0\} = \sigma \{X(t+S) : t \geqslant 0\}$$

as the information from the future. The time-homogeneous Markov property can be expressed as follows:

$$\mathbb{E}_x \left[f(X(s+t)) \mid \mathcal{F}_s \right] = \mathbb{E}_x \left[f(X(s+t)) \mid \sigma(X(s)) \right] = \mathbb{E}_{X(s)} \left[f(X(t)) \right], \tag{1.338}$$

$\mathbb{P}_x$-almost surely for all $f \in C_b(E)$ and for all s and $t \geqslant 0$. The strong Markov property can be expressed as follows:

$$\mathbb{E}_x \left[Y \circ \vartheta_S | \mathcal{F}_S \right] = \mathbb{E}_{X(S)} \left[Y \right], \ \mathbb{P}_x\text{-almost surely} \tag{1.339}$$

on the event $\{S < \infty\}$, for all bounded random variables Y, for all stopping times S, and for all $x \in E$. One can prove that under the "cadlag" property events like $\{X(S) \in B,\ S < \infty\}$, B Borel, are $\mathcal{F}_S$-measurable. The passage from (1.339) to (1.338) is easy: put $Y = f(X(t))$ and $S(\omega) = s$, $\omega \in \Omega$. The other way around is much more intricate and uses the cadlag property of the process $\{X(t) : t \geqslant 0\}$. In this procedure the stopping time S is approximated by a decreasing sequence of discrete stopping times $(S_n = 2^{-n} \lceil 2^n S \rceil : n \in \mathbb{N})$. The equality

$$\mathbb{E}_x \left[Y \circ \vartheta_{S_n} | \mathcal{F}_{S_n} \right] = \mathbb{E}_{X(S_n)} \left[Y \right], \ \mathbb{P}_x\text{-almost surely},$$

is a consequence of (2) for a fixed time. Let n tend to infinity in (1.5) to obtain (1.339). The "strong Markov property" can be extended to the "strong time dependent Markov property":

$$\mathbb{E}_x\left[Y\left(S+T\circ\vartheta_S,\vartheta_S\right)\,\middle|\,\mathcal{F}_S\right](\omega)=\mathbb{E}_{X\left(S(\omega)\right)}\left[\omega'\mapsto Y\left(S(\omega)+T\left(\omega'\right),\omega'\right)\right], \tag{1.340}$$

$\mathbb{P}_x$-almost surely on the event $\{S<\infty\}$. Here $Y:[0,\infty)\times\Omega\to\mathbb{C}$ is a bounded random variable. The cartesian product $[0,\infty)\times\Omega$ is supplied with the product field $\mathcal{B}_{[0,\infty)}\otimes\mathcal{F}$; $\mathcal{B}_{[0,\infty)}$ is the Borel field of $[0,\infty)$ and $\mathcal{F}$ is (some extension of) $\sigma\left(X(u):u\geqslant 0\right)$. Important stopping times are "hitting times", or times related to hitting times:

$$\begin{aligned}T_U&=\inf\left\{s>0:X(s)\in E^{\triangle}\backslash U\right\},\quad\text{and}\\ S&=\inf\left\{s>0:\int_0^s 1_{E\backslash U}(X(u))du>0\right\},\end{aligned}$$

where U is some open (or Borel) subset of $E^{\triangle}$. This kind of stopping times have the extra advantage of being *terminal* stopping times, i.e. $t+S\circ\vartheta_t=S$ $\mathbb{P}_x$-almost surely on the event $\{S>t\}$. A similar statement holds for the *hitting time* T_U. The time S is called the *penetration time* of $E\backslash U$. Let $p:E\to[0,\infty)$ be a Borel measurable function. Stopping times of the form

$$S_\xi=\inf\left\{s>0:\int_0^s p\big(X(u)\big)du>\xi\right\}$$

serve as a *stochastic time change*, because they enjoy the equality: $S_\xi+S_\eta\circ\vartheta_{S_\xi}=S_{\xi+\eta}$, $\mathbb{P}_x$-almost surely on the event $\{S_\xi<\infty\}$. As a consequence operators of the form $\mathcal{S}(\xi)f(x):=\mathbb{E}_x\left[f\left(X\left(S_\xi\right)\right)\right]$, f a bounded Borel function, possess the semigroup property. Also notice that $S_0=0$, provided that the function p is strictly positive.

PART 2

Strong Markov Processes

This page is intentionally left blank

Chapter 2

Strong Markov processes on Polish spaces

In this chapter we describe time-dependent strong Markov processes with a Polish space as state space. As indicated in Chapter 1 stochastic differential equations in a Banach space often give rise to Markov processes with a separable Banach space as state space. In our theory we are mainly interested in the Markovian behavior of our process. In addition we consider the corresponding martingale problem and the problem of unique Markov extensions. As highlights we mention the theorems 2.9 through 2.13. In order to establish these general results a study of the strict topology is required as well as a precise knowledge of measures on Polish spaces. These topics also are included in this chapter.

2.1 Strict topology

Throughout this book E stands for a complete metrizable separable topological space, i.e. E is a Polish space. A recent book which among other things treats Polish spaces is [Kanovei (2008)]. The Borel field of E is denoted by $\mathcal{E}$. We write $C_b(E)$ for the space of all complex valued bounded continuous functions on E. The space $C_b(E)$ is equipped with the supremum norm: $\|f\|_\infty = \sup_{x\in E} |f(x)|$, $f \in C_b(E)$. The space $C_b(E)$ will be endowed with a second topology which will be used to describe the continuity properties. This second topology, which is called the strict topology, is denoted as $\mathcal{T}_\beta$-topology. The strict topology is generated by the semi-norms of the form p_u, where u varies over $H(E)$, and where $p_u(f) = \sup_{x\in E} |u(x)f(x)| = \|uf\|_\infty$, $f \in C_b(E)$. Here a function u belongs to $H(E)$ if u is bounded and if for every real number $\alpha > 0$ the set $\{|u| \geqslant \alpha\} = \{x \in E : |u(x)| \geqslant \alpha\}$ is contained in a compact subset of E. It is noticed that Buck [Buck (1958)] was the first author who introduced the notion of strict topology (in the locally compact setting). He used the notation β instead of $\mathcal{T}_\beta$.

Remark 2.1. Let $H^+(E)$ be the collection of those functions $u \in H(E)$ with the following properties: $u \geqslant 0$ and for every $\alpha > 0$ the set $\{u \geqslant \alpha\}$ is a compact subset of E. Then every function $u \in H^+(E)$ is bounded, and the strict topology is also generated by semi-norms of the form $\{p_u : u \in H^+(E)\}$. Every $u \in H^+(E)$ attains its supremum at some point $x \in E$. Moreover a sequence $(f_n)_{n\in\mathbb{N}}$ converges for the strict topology to a function $f \in C_b(E)$ if and only if it is uniformly bounded and if for every compact subset K of E the equality $\lim_{m\to\infty} \sup_{n\geqslant m} \sup_{x\in K} |f_n(x) - f_m(x)| = 0$ holds. Since $\mathcal{T}_\beta$-convergent sequences are $\mathcal{T}_\beta$-bounded, from Proposition 2.1 below it follows that a $\mathcal{T}_\beta$-convergent sequence is uniformly bounded. The same conclusion is true for $\mathcal{T}_\beta$-Cauchy sequences. Moreover, a $\mathcal{T}_\beta$-Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ converges to a bounded function f. Such a sequence converges uniformly on compact subsets of the space E. Since the space E is Polish, it follows that the limit function f is continuous. Consequently, the space $(C_b(E), \mathcal{T}_\beta)$ is sequentially complete. Observe that continuity properties of functions $f \in C_b(E)$ can be formulated in terms of convergent sequences in E which are contained in compact subsets of E. The topology of uniform convergence on $C_b(E)$ is denoted by $\mathcal{T}_u$. In the sixties Conway (see [Conway (1966, 1967)]) proved that the strict topology $\mathcal{T}_\beta$ is the Mackey topology for the duality of $C_b(E)$ and the space of bounded complex Borel measures on E. This means that $\mathcal{T}_\beta$ is the finest locally convex topology on $C_b(E)$ for which the dual is given by the space $\mathcal{M}(E)$, the space of bounded complex Borel measures on $\mathcal{M}(E)$; for more details see e.g. [Sentilles (1970)].

2.1.1 *Theorem of Daniell-Stone*

In Proposition 2.2 below we need the following theorem. It says that an abstract integral is a concrete integral relative to a σ-additive measure. Theorem 2.1 will be applied with $S = E$, $H = C_b^+$, the collection of non-negative functions in $C_b(E)$, and for $I : C_b^+ \to [0,\infty)$ we take the restriction to C_b^+ of a non-negative linear functional defined on $C_b(E)$ which is continuous with respect to the strict topology.

Theorem 2.1 (Theorem of Daniell-Stone). *Let S be any set, and let H be a non-empty collection of functions on S with the following properties:*

(1) If f and g belong to H, then the functions $f+g$, $f \vee g$ and $f \wedge g$ belong to H as well;
(2) If $f \in H$ and α is a non-negative real number, then αf, $f \wedge \alpha$, and $(f-\alpha)^+ = (f-\alpha) \vee 0$ belong to H;
(3) If f, $g \in H$ are such that $f \leqslant g \leqslant \mathbf{1}$, then $g - f$ belongs to H.

Let $I : H \to [0, \infty]$ be an abstract integral in the sense that I is a mapping which possesses the following properties:

(4) If f and g belong to H, then $I(f+g) = I(f) + I(g)$;
(5) If $f \in H$ and $\alpha \geqslant 0$, then $I(\alpha f) = \alpha I(f)$;
(6) If $(f_n)_{n\in\mathbb{N}}$ is a sequence in H which increases pointwise to $f \in H$, then $I(f_n)$ increases to $I(f)$.

Then there exists a non-negative σ-additive measure μ on the σ-field generated by H, which is denoted by $\sigma(H)$, such that $I(f) = \int f d\mu$, for $f \in H$. If there exists a countable family of functions $(f_n)_{n\in\mathbb{N}} \subset H$ such that $I(f_n) < \infty$ for all $n \in \mathbb{N}$, and such that $S = \bigcup_{n=1}^{\infty} \{f_n > 0\}$, then the measure μ is unique.

Proof. Define the collection H^* of functions on S as follows. A function $f : S \to [0, \infty]$ belongs to H^* provided there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset H$ which increases pointwise to f. Then the subset H^* has the properties (1) and (2) with H^* instead of H. Define the mapping $I^* : H^* \to [0, \infty]$ by

$$I^*(f) = \lim_{n\to\infty} I(f_n), \quad f \in H^*,$$

where $(f_n)_{n\in\mathbb{N}} \subset H$ is a sequence which pointwise increases to f. The definition does not depend on the choice of the increasing sequence $(f_n)_{n\in\mathbb{N}} \subset H$. In fact let $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ be sequences in H which both increase to $f \in H^*$. Then by (6) we have

$$\begin{aligned}\lim_{n\to\infty} I(f_n) &= \sup_{n\in\mathbb{N}} I(f_n) = \sup_{n\in\mathbb{N}} \sup_{m\in\mathbb{N}} I(f_n \wedge g_m) = \sup_{m\in\mathbb{N}} \sup_{n\in\mathbb{N}} I(f_n \wedge g_m) \\ &= \sup_{m\in\mathbb{N}} I(g_m) = \lim_{m\to\infty} I(g_m). \end{aligned} \tag{2.1}$$

From (2.1) it follows that I^* is well-defined. The functional $I^* : H^* \to [0, \infty]$ has the properties (4), (5), and (6) (somewhat modified) with H^* instead of H and I replaced by I^*. In fact the correct version of (6) for H^* reads as follows:

(6*) Let $(f_n)_{n\in\mathbb{N}}$ be a sequence H^* which increases pointwise to a function f. Then $f \in H^*$, and $I^*(f_n)$ increases to $I^*(f)$.

We also have the following assertion:

(3*) Let f and $g \in H^*$ be such that $f \leqslant g$. Then $I^*(f) \leqslant I^*(g)$.

We first prove (3*) if f and g belong to H and $f \leqslant g$. From (6), (3) and (4) we get

$$\begin{aligned} I(g) &= \sup_{m\in\mathbb{N}} I\left(g \wedge m\right) = \sup_{m\in\mathbb{N}} \left(I\left(g \wedge m - f \wedge m\right) + I\left(f \wedge m\right)\right) \\ &\geqslant \sup_{m\in\mathbb{N}} I\left(f \wedge m\right) = I\left(f\right). \end{aligned} \tag{2.2}$$

Here we used the fact that by (3) the functions $g \wedge m - f \wedge m$, $m \in \mathbb{N}$, belong to H. Next let f and g be functions in H such that $f \leqslant g$. Then there exist increasing sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in H such that f_n converges pointwise to $f \in H^*$ and g_n to $g \in H^*$. Then

$$I^*\left(f\right) = \sup_{n\in\mathbb{N}} I\left(f_n\right) \leqslant \sup_{n\in\mathbb{N}} I\left(f_n \vee g_n\right) = I^*\left(g\right). \tag{2.3}$$

Next we prove (6*). Let $(f_n)_{n\in\mathbb{N}}$ be a pointwise increasing sequence in H^*, and put $f = \sup_{n\in\mathbb{N}} f_n$. Choose for every $n \in \mathbb{N}$ an increasing sequence $(f_{n,m})_{m\in\mathbb{N}} \subset H$ such that $\sup_{m\in\mathbb{N}} f_{n,m} = f_n$. Define the functions g_m, $m \in \mathbb{N}$, by

$$g_m = f_{1,m} \vee f_{2,m} \vee \cdots \vee f_{m,m}.$$

Then $g_{m+1} \geqslant g_m$ and $g_m \in H$ for all $m \in \mathbb{N}$. In addition, we have

$$\sup_{m\in\mathbb{N}} g_m = \sup_{m\in\mathbb{N}} \max_{1\leqslant n\leqslant m} f_{n,m} = \sup_{n\in\mathbb{N}} \sup_{m\geqslant n} f_{n,m} = \sup_{n\in\mathbb{N}} f_n = f. \tag{2.4}$$

Hence $f \in H^*$. For $1 \leqslant n \leqslant m$ the inequalities $f_{n,m} \leqslant f_n \leqslant f_m$ hold pointwise, and hence $g_m \leqslant f_m$. From (3*) we infer

$$I^*\left(f\right) = \sup_{m\in\mathbb{N}} I\left(g_m\right) = \sup_{m\in\mathbb{N}} I^*\left(g_m\right) \leqslant \sup_{m\in\mathbb{N}} I^*\left(f_m\right) \leqslant I^*\left(f\right), \tag{2.5}$$

and thus $\sup_{m\in\mathbb{N}} I^*\left(f_m\right) = I^*\left(f\right)$.

Next we will get closer to measure theory. Therefore we define the collection $\mathcal{G}$ of subsets of S by $\mathcal{G} = \{G \subset S : \mathbf{1}_G \in H^*\}$, and the mapping $\mu : \mathcal{G} \to [0, \infty]$ by $\mu\left(G\right) = I^*\left(\mathbf{1}_G\right)$, $G \in \mathcal{G}$. The mapping μ possesses the following properties:

(1′) If the subsets G_1 and G_2 belong to $\mathcal{G}$, then the same is true for the subsets $G_1 \bigcap G_2$ and $G_1 \bigcup G_2$;

(2′) $\varnothing \in \mathcal{G}$;

(3′) If the subsets G_1 and G_2 belong to $\mathcal{G}$ and if $G_1 \subset G_2$, then $\mu\left(G_1\right) \leqslant \mu\left(G_2\right)$;

(4′) If the subsets G_1 and G_2 belong to $\mathcal{G}$, then the following strong additivity holds: $\mu(G_1 \bigcap G_2) + \mu(G_1 \bigcup G_2) = \mu(G_1) + \mu(G_2)$;

(5′) $\mu(\varnothing) = 0$;

(6′) If $(G_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{G}$ such that $G_{n+1} \supset G_n$, $n \in \mathbb{N}$, then $\bigcup_{n\in\mathbb{N}} G_n$ belongs to $\mathcal{G}$ and $\mu(\bigcup_{n\in\mathbb{N}} G_n) = \sup_{n\in\mathbb{N}} \mu(G_n)$.

These properties are more or less direct consequences of the corresponding properties of I^*: (1*)–(6*).

Using the mapping μ we will define an exterior or outer measure μ^* on the collection of all subsets of S. Let A be any subset of S. Then we put $\mu^*(A) = \infty$ if for no $G \in \mathcal{G}$ we have $A \subset G$, and we write $\mu^*(A) = \inf\{\mu(G) : G \in \mathcal{G}, G \supset A\}$, if $A \subset G_0$ for some $G_0 \in \mathcal{G}$. Then μ^* has the following properties:

(i) $\mu^*(\varnothing) = 0$;

(ii) $\mu^*(A) \geqslant 0$, for all subsets A of S;

(iii) $\mu^*(A) \leqslant \mu^*(B)$, whenever A and B are subsets of S for which $A \subset B$;

(iv) $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n)$ for any sequence $(A_n)_{n\in\mathbb{N}}$ of subsets of S.

The assertions (i), (ii) and (iii) follow directly from the definition of μ^*. In order to prove (iv) we choose a sequence $(A_n)_{n\in\mathbb{N}}$, $A_n \subset S$, such that $\mu^*(A_n) < \infty$ for all $n \in \mathbb{N}$. Fix $\varepsilon > 0$, and choose for every $n \in \mathbb{N}$ an subset G_n of S which belongs to $\mathcal{G}$ and which has the following properties: $A_n \subset G_n$ and $\mu(G_n) \leqslant \mu^*(A_n) + \varepsilon 2^{-n}$. By the equality $\bigcup_{n\in\mathbb{N}} G_n = \bigcup_{m\in\mathbb{N}} \bigcup_{n=1}^{m} G_n$ we see that $\bigcup_{n\in\mathbb{N}} G_n$ belongs to $\mathcal{G}$. From the properties of an exterior measure we infer the following sequence of inequalities:

$$\begin{aligned}
\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) &\leqslant \mu^*\left(\bigcup_{n\in\mathbb{N}} G_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} G_n\right) = \sup_{m\in\mathbb{N}} \mu\left(\bigcup_{n=1}^{m} G_n\right) \\
&= \sup_{m\in\mathbb{N}} I^*\left(\mathbf{1}_{\bigcup_{n=1}^{m} G_n}\right) \leqslant \sup_{m\in\mathbb{N}} I^*\left(\sum_{n=1}^{m} \mathbf{1}_{G_n}\right) = \sup_{m\in\mathbb{N}} \sum_{n=1}^{m} I^*\left(\mathbf{1}_{G_n}\right) \\
&= \sup_{m\in\mathbb{N}} \sum_{n=1}^{m} \mu(G_n) \leqslant \sum_{n=1}^{\infty} \left(\mu^*(A_n) + \varepsilon 2^{-n}\right) = \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon. \qquad (2.6)
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary we see that $\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n)$. Hence Assertion (iv) follows.

Next we consider the σ-field $\mathcal{D}$ which is associated to the exterior measure μ^*, and which is defined by

$$\mathcal{D} = \left\{A \subset S : \mu^*(D) \geqslant \mu^*\left(A \bigcap D\right) + \mu^*\left(A^c \bigcap D\right) \quad \text{for all } D \subset S\right\}$$

$$= \Big\{ A \subset S : \mu(D) \geqslant \mu^*\left(A \bigcap D\right) + \mu^*\left(A^c \bigcap D\right) \\ \text{for all } D \in \mathcal{G} \text{ with } \mu(D) < \infty \Big\}. \tag{2.7}$$

Here we wrote $A^c = S \backslash A$ for the complement of A in S. The reader is invited to check the equality in (2.7). According to Caratheodory's theorem the exterior measure μ^* restricted to the σ-field $\mathcal{D}$ is a σ-additive measure. We will prove that $\mathcal{D}$ contains $\mathcal{G}$. Therefore pick $G \in \mathcal{G}$, and consider for $D \in \mathcal{G}$ for which $\mu(D) < \infty$ the equality

$$\mu^*\left(G \bigcap D\right) + \mu^*\left(G^c \bigcap D\right) = \mu\left(G \bigcap D\right) + \inf\left\{\mu(U) : U \in \mathcal{G}, U \supset G^c \bigcap D\right\}. \tag{2.8}$$

Choose $h \in H^*$ in such that $h \geqslant \mathbf{1}_{G^c \bigcap D}$. For $0 < \alpha < 1$ we have

$$\mathbf{1}_{G^c \bigcap D} \leqslant \mathbf{1}_{\{h > \alpha\}} \leqslant \frac{1}{\alpha} h.$$

Since $\mathbf{1}_{\{h>\alpha\}} = \sup_{m \in \mathbb{N}} \mathbf{1} \wedge (m(h-\alpha)^+)$ we see that the set $\{h > \alpha\}$ is a member of $\mathcal{G}$. It follows that $I^*(h) \geqslant \alpha\mu(\{h > \alpha\}) \geqslant \alpha\mu^*(G^c \bigcap D)$, and hence

$$\begin{aligned} \mu^*\left(G^c \bigcap D\right) &\leqslant \inf\left\{I^*(h) : h \geqslant \mathbf{1}_{G^c \bigcap D}, h \in H^*\right\} \\ &\leqslant \inf\left\{I^*(\mathbf{1}_U) : U \supset G^c \bigcap D, U \in \mathcal{G}\right\} = \mu^*\left(G^c \bigcap D\right). \end{aligned} \tag{2.9}$$

From (2.9) the equality

$$\mu^*\left(G^c \bigcap D\right) = \inf\left\{I^*(h) : h \geqslant \mathbf{1}_{G^c \bigcap D}, h \in H^*\right\}$$

follows. Next choose the increasing sequences $(f_n)_{n \in \mathbb{N}}$ and $(g_n)_{n \in \mathbb{N}}$ in such a way that the sequence f_n increases to $\mathbf{1}_D$ and g_n increases to $\mathbf{1}_G$. Define the functions h_n, $n \in \mathbb{N}$, by

$$h_n = \mathbf{1}_D - f_n \wedge g_n = \sup_{m \geqslant n} \left\{(f_m - f_n) + (f_n - f_n \wedge g_n)\right\}.$$

Since the functions $f_m - f_n$, $m \geqslant n$, and $f_n - f_n \wedge g_n$ belong to H we see that h_n belongs to H^*. Hence we get:

$$\infty > \mu(D) = I^*(\mathbf{1}_D) = I^*(h_n) + I^*(f_n \wedge g_n) = I^*(h_n) + I(f_n \wedge g_n). \tag{2.10}$$

In addition we have $h_n \geqslant \mathbf{1}_{G^c \bigcap D}$. Consequently,

$$\begin{aligned} &\mu^*\left(G \bigcap D\right) + \mu^*\left(G^c \bigcap D\right) \\ &\leqslant \mu\left(G \bigcap D\right) + \inf_{n \in \mathbb{N}} I^*(h_n) \end{aligned}$$

$$= \mu\left(G\bigcap D\right) + \mu(D) - \sup_{n\in\mathbb{N}} I(f_n \wedge g_n)$$
$$= \mu\left(G\bigcap D\right) + \mu(D) - \mu\left(G\bigcap D\right) = \mu(D). \tag{2.11}$$

The equality in (2.11) proves that the σ-field $\mathcal{D}$ contains the collection $\mathcal{G}$, and hence that the mapping μ, which originally was defined on $\mathcal{G}$, is in fact the restriction to $\mathcal{G}$ of a genuine measure defined on the σ-field generated by H. This restriction is again called μ.

We will show the equality $I(f) = \int f d\mu$ for all $f \in H$. For $f \in H$ we have

$$\int f d\mu = \int_0^\infty \mu\{f > \xi\} d\xi = \int_0^\infty I^*\left(\mathbf{1}_{\{f>\xi\}}\right) d\xi$$
$$= \sup_{n\in\mathbb{N}} \frac{1}{2^n}\sum_{j=1}^{n2^n} I^*\left(\mathbf{1}_{\{f>j2^{-n}\}}\right) = \sup_{n\in\mathbb{N}} I^*\left(\frac{1}{2^n}\sum_{j=1}^{n2^n}\mathbf{1}_{\{f>j2^{-n}\}}\right)$$
$$= I^*\left(x \mapsto \int_0^\infty \mathbf{1}_{\{f>\xi\}}(x) d\xi\right) = I^*(f) = I(f). \tag{2.12}$$

Finally we will prove the uniqueness of the measure μ. Let μ_1 and μ_2 be two measures on $\sigma(H)$ with the property that $I(f) = \int f d\mu_1 = \int f d\mu_2$ for all $f \in H$. Under the extra condition in Theorem 2.1 that there exist countable many functions $(f_n)_{n\in\mathbb{N}}$ such that $I(f_n) < \infty$ for all $n \in \mathbb{N}$ and such that $S = \bigcup_{n=1}^\infty \{f_n > 0\}$ we shall show that $\mu_1(B) = \mu_2(B)$ for all $B \in \sigma(H)$. Therefore we fix a function $f \in H$ for which $I(f) < \infty$. Then the collection $\left\{B \in \sigma(H) : \int_B f d\mu_1 = \int_B f d\mu_2\right\}$ is a Dynkin system containing all sets of the form $\{g > \beta\}$ with $g \in H$ and $\beta > 0$. Fix $\xi > 0$, $\beta > 0$ and $g \in H$. Then the functions $g_{m,n} := \min\left(m(g-\beta)^+ \wedge \mathbf{1}, n(f-\xi)^+ \wedge \mathbf{1}\right)$, m, $n \in \mathbb{N}$, belong to H. Then we have

$$\mu_1\left[\{g > \beta\}\bigcap\{f > \xi\}\right] = \lim_{m\to\infty}\lim_{n\to\infty}\int g_{m,n} d\mu_1 = \lim_{m\to\infty}\lim_{n\to\infty} I(g_{m,n})$$
$$= \lim_{m\to\infty}\lim_{n\to\infty}\int g_{m,n} d\mu_2 = \mu_2\left[\{g > \beta\}\bigcap\{f > \xi\}\right]. \tag{2.13}$$

Integration of the extreme terms in (2.13) with respect to the Lebesgue measure $d\xi$ shows the equality $\int_{\{g>\beta\}} f d\mu_1 = \int_{\{g>\beta\}} f d\mu_2$. It follows that the collection $\left\{B \in \sigma(H) : \int_B f d\mu_1 = \int_B f d\mu_2\right\}$ contains all sets of the form $\{g > \beta\}$ where $g \in H$ and $\beta > 0$. Such collection of sets is closed under finite intersection. Since the Dynkin system generated by a collection of subsets which is closed under finite intersections coincides with the σ-field generated by such a set, we infer the equality

$$\left\{B \in \sigma(H) : \int_B f d\mu_1 = \int_B f d\mu_2\right\} = \sigma(H).$$

Such an argument may be called a "Dynkin argument"; the Monotone Class Theorem generalizes such an argument. See Remark 2.16 on the π-λ theorem as well. The same argument applies with $(nf) \wedge \mathbf{1}$ replacing f. By letting n tend to ∞ this shows the equality

$$\sigma(H) = \left\{B \in \sigma(H) : \mu_1 \left[B \bigcap \{f > 0\}\right] = \mu_2 \left[B \bigcap \{f > 0\}\right]\right\}. \tag{2.14}$$

Since the set H is closed under taking finite maxima, $I(f \vee g) \leqslant I(f) + I(g) < \infty$ whenever $I(f)$ and $I(g)$ are finite, and $S = \bigcup_{n=1}^{\infty} \{f_n > 0\}$ with $I(f_n) < \infty$, $n \in \mathbb{N}$, we see that

$$\begin{aligned} \mu_1(B) &= \lim_{n\to\infty} \mu_1 \left[B \bigcap \left\{\max_{1\leqslant j\leqslant n} f_j > 0\right\}\right] \\ &= \lim_{n\to\infty} \mu_2 \left[B \bigcap \left\{\max_{1\leqslant j\leqslant n} f_j > 0\right\}\right] = \mu_2(B) \end{aligned} \tag{2.15}$$

for $B \in \sigma(H)$. □

This finishes the proof of Theorem 2.1.

2.1.2 *Measures on Polish spaces*

Our first proposition says that the identity mapping $f \mapsto f$ sends $\mathcal{T}_\beta$-bounded subsets of $C_b(E)$ to $\|\cdot\|_\infty$-bounded subsets.

Proposition 2.1. *Every $\mathcal{T}_\beta$-bounded subset of $C_b(E)$ is $\|\cdot\|_\infty$-bounded. On the other hand the identity is not a continuous operator from $(C_b(E), \mathcal{T}_\beta)$ to $(C_b(E), \|\cdot\|_\infty)$, provided that E itself is not compact.*

Proof. Let $B \subset C_b(E)$ be $\mathcal{T}_\beta$-bounded. If B were not uniformly bounded, then there exist sequences $(f_n)_{n\in\mathbb{N}} \subset B$ and $(x_n)_{n\in\mathbb{N}} \subset E$ such that $|f_n(x_n)| \geqslant n^2$, $n \in \mathbb{N}$. Put $u(x) = \sum_{n=1}^{\infty} \frac{1}{n} \mathbf{1}_{x_n}$. Then the function u belongs to $H(E)$, but $\sup_{f\in B} p_u(f) \geqslant \sup_{n\in\mathbb{N}} p_u(f_n) \geqslant \sup_{n\in\mathbb{N}} u(x_n)|f(x_n)| \geqslant \sup_{n\in\mathbb{N}} n = \infty$. The latter shows that the set B is not $\mathcal{T}_\beta$-bounded. By contra-position it follows that $\mathcal{T}_\beta$-bounded subsets are uniformly bounded.

Next suppose that E is not compact. Let u be any function in $H(E)$. Then $\lim_{n\to\infty} u(x_n) = 0$. If the imbedding $(C_b(E), \mathcal{T}_\beta) \to (C_b(E), \mathcal{T}_u)$ were continuous, then there would exist a function $u \in H^+(E)$ such that $\|f\|_\infty \leqslant \|uf\|_\infty$ for all $f \in C_b(E)$. Let K be a compact subset of E such that $0 \leqslant u(x) \leqslant \frac{1}{2}$ for $x \notin K$. Since $1 \leqslant \|u\|_\infty = u(x_0)$ for some $x_0 \in E$, and since by assumption E is not compact we see that $K \neq E$. Choose an open neighborhood O of K, $O \neq E$, and a function $f \in C_b(E)$ such that

$\mathbf{1} - \mathbf{1}_O \leqslant f \leqslant \mathbf{1} - \mathbf{1}_K$. In particular, it follows that $f = 1$ outside of O, and $f = 0$ on K. Then $1 = \|f\|_\infty \leqslant \|uf\|_\infty \leqslant \sup_{x \notin K} |u(x)f(x)| \leqslant \frac{1}{2} \|f\|_\infty \leqslant \frac{1}{2}$. Clearly, this is a contradiction.

This concludes the proof of Proposition 2.1. □

The following proposition shows that the dual of the space $(C_b(E), \mathcal{T}_\beta)$ coincides with the space of all complex Borel measures on $\mathcal{E}$.

Proposition 2.2.

(1) Let μ be a complex Borel measure on E. Then there exists a function $u \in H(E)$ such that $\left|\int f d\mu\right| \leqslant p_u(f)$ for all $f \in C_b(E)$.
(2) Let $\Lambda : C_b(E) \to \mathbb{C}$ be a linear functional on $C_b(E)$ which is continuous with respect to the strict topology. Then there exists a unique complex measure μ on $\mathcal{E}$ such that $\Lambda(f) = \int f d\mu$, $f \in C_b(E)$.

Proof. (1). Since on a Polish space every bounded Borel measure is inner-regular, there exists an increasing sequence of compact subsets $(K_n)_{n\in\mathbb{N}}$ in E with $K_0 = \varnothing$ such that $|\mu|(E\backslash K_n) \leqslant 2^{-2n-2} |\mu|(E)$, $n \in \mathbb{N}$. Fix $f \in C_b(E)$. Then we have

$$\begin{aligned}
\left|\int f d\mu\right| &\leqslant \sum_{j=0}^{\infty} \left|\int_{K_{j+1}\backslash K_j} f d\mu\right| \leqslant \sum_{j=0}^{\infty} \int_{K_{j+1}\backslash K_j} |f|\, d|\mu| \\
&\leqslant \sum_{j=0}^{\infty} \left\|\mathbf{1}_{K_{j+1}\backslash K_j} f\right\|_\infty |\mu| (K_{j+1}\backslash K_j) \\
&\leqslant \sum_{j=0}^{\infty} \left\|\mathbf{1}_{K_{j+1}\backslash K_j} f\right\|_\infty |\mu| (E\backslash K_j) \\
&\leqslant \sum_{j=0}^{\infty} 2^{-2j-2} \left\|\mathbf{1}_{K_{j+1}\backslash K_j} f\right\|_\infty |\mu| (E) \\
&\leqslant \sum_{j=0}^{\infty} 2^{-2j-2} 2^{j+1} \|uf\|_\infty \leqslant \|uf\|_\infty
\end{aligned} \tag{2.16}$$

where $u(x) = \sum_{j=1}^{\infty} 2^{-j} \mathbf{1}_{K_j}(x) |\mu| (E)$.

(2). We decompose the functional Λ into a combination of four positive functionals: $\Lambda = (\Re\Lambda)^+ - (\Re\Lambda)^- + i(\Im\Lambda)^+ - i(\Im\Lambda)^-$ where the linear functionals $(\Re\Lambda)^+$ and $(\Re\Lambda)^-$ are determined by their action on positive functions $f \in C_b(E)$:

$$(\Re\Lambda)^+(f) = \sup\{\Re(\Lambda(g)) : 0 \leqslant g \leqslant f, g \in C_b(E)\}, \quad \text{and}$$
$$(\Re\Lambda)^-(f) = \sup\{\Re(-\Lambda(g)) : 0 \leqslant g \leqslant f, g \in C_b(E)\}.$$

Similar expressions can be employed for the action of $(\Im\Lambda)^+$ and $(\Im\Lambda)^-$ on functions $f \in C_b^+$. Since the complex linear functional $\Lambda : C_b(E) \to \mathbb{C}$ is $\mathcal{T}_\beta$-continuous there exists a function $u \in H^+(E)$ such that $|\Lambda(f)| \leqslant \|uf\|_\infty$ for all $f \in C_b(E)$. Then it easily follows that $\left|(\Re\Lambda)^+(f)\right| \leqslant \|uf\|_\infty$ for all real-valued functions in $C_b(E)$, and $\left|(\Re\Lambda)^+(f)\right| \leqslant \sqrt{2}\,\|uf\|_\infty$ for all $f \in C_b(E)$, which in general take complex values. Similar inequalities hold for $(\Re\Lambda)^-(f)$, $(\Im\Lambda)^+(f)$, and $(\Im\Lambda)^-(f)$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions in $C_b^+(E)$ which pointwise increases to a function $f \in C_b^+(E)$. Then $\lim_{n\to\infty}\Lambda(f_n) = \Lambda(f)$. This can be seen as follows. Put $g_n = f - f_n$, and fix $\varepsilon > 0$. Then the sequence $(g_n)_{n\in\mathbb{N}}$ decreases pointwise to 0. Moreover it is dominated by f. Choose a strictly positive real number α in such a way that $\alpha\|f\|_\infty \leqslant \varepsilon$. Then it follows that

$$\begin{aligned} |\Lambda(g_n)| \leqslant \|ug_n\|_\infty &= \max\left(\left\|u\mathbf{1}_{\{u\geqslant\alpha\}}g_n\right\|_\infty, \left\|u\mathbf{1}_{\{u<\alpha\}}g_n\right\|_\infty\right) \\ &\leqslant \max\left(\|u\|_\infty\left\|\mathbf{1}_{\{u\geqslant\alpha\}}g_n\right\|_\infty, \alpha\|f\|_\infty\right) \leqslant \varepsilon \end{aligned} \tag{2.17}$$

where N chosen so large that $\|u\|_\infty\left\|\mathbf{1}_{\{u\geqslant\alpha\}}g_n\right\|_\infty \leqslant \varepsilon$ for $n \geqslant N$. By Dini's lemma such a choice of N is possible. An application of Theorem 2.1 then yields the existence of measures μ_j, $1 \leqslant j \leqslant 4$, defined on the Baire field of E such that $(\Re\Lambda)^+(f) = \int f d\mu_1$, $(\Re\Lambda)^-(f) = \int f d\mu_2$, $(\Im\Lambda)^+(f) = \int f d\mu_3$, and $(\Im\Lambda)^-(f) = \int f d\mu_4$ for $f \in C_b(E)$. It follows that $\Lambda(f) = \int f d\mu_1 - \int f d\mu_2 + i\int f d\mu_3 - i\int f d\mu_4 = \int f d\mu$ for $f \in C_b(E)$. Here $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$ and each measure μ_j, $1 \leqslant j \leqslant 4$, is finite and positive. Since the space E is Polish it follows that Baire field coincides with the Borel field, and hence the measure μ is a complex Borel measure.

This concludes the proof of Proposition 2.2. □

The next corollary gives a sequential continuity characterization of linear functionals which belong to the space $(C_b(E), \mathcal{T}_\beta)^*$, the topological dual of the space $C_b(E)$ endowed with the strict topology. We say that a sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ converges for the strict topology to $f \in C_b(E)$ if $\lim_{n\to\infty}\|u(f - f_n)\|_\infty = 0$ for all functions $u \in H^+(E)$. It follows that a sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ converges to a function $f \in C_b(E)$ with respect to the strict topology if and only if this sequence is uniformly bounded and $\lim_{n\to\infty}\|\mathbf{1}_K(f - f_n)\|_\infty = 0$ for all compact subsets K of E.

Corollary 2.1. *Let $\Lambda : C_b(E) \to \mathbb{C}$ be a linear functional. Then the following assertions are equivalent:*

(1) The functional Λ belongs to $(C_b(E), \mathcal{T}_\beta)^*$;

(2) $\lim_{n\to\infty} \Lambda(f_n) = 0$ *whenever* $(f_n)_{n\in\mathbb{N}}$ *is a sequence in* $C_b^+(E)$ *which converges to the zero-function for the strict topology;*

(3) There exists a finite constant $C \geqslant 0$ *such that* $|\Lambda(f)| \leqslant C\|f\|_\infty$ *for all* $f \in C_b(E)$*, and* $\lim_{n\to\infty} \Lambda(g_n) = 0$ *whenever* $(g_n)_{n\in\mathbb{N}}$ *is a sequence in* $C_b^+(E)$ *which is dominated by a sequence* $(f_n)_{n\in\mathbb{N}}$ *in* $C_b^+(E)$ *which decreases pointwise to 0;*

(4) There exists a finite constant $C \geqslant 0$ *such that* $|\Lambda(f)| \leqslant C\|f\|_\infty$ *for all* $f \in C_b(E)$*, and* $\lim_{n\to\infty} \Lambda(f_n) = 0$ *whenever* $(f_n)_{n\in\mathbb{N}}$ *is a sequence in* $C_b^+(E)$ *which decreases pointwise to 0;*

(5) There exists a complex Borel measure μ *on* E *such that* $\Lambda(f) = \int f d\mu$ *for all* $f \in C_b(E)$*.*

In (3) we say that a sequence $(g_n)_{n\in\mathbb{N}}$ in $C_b^+(E)$ is dominated by a sequence $(f_n)_{n\in\mathbb{N}}$ if $g_n \leqslant f_n$ for all $n \in \mathbb{N}$. A functional $\Lambda : C_b(E) \to \mathbb{C}$ with the property that for every sequence $(f_n)_{n\in\mathbb{N}}$ which decreases pointwise to zero the inequality $\lim_{n\to\infty} \Lambda(f_n) = 0$ is called a σ-smooth functional in [Varadarajan (1961, 1999)].

Proof. (1) $\Longrightarrow$ (2). First suppose that Λ belongs to $(C_b(E), \mathcal{T}_\beta)^*$. Then there exists a function $u \in H^+(E)$ such that $|\Lambda(f)| \leqslant \|uf\|_\infty$ for all $f \in C_b(E)$. Hence, if the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E)$ converges to zero for the strict topology, then $\lim_{n\to\infty} \|uf_n\|_\infty = 0$, and so $\lim_{n\to\infty} \Lambda(f_n) = 0$. This proves the implication (1) $\Longrightarrow$ (2).

(2) $\Longrightarrow$ (3). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $C_b(E)$ which converges to 0 for the uniform topology. From (2) it follows that the sequences $\left((\Re f_n)^+\right)_{n\in\mathbb{N}}$, $\left((\Re f_n)^-\right)_{n\in\mathbb{N}}$, $\left((\Im f_n)^+\right)_{n\in\mathbb{N}}$, and $\left((\Im f_n)^-\right)_{n\in\mathbb{N}}$ converge to 0 for the strict topology $\mathcal{T}_\beta$, and hence $\lim_{n\to\infty} \Lambda(f_n) = 0$. Consequently, the functional $\Lambda : C_b(E) \to \mathbb{C}$ is continuous if $C_b(E)$ is equipped with the uniform topology, and hence there exists a finite constant $C \geqslant 0$ such that $|\Lambda(f)| \leqslant C\|f\|_\infty$ for all $f \in C_b(E)$. If $(f_n)_{n\in\mathbb{N}}$ is a sequence in $C_b^+(E)$ which decreases to 0, then by Dini's lemma it converges uniformly on compact subsets of E to 0. Moreover, it is uniformly bounded, and hence it converges to 0 for the strict topology. If the sequence $(g_n)_{n\in\mathbb{N}} \subset C_b^+(E)$ is such that $g_n \leqslant f_n$. Then the sequence $(g_n)_{n\in\mathbb{N}}$ converges to 0 for the strict topology. Assertion (2) implies that $\lim_{n\to\infty} \Lambda(g_n) = 0$.

(3) $\Longrightarrow$ (4). This implication is trivial.

(3) $\Longrightarrow$ (5). The boundedness of the functional Λ, i.e. the inequality $|\Lambda(f)| \leqslant C\|f\|_\infty$, $f \in C_b(E)$, enables us to write Λ in the form

$$\Lambda = \Lambda_1 - \Lambda_2 + i\Lambda_3 - i\Lambda_4$$

in such a way that $\Lambda_1 = (\Re\Lambda)^+$, $\Lambda_2 = (\Re\Lambda)^-$, $\Lambda_3 = (\Im\Lambda)^+$, and $\Lambda_4 = (\Im\Lambda)^-$. From the definitions of these functionals (see the proof of Assertion (2) in Proposition 2.2) Assertion (3) implies that $\lim_{n\to\infty} \Lambda_j(f_n) = 0$, $1 \leqslant j \leqslant 4$, whenever the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E)$ decreases to 0. From Theorem 2.1 we infer that each functional Λ_j, $1 \leqslant j \leqslant 4$, can be represented by a Borel measure μ_j: $\Lambda_j(f) = \int f d\mu_j$, $1 \leqslant j \leqslant 4$, $f \in C_b(E)$. It follows that $\Lambda(f) = \int f d\mu$, $f \in C_b(E)$, where $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$.

(4) $\Longrightarrow$ (5). From the apparently weaker hypotheses in Assertion (4) compared to (3) we still have to prove that the functionals Λ_j, $1 \leqslant j \leqslant 4$, as described in the implication (3) $\Longrightarrow$ (5) have the property that $\lim_{n\to\infty} \Lambda_j(f_n) = 0$ whenever the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E)$ decreases pointwise to 0. We will give the details for the functional $\Lambda_1 = (\Re\Lambda)^+$. This suffices because $\Lambda_2 = (\Re(-\Lambda))^+$, $\Lambda_3 = (\Re(-i\Lambda))^+$, and $\Lambda_4 = (\Re(i\Lambda))^+$. So let the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E)$ decreases pointwise to 0. Fix $\varepsilon > 0$, and choose $0 \leqslant g_1 \leqslant f_1$, $g_1 \in C_b(E)$, in such a way that

$$\Lambda_1(f_1) = (\Re\Lambda)^+(f_1) \leqslant \Re(\Lambda(g_1)) + \frac{1}{2}\varepsilon. \tag{2.18}$$

Then we choose a sequence of functions $(u_k)_{k\in\mathbb{N}} \subset C_b^+(E)$ such that $g_1 = \sup_{n\in\mathbb{N}} \sum_{k=1}^n u_k = \sum_{k=1}^\infty u_k$ (which is a pointwise increasing limit), and such that $u_k \leqslant f_k - f_{k+1}$, $k \in \mathbb{N}$. In Lemma 2.1 below we will show that such a decomposition is possible. Then $g_1 - \sum_{k=1}^n u_k$ decreases pointwise to 0, and hence by (4) we have

$$\Re\Lambda(g_1) \leqslant \Re\Lambda\left(\sum_{k=1}^n u_k\right) + \frac{1}{2}\varepsilon, \quad \text{for } n \geqslant n_\varepsilon. \tag{2.19}$$

From (2.18) and (2.19) we infer for $n \geqslant n_\varepsilon$ the inequality

$$\begin{aligned}\Lambda_1(f_1) &= (\Re\Lambda)^+(f_1) \leqslant \Re(\Lambda(g_1)) + \frac{1}{2}\varepsilon \\ &\leqslant \Re\Lambda\left(\sum_{k=1}^n u_k\right) + \varepsilon = \sum_{k=1}^n \Re\Lambda(u_k) + \varepsilon \leqslant \sum_{k=1}^n (\Re\Lambda)^+(f_k - f_{k+1}) + \varepsilon \\ &= \sum_{k=1}^n \Lambda_1(f_k - f_{k+1}) + \varepsilon = \Lambda_1(f_1) - \Lambda_1(f_{n+1}) + \varepsilon. \end{aligned} \tag{2.20}$$

From (2.20) we deduce $\Lambda_1(f_n) \leqslant \varepsilon$ for $n \geqslant n_\varepsilon + 1$. Since $\varepsilon > 0$ was arbitrary, this shows $\lim_{n\to\infty} \Lambda_1(f_n) = 0$. This is true for the other linear functionals Λ_2, Λ_3 and Λ_4 as well. As in the proof of the implication (3) $\Longrightarrow$ (5) from Theorem 2.1 it follows that each functional Λ_j, $1 \leqslant j \leqslant 4$, can be represented by a Borel measure μ_j: $\Lambda_j(f) = \int f d\mu_j$, $1 \leqslant j \leqslant 4$, $f \in C_b(E)$. It follows that $\Lambda(f) = \int f d\mu$, $f \in C_b(E)$, where $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$.

(5) $\Longrightarrow$ (1). The proof of Assertion (1) in Proposition 2.2 then shows that the functional Λ belongs to $(C_b(E), \mathcal{T}_\beta)^*$.

This proves Corollary 2.1. □

Lemma 2.1. *Let the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E)$ decrease pointwise to 0, and $0 \leqslant g \leqslant f_1$ be a continuous function. Then there exists a sequence of continuous functions $(u_k)_{k\in\mathbb{N}}$ such that $0 \leqslant u_k \leqslant f_k - f_{k+1}$, $k \in \mathbb{N}$, and such that $g = \sup_{n\in\mathbb{N}} \sum_{k=1}^n u_k = \sum_{k=1}^\infty u_k$ which is a pointwise monotone increasing limit.*

Proof. We write $g = v_1 = u_1 + v_2 = \sum_{k=1}^n u_k + v_{n+1}$, and $v_{n+1} = u_{n+1} + v_{n+2}$ where $u_1 = g \wedge (f_1 - f_2)$, $u_{n+1} = v_{n+1} \wedge (f_{n+1} - f_{n+2})$, and $v_{n+2} = v_{n+1} - u_{n+1}$. Then $0 \leqslant v_{n+1} \leqslant v_n \leqslant f_n$. Since the sequence $(f_n)_{n\in\mathbb{N}}$ decreases to 0, the sequence $(v_n)_{n\in\mathbb{N}}$ also decreases to 0, and thus $g = \sup_{n\in\mathbb{N}} \sum_{k=1}^n u_k$.

The latter shows Lemma 2.1. □

In the sequel we write $\mathcal{M}(E)$ for the complex vector space of all complex Borel measures on the Polish space E. The space is supplied with the weak topology $\sigma(\mathcal{E}, C_b(E))$. We also write $\mathcal{M}^+(E)$ for the convex cone of all positive (= non-negative) Borel measures in $\mathcal{M}(E)$. The notation $\mathcal{M}_1^+(E)$ is employed for all probability measures in $\mathcal{M}^+(E)$, and $\mathcal{M}_{\leqslant 1}^+(E)$ stands for all sub-probability measures in $\mathcal{M}^+(E)$. We identify the space $\mathcal{M}(E)$ and the space $(C_b(E), \mathcal{T}_\beta)^*$.

Theorem 2.2. *Let M be a subset of $\mathcal{M}(E)$ with the property that for every sequence $(\Lambda_n)_{n\in\mathbb{N}}$ in M there exists a subsequence $(\Lambda_{n_k})_{k\in\mathbb{N}}$ such that*

$$\lim_{k\to\infty} \sup_{0\leqslant f\leqslant \mathbf{1}} \Re\left(i^\ell \Lambda_{n_k}(f)\right) = \sup_{0\leqslant f\leqslant \mathbf{1}} \Re\left(i^\ell \Lambda(f)\right), \quad 0 \leqslant \ell \leqslant 3,$$

for some $\Lambda \in \mathcal{M}(E)$. Then M is a relatively weakly compact subset of $\mathcal{M}(E)$ if and only if it is equi-continuous viewed as a subset of the dual space of $(C_b(E), \mathcal{T}_\beta)^*$.

Proof. First suppose that M is relatively weakly compact. Since the weak topology on $\mathcal{M}(E)$ restricted to compact subsets is metrizable and separable, the weak closure of M is bounded for the variation norm. Without loss of generality we may and do assume that M itself is weakly compact. Fix $f \in C_b(E)$, $f \geqslant 0$. Consider the mapping $\Lambda \mapsto (\Re\Lambda)^+(f)$,

$\Lambda \in \mathcal{M}(E)$. Here we identify $\Lambda = \Lambda_\mu \in (C_b(E), \mathcal{T}_\beta)^*$ and the corresponding complex Borel measure $\mu = \mu_\Lambda$ given by the equality $\Lambda(g) = \int g d\mu$, $g \in C_b(E)$. The mapping $\Lambda \mapsto (\Re\Lambda)^+(f)$, $\Lambda \in \mathcal{M}(E)$, is weakly continuous. This can be seen as follows. Suppose $\Lambda_n(g) \to \Lambda(g)$ for all $g \in C_b(E)$. Then $(\Re\Lambda_n)^+(f) \geqslant \Re\Lambda_n(g)$ for all $0 \leqslant g \leqslant f$, $g \in C_b(E)$, and hence

$$\liminf_{n\to\infty} (\Re\Lambda_n)^+(f) \geqslant \liminf_{n\to\infty} \Re\Lambda_n(g) = (\Re\Lambda)(g).$$

It follows that

$$\liminf_{n\to\infty} (\Re\Lambda_n)^+(f) \geqslant \sup_{0\leqslant g\leqslant f} (\Re\Lambda)(g) = (\Re\Lambda)^+(f).$$

Since $\lim_{n\to\infty} (\Re\Lambda_n)^+(\mathbf{1}) = (\Re\Lambda)^+(\mathbf{1})$ we also have

$$\liminf_{n\to\infty} (\Re\Lambda_n)^+(\mathbf{1}-f) \geqslant \sup_{0\leqslant g\leqslant \mathbf{1}-f} (\Re\Lambda)(g) = (\Re\Lambda)^+(\mathbf{1}-f).$$

Hence we see $\limsup_{n\to\infty} (\Re\Lambda_n)^+(f) \leqslant (\Re\Lambda)^+(f)$, which completes the proof of Theorem 2.2. □

In what follows we write $\mathcal{K}(E)$ for the collection of compact subsets of E. Theorem 2.3 gives an alternative description of a tight family of measures: see Definition 2.1 below as well.

Theorem 2.3. *Let M be a subset of $\mathcal{M}(E)$. Then the following assertions are equivalent:*

(a) For every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ which decreases pointwise to the zero function the equality $\inf_{n\in\mathbb{N}} \sup_{\mu\in M} \int f_n d|\mu| = 0$ holds;

(b) The equality $\inf_{K\in\mathcal{K}(E)} \sup_{\mu\in M} |\mu|(E\backslash K) = 0$ holds, and $\sup_{\mu\in M} |\mu|(E) < \infty$;

(c) There exists a function $u \in H^+(E)$ such that for all $f \in C_b(E)$ and for all $\mu \in M$ the inequality $\left|\int f d\mu\right| \leqslant \|uf\|_\infty$ holds.

Moreover, if $M \subset \mathcal{M}(E)$ satisfies one of the equivalent conditions (a), (b) or (c), then M is relatively weakly compact.

Let $\Lambda : C_b(E) \to \mathbb{C}$ be a linear functional such that $\inf_{n\in\mathbb{N}} |\Lambda|(f_n) = 0$ for every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ which decreases pointwise to zero. Here the linear functional $|\Lambda|$ is defined in such a way that $|\Lambda|(f) = \sup\{|\Lambda(v)| : |v| \leqslant f,\, v \in C_b(E)\}$ for all $f \in C_b^+(E)$. Then by Corollary 2.1 there exists a complex Borel measure μ such that $\Lambda(f) = \int f d\mu$ for all $f \in C_b(E)$. The positive Borel measure $|\mu|$ is such that $|\Lambda|(f) = \int f d|\mu|$ for all $f \in C_b(E)$.

Proof. (a) $\Longrightarrow$ (b). By choosing the sequence $f_n = n^{-1}\mathbf{1}$ we see that $\sup_{\mu\in M} |\mu|(E) < \infty$. Next let ρ be a metric on E for which it is a Polish space, let $(x_n)_{n\in\mathbb{N}}$ be a dense sequence in E, and put

$$B_{k,n} = \left\{x \in E : \rho(x, x_k) \leqslant 2^{-n}\right\}.$$

Choose continuous functions $w_{k,n} \in C_b(E)$ such that $\mathbf{1}_{B^c_{k,n}} \leqslant w_{k,n} \leqslant \mathbf{1}_{B^c_{k,n+1}}$. Put $v_{\ell,n} = \min_{1\leqslant k\leqslant \ell} w_{k,n}$. Then for every $n \in \mathbb{N}$ the sequence $\ell \mapsto v_{\ell,n}$ decreases pointwise to zero. So for given $\varepsilon > 0$ and for given $n \in \mathbb{N}$ there exists $\ell_n(\varepsilon)$ such that $\int v_{\ell_n(\varepsilon),n} d|\mu| \leqslant \varepsilon 2^{-n}$ for all $\mu \in M$. It follows that $|\mu|\left(\bigcap_{k=1}^{\ell_n(\varepsilon)} B^c_{k,n}\right) \leqslant \varepsilon 2^{-n}$, and hence

$$|\mu|\left(\bigcup_{n=1}^{\infty}\bigcap_{k=1}^{\ell_n(\varepsilon)} B^c_{k,n}\right) \leqslant \varepsilon, \quad \mu \in M.$$

Put $K(\varepsilon) = \bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\ell_n(\varepsilon)} B_{k,n}$. Then $K(\varepsilon)$ is closed, and thus complete, and completely ρ-bounded. Hence it is compact. Moreover, $|\mu|(E\backslash K(\varepsilon)) \leqslant \varepsilon$ for all $\mu \in M$. Hence (b) follows from (a).

(b) $\Longrightarrow$ (c). This proof follows the lines of proof of Assertion (1) of Proposition 2.2. Instead of considering just one measure we now have a family of measures M.

(c) $\Longrightarrow$ (a). Essentially speaking this is a consequence of Dini's lemma. Here we use the following fact. If for some $\mu \in \mathcal{M}(E)$ the inequality $\left|\int f d\mu\right| \leqslant \|uf\|_\infty$ holds for all $f \in C_b(E)$, then we also have $\left|\int f d|\mu|\right| \leqslant \|uf\|_\infty$ for all $f \in C_b(E)$. Fix $\alpha > 0$. If $(f_n)_{n\in\mathbb{N}}$ is any sequence in $C_b^+(E)$ which decreases pointwise to zero, then for $\mu \in M$ we have the following estimate

$$\begin{aligned}\int f_n d|\mu| &\leqslant \max\left(\left\|u\mathbf{1}_{\{u\geqslant\alpha\}}f_n\right\|_\infty, \left\|u\mathbf{1}_{\{u<\alpha\}}f_n\right\|_\infty\right)\\ &\leqslant \max\left(\|u\|_\infty \sup_{x\in\{u\geqslant\alpha\}} f_n(x), \alpha\sup_{x\in E} f_n(x)\right)\\ &\leqslant \max\left(\|u\|_\infty \sup_{x\in\{u\geqslant\alpha\}} f_n(x), \alpha\sup_{x\in E} f_1(x)\right). \end{aligned} \tag{2.21}$$

Because of the fact that the set $\{u \geqslant \alpha\}$ is contained in a compact subset of E from (2.21) and Dini's lemma we deduce that $\inf_{n\in\mathbb{N}}\sup_{\mu\in M}\int f_n d|\mu| \leqslant \alpha \sup_{x\in E} f_1(x)$ for all $\alpha > 0$. Consequently, (a) follows.

Finally we prove that if M satisfies (c), then M is relatively weakly compact. First observe that $\mu \in M$ implies $|\mu|(E) \leqslant \|u\|_\infty$. So the subset

M is uniformly bounded, and since E is a Polish space, the same is true for the ball $\{\mu \in \mathcal{M}(E) : |\mu|(E) \leqslant \|u\|_\infty\}$ endowed with the weak topology. Therefore, if $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in M it contains a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ such that $\Lambda(f) := \lim_{k\to\infty} \int f d\mu_{n_k}$ exists for all $f \in C_b(E)$. Then it follows that $|\Lambda(f)| \leqslant \|uf\|_\infty$ for all $f \in C_b(E)$. Consequently, the linear functional Λ can be represented as a measure: $\Lambda(f) = \int f d\mu$, $f \in C_b(E)$. It follows that the weak closure of the set M is weakly compact.

This completes the proof of Theorem 2.3. □

The following result generalizes Theorem 2.3 to open subsets of E.

Theorem 2.4. *Let M be a subset of $\mathcal{M}(E)$, and let O be an open subset of E. Then the following assertions are equivalent:*

(a) For every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$, which decreases pointwise to zero on the open subset O, the equality $\inf_{n\in\mathbb{N}} \sup_{\mu\in M} \int f_n d|\mu| = 0$ holds;

(b) The equality $\inf_{K\subset O,\, K\in\mathcal{K}(O)} \sup_{\mu\in M} |\mu|(E\backslash K) = 0$ holds, and $\sup_{\mu\in M} |\mu|(E) < \infty$;

(c) There exists a function $u \in H^+(O)$ such that for all $f \in C_b(E)$ and for all $\mu \in M$ the inequality $\left|\int f d\mu\right| \leqslant \|uf\|_\infty$ holds.

(d′) For every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ which decreases pointwise to $\mathbf{1}_{E\backslash O}$ the equality $\inf_{n\in\mathbb{N}} \sup_{\mu\in M} \int f_n d|\mu| = 0$ holds.

Moreover, if $M \subset \mathcal{M}(E)$ satisfies one of the equivalent conditions (a), (b) or (c), then $M\big|_O := \left\{\mu\big|_O : \mu \in M\right\}$ is relatively weakly compact in $\mathcal{M}(O)$.

Proof. An analysis of the proof of Theorem 2.3, adapted to a genuine open subset O instead of E will reveal this equivalence. The balls have to be taken relative to a metric which makes O a Polish space: see the proof of (a) $\Longrightarrow$ (b). The constructed functions f_n are identically one on $E\backslash O$. These arguments suffice to prove Theorem 2.4. □

In the terminology of Varadajan [Varadarajan (1961, 1999)] a functional $\Lambda : C_b(E) \to \mathbb{C}$ is called *smooth* if for every sequence $(f_n)_{n\in\mathbb{N}}$ which decreases pointwise to zero the following equality holds: $\lim_{n\to\infty} \Lambda(f_n) = 0$. So in the following definition we could have said that a family of measures M which satisfies one of the conditions in Theorem 2.4 is uniformly σ-smooth on the open set O instead of "a tight family".

Definition 2.1. A family of complex measures $M \subset \mathcal{M}(E)$ is called tight if it satisfies one of the equivalent conditions in Theorem 2.3 with $O = E$. Let $\widetilde{M}$ be a collection of linear functionals on $C_b(E)$ which are continuous for the strict topology. Then each $\Lambda \in \widetilde{M}$ can be represented by a measure: $\Lambda(f) = \int f d\mu_\Lambda$, $f \in C_b(E)$. The collection $\widetilde{M}$ of linear functionals is called tight, provided the same is true for the family $M = \left\{\mu_\Lambda : \Lambda \in \widetilde{M}\right\}$.

Remark 2.2. In fact if M satisfies (a) in Theorem 2.4, then M satisfies Dini's condition in the sense that a sequence of functions $\mu \mapsto |\mu|(f_n)$ which decreasing pointwise to zero in fact converges uniformly on M. Assertion (b) says that the family M is tight in the usual sense as it can be found in the standard literature. Assertion (c) says that the family M is equicontinuous for the strict topology.

The following corollary says that if for M in Theorem 2.3 we choose a collection of positive measures, then the family M is tight if and only if it is relatively weakly compact. Compare these results with Stroock [Stroock (2000)].

Corollary 2.2. *Let M be a collection of positive Borel measures. Then the following assertions are equivalent:*

(a) The collection M is relatively weakly compact.
(b) The collection M is tight in the sense that $\sup_{\mu\in M}\mu(E) < \infty$ *and* $\inf_{K\in\mathcal{K}(E)}\sup_{\mu\in M}\mu(E\backslash K) = 0$.
(c) There exists a function $u \in H^+(E)$ such that $\left|\int f d\mu\right| \leqslant \|uf\|_\infty$ for all $\mu \in M$ and for all $f \in C_b(E)$.

Remark 2.3. Suppose that the collection M in Corollary 2.2 consists of probability measures and is closed with respect to the Lévy-Prohorov metric. If M satisfies one of the equivalent conditions in Corollary 2.2, then it is a weakly compact subset of $P(E)$, the collection of Borel probability measures on E. For probability measures μ and ν the Lévy-Prohorov metric $d_{\rm LP}(\mu,\nu)$ may be defined by

$$d_{\rm LP}(\mu,\nu) = \inf\{\varepsilon > 0 : \mu(A) \leqslant \nu(A^\varepsilon) + \varepsilon\} \leqslant 1. \tag{2.22}$$

For a subset $A \subseteq E$, define the ε-neighborhood of A by

$$A^\varepsilon := \{x \in E : \text{there exists } y \in A \text{ such that } d(x,y) < \varepsilon\} = \bigcup_{y\in A} B(y,\varepsilon)$$

where $B(y,\varepsilon)$ is the open ball of radius ε centered at y. For more details see Definition 3.2 in Chapter 3.

Proof. Corollary 2.2 follows more or less directly from Theorem 2.3. Let M be as in Corollary 2.2, and $(f_n)_{n\in\mathbb{N}}$ be a sequence in $C_b(E)$ which decreases to the zero function. Then observe that the sequence of functions $\mu \mapsto \int f_n d|\mu| = \int f_n d\mu$, $\mu \in M$, decreases pointwise to zero. Each of these functions is weakly continuous. Hence, if M is relatively weakly compact, then Dini's lemma implies that this sequence converges uniformly on M to zero. It follows that assertion (a) in Corollary 2.2 implies assertion (a) in Theorem 2.3. So we see that in Corollary 2.2 the following implications are valid: (a) $\Longrightarrow$ (b), and (b) $\Longrightarrow$ (c). If $M \subset \mathcal{M}^+(E)$ satisfies (c), then Theorem 2.3 implies that M is relatively weakly compact. This means that the assertions (a), (b) and (c) in Corollary 2.2 are equivalent. □

We will also need the following theorem.

Theorem 2.5. *Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}(E)$ be a tight sequence (see Definition 2.1) with the property that $\Lambda(f) := \lim_{n\to\infty} \int f d\mu_n$ exists for all $f \in C_b(E)$. Let $\Phi \subset C_b(E)$ be a family of functions which is equi-continuous and bounded. Then Λ can be represented as a complex Borel measure μ, and*

$$\lim_{n\to\infty} \sup_{\varphi\in\Phi} \left| \int \varphi d\mu_n - \int \varphi d\mu \right| = 0.$$

Remark 2.4. According to the Theorem of Arzela-Ascoli an equi-continuous and uniformly bounded family of functions restricted to a compact subset K is relatively compact in $C_b(K)$.

Proof. The fact that the linear functional Λ can be represented by a Borel measure follows from Corollary 2.1 and Theorem 2.3. Assume to arrive at a contradiction that

$$\limsup_{n\to\infty} \sup_{\varphi\in\Phi} \left| \int \varphi d\mu_n - \int \varphi d\mu \right| > 0.$$

Then there exist $\varepsilon > 0$, a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$, and a sequence $(\varphi_k)_{k\in\mathbb{N}} \subset \Phi$ such that

$$\left| \int \varphi_k d\mu_{n_k} - \int \varphi_k d\mu \right| > \varepsilon, \quad k \in \mathbb{N}. \tag{2.23}$$

Choose a compact subset of E in such a way that

$$\sup_{\varphi\in\Phi} \|\varphi\|_\infty \times \sup_{n\in\mathbb{N}} |\mu_n| (E\backslash K) \leqslant \frac{\varepsilon}{16}. \tag{2.24}$$

By the Bolzano-Weierstrass theorem for bounded equi-continuous families of functions, there exists a continuous function $\varphi^K \in C(K)$ and a subsequence of the sequence $(\varphi_k)_{k\in\mathbb{N}}$, which we call again $(\varphi_k)_{k\in\mathbb{N}}$, such that

$$\lim_{k\to\infty} \sup_{x\in K} \left|\varphi_k(x) - \varphi^K(x)\right| = 0. \tag{2.25}$$

By Tietze's extension theorem there exists a continuous function $\varphi \in C_b(E)$ such that φ restricted to K coincides with φ^K and such that $|\varphi| \leqslant 2 \sup_{\psi \in \Phi} \|\psi\|_\infty$. From (2.25) it follows there exists $k_\varepsilon \in \mathbb{N}$ such that for $k \geqslant k_\varepsilon$ the inequality

$$\sup_{n \in \mathbb{N}} |\mu_n| (E) \left\| \mathbf{1}_K (\varphi_k - \varphi) \right\|_\infty \leqslant \frac{\varepsilon}{8}. \tag{2.26}$$

From (2.24)and (2.26) we obtain the following estimate:

$$\begin{aligned}
&\left| \int \varphi_k d\mu_{n_k} - \int \varphi_k d\mu \right| \\
&\leqslant \left| \int_K (\varphi_k - \varphi) \, d\mu_{n_k} - \int_K (\varphi_k - \varphi) \, d\mu \right| \\
&\quad + \left| \int_{E \setminus K} (\varphi_k - \varphi) \, d\mu_{n_k} - \int_{E \setminus K} (\varphi_k - \varphi) \, d\mu \right| + \left| \int \varphi d\mu_{n_k} - \int \varphi d\mu \right| \\
&\leqslant \left\| \mathbf{1}_K (\varphi_k - \varphi) \right\|_\infty \left(|\mu_{n_k}| (K) + |\mu| (K) \right) \\
&\quad + 4 \sup_{\psi \in \Phi} \|\psi\|_\infty \left(|\mu_{n_k}| (E \setminus K) + |\mu| (E \setminus K) \right) + \left| \int \varphi d\mu_{n_k} - \int \varphi d\mu \right| \\
&\leqslant 2 \left\| \mathbf{1}_K (\varphi_k - \varphi) \right\|_\infty \sup_{k \in \mathbb{N}} |\mu_{n_k}| (K) \\
&\quad + 8 \sup_{\psi \in \Phi} \|\psi\|_\infty \sup_{k \in \mathbb{N}} |\mu_{n_k}| (E \setminus K) + \left| \int \varphi d\mu_{n_k} - \int \varphi d\mu \right| \\
&\leqslant \frac{3}{4} \varepsilon + \left| \int \varphi d\mu_{n_k} - \int \varphi d\mu \right|.
\end{aligned} \tag{2.27}$$

Since $\lim_{n \to \infty} \left| \int \varphi d\mu_n - \int \varphi d\mu \right| = 0$ the equality in (2.27) implies

$$\left| \int \varphi_k d\mu_{n_k} - \int \varphi_k d\mu \right| < \varepsilon \tag{2.28}$$

for k large enough. The conclusion in (2.28) contradicts our assumption in (2.23).

This proves Theorem 2.5. □

Occasionally we will need the following version of the Banach-Alaoglu theorem; see e.g. Theorem 8.4. We use the notation $\langle f, \mu \rangle = \int_E f(x) \, d\mu(x)$, $f \in C_b(E)$, $\mu \in M(E)$. For a proof of the following theorem we refer to e.g. [Rudin (1991)]. Notice that any $\mathcal{T}_\beta$-equi-continuous family of measures is contained in B_u for some $u \in H(E)$. Here B_u is the collection defined in (2.29) below.

Theorem 2.6. *(Banach-Alaoglu) Let u be a function in $H(E)$, and define the subset B_u of $M(E)$ by*

$$B_u = \{\mu \in M(E) : |\langle f, \mu\rangle| \leqslant \|uf\|_\infty \quad \text{for all} \quad f \in C_b(E)\}. \tag{2.29}$$

Then B_u is $\sigma(M(E), C_b(E))$-compact.

Since the space $(C_b(E), \mathcal{T}_\beta)$ is separable, it follows that for every sequence $(\mu_n)_{n\in\mathbb{N}}$ in B_u there exists a measure $\mu \in M(E)$ and a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \langle f, \mu_{n_k}\rangle = \langle f, \mu\rangle$ for all $f \in C_b(E)$.

Instead of "$\sigma(M(E), C_b(E))$"-convergence we often write "weak*-convergence", which is a functional analytic term. In a probabilistic context people usually write "weak convergence". Another term which is in use is "convergence relative to the Bernoulli topology": see e.g. [Bloom and Heyer (1995)] and [Berg and Forst (1975)].

2.1.3 *Integral operators on the space of bounded continuous functions*

We insert a short digression to operator theory. Let E_1 and E_2 be two Polish spaces, and let $T : C_b(E_1) \to C_b(E_2)$ be a linear operator with the property that its absolute value $|T| : C_b(E_1) \to C_b(E_2)$ determined by the equality

$$|T|(f) = \sup\{|Tg| : |g| \leqslant f\}, \quad f \in C_b(E_1),\ f \geqslant 0,$$

is well-defined and acts as a linear operator from $C_b(E_1)$ to $C_b(E_2)$. Endow the spaces $C_b(E_1)$ and $C_b(E_2)$ with the strict topology, and let the symbol $\mathcal{L}(C_b(E_1), C_b(E_2))$ denote the space of linear operators which are continuous for the respective strict topologies.

Definition 2.2. A family of linear operators $\{T_\alpha : \alpha \in A\}$, where every $T_\alpha \in \mathcal{L}(C_b(E_1), C_b(E_2))$ is called equi-continuous for the strict topology if for every $v \in H(E_2)$ there exists $u \in H(E_1)$ such that the inequality $\|vT_\alpha f\|_\infty \leqslant \|uf\|_\infty$ holds for all $\alpha \in A$ and for all $f \in C_b(E_1)$.

So the notion "equi-continuous for the strict topology" has a functional analytic flavor.

Definition 2.3. A family of linear operators $\{T_\alpha : \alpha \in A\}$, where every T_α belongs to $\mathcal{L}(C_b(E_1), C_b(E_2))$, is called tight if for every compact subset K of E_2 the family of functionals $\{\Lambda_{\alpha,x} : \alpha \in A,\ x \in K\}$ is tight in the sense of Definition 2.1. Here the functional $\Lambda_{\alpha,x} : C_b(E_1) \to \mathbb{C}$ is defined by

$\Lambda_{\alpha,x}(f) = T_\alpha f(x)$, $f \in C_b(E_1)$. Its absolute value $|\Lambda_{\alpha,x}|$ has then the property that $|\Lambda_{\alpha,x}|(f) = |T_\alpha| f(x)$, $f \in C_b(E_1)$.

The following theorem says that a tight family of operators $\{T_\alpha : \alpha \in A\}$ is equi-continuous for the strict topology and vice versa. Both spaces E_1 and E_2 are supposed to be Polish.

Theorem 2.7. *Let A be some index set, and let for every $\alpha \in A$ the mapping $T_\alpha : C_b(E_1) \to C_b(E_2)$ be a linear operator, which is continuous for the uniform topology. Suppose that the family $\{T_\alpha : \alpha \in A\}$ is tight. Then for every $v \in H(E_2)$ there exists $u \in H(E_1)$ such that*

$$\|vT_\alpha f\|_\infty \leqslant \|uf\|_\infty, \quad \textit{for every } \alpha \in A \textit{ and for all } f \in C_b(E_1). \tag{2.30}$$

Conversely, if the family $\{T_\alpha : \alpha \in A\}$ is equi-continuous in the sense that for every $v \in H(E_2)$ there exists $u \in H(E_1)$ such that (2.30) is satisfied. Then the family $\{T_\alpha : \alpha \in A\}$ is tight.

If the family $\{T_\alpha : \alpha \in A\}$ satisfies (2.30), then the family $\{|T_\alpha| : \alpha \in A\}$ satisfies the same inequality with $|T_\alpha|$ instead of T_α. The argument to see this goes in more or less the same way as we will prove the first part of Proposition 2.7 below. Fix $f \in C_b(E_1)$, $\alpha \in A$, and $x \in E_1$, and let the functions $u \in H(E_1)$ and $v \in H(E_2)$ be such that (2.30) is satisfied. Choose $\vartheta \in [-\pi, \pi]$ in such a way that

$$|v(x)\,|T_\alpha|\,(f)(x)| = |v(x)|\,|T_\alpha|\left(\Re\left(e^{i\vartheta}f\right)\right)(x) \leqslant |v(x)|\,|T_\alpha|\left(\Re\left(e^{i\vartheta}f\right)^+\right)(x)$$

(definition of $|T_\alpha|$)

$$\begin{aligned} &= \sup\left\{|v(x)T_\alpha g(x)| : |g| \leqslant \Re\left(e^{i\vartheta}f\right)^+\right\} \\ &\leqslant \sup\left\{\|ug\|_\infty : |g| \leqslant \Re\left(e^{i\vartheta}f\right)^+\right\} \leqslant \|uf\|_\infty . \end{aligned} \tag{2.31}$$

From (2.31) we see that the inequality in (2.30) is also satisfied for the operators $|T_\alpha|$, $\alpha \in A$.

Corollary 2.3. *Like in Theorem 2.7 let A be some index set, and let for every $\alpha \in A$ the mapping $T_\alpha : C_b(E_1) \to C_b(E_2)$ be a positivity preserving linear operator. Then the family $\{T_\alpha : \alpha \in A\}$ is $\mathcal{T}_\beta$-equi-continuous if and only if for every sequence $(\psi_m)_{m\in\mathbb{N}}$ which decreases pointwise to 0, the sequence $\{T_\alpha(\psi_m f) : m \in \mathbb{N}\}$ decreases pointwise to 0 uniformly in $\alpha \in A$.*

Proof. [Proof of Corollary 2.3.] Choose $v \in H^+(E)$. The proof follows by considering the family of functionals $\Lambda_{\alpha,x} : C_b(E) \to \mathbb{C}$, $\alpha \in A$, $x \in E$, defined by $\Lambda_{\alpha,x}f(x) = v(x)T_\alpha f(x)$, $f \in C_b(E)$. If the family $\{T_\alpha : \alpha \in A\}$ is $\mathcal{T}_\beta$-equi-continuous, then the family $\{\Lambda_{\alpha,x} : \alpha \in A, x \in E\}$ is tight. For example, it then easily follows that $\{\Lambda_{v,\alpha,x}f_m : \alpha \in A, x \in E\}$ converges uniformly in $\alpha \in A$, $x \in E$, to 0, provided that the sequence $(f_m)_{m\in\mathbb{N}}$ decreases pointwise to 0. Conversely, suppose that for any given $v \in H^+(E)$, and for any sequence of functions $(f_m)_{m\in\mathbb{N}} \subset C_b(E)$ which decreases pointwise to 0, the sequence $\{\Lambda_{v,\alpha,x}f_m : \alpha \in A, x \in E\}_{m\in\mathbb{N}}$ converges uniformly to 0. Then the family $\{\Lambda_{\alpha,x} : \alpha \in A, x \in E\}$ is tight: see Theorem 2.7. This completes the proof of Corollary 2.3. □

Proof. [Proof of Theorem 2.7.] Like in Definition 2.3 the functionals $\Lambda_{\alpha,x}$, $\alpha \in A$, $x \in E_1$, are defined by $\Lambda_{\alpha,x}(f) = [T_\alpha f](x)$, $f \in C_b(E_1)$. First we suppose that the family $\{T_\alpha : \alpha \in A\}$ is tight. Let $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E_1)$ be sequence of continuous functions which decreases pointwise to zero, and let $v \in H(E_2)$ be arbitrary. Since the family $\{T_\alpha : \alpha \in A\}$ is tight, it follows that, for every compact subset K the collection of functionals $\{\Lambda_{\alpha,x} : \alpha \in A, x \in K\}$ is tight. Then, since the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E_1)$ decreases pointwise to zero, we have

$$\lim_{n\to\infty} \sup_{\alpha\in A,\, x\in K} |\Lambda_{\alpha,x}|(f_n) = 0 \quad \text{for every compact subset } K \text{ of } E_1. \tag{2.32}$$

From (2.32) it follows that $\lim_{n\to\infty} \sup_{\alpha\in A,\, x\in K} |v(x)|\,|\Lambda_{\alpha,x}|(f_n) = 0$. Hence the family of functionals $\{|v(x)|\,\Lambda_{\alpha,x} : \alpha \in A, x \in E_1\}$ is tight. By Theorem 2.3 (see Definition 2.1 as well) it follows that there exists a function $u \in H(E_1)$ such that

$$|v(x)\,[T_\alpha f](x)| = |v(x)\Lambda_{\alpha,x}(f)| \leqslant \|uf\|_\infty \tag{2.33}$$

for all $f \in C_b(E_1)$, for all $x \in E$ and for all $\alpha \in A$. The inequality in (2.33) implies the equi-continuity property (2.30).

Next let the family $\{T_\alpha : \alpha \in A\}$ be equi-continuous in the sense that it satisfies inequality (2.30). Then the same inequality holds for the family $\{|T_\alpha| : \alpha \in A\}$; the argument was given just prior to the proof of Theorem 2.7. Let K be any compact subset of E_1 and let $(f_n)_{n\in\mathbb{N}} \subset C_b^+(E_1)$ be a sequence which decreases to zero. Then there exists a function $u \in H(E_1)$ such that

$$\sup_{\alpha\in A,\, x\in K} [|T_\alpha|\, f_n](x) = \|\mathbf{1}_K\, |T_\alpha|\, f_n\|_\infty \leqslant \|uf_n\|_\infty\,. \tag{2.34}$$

From (2.34) it readily follows that $\lim_{n\to\infty} \sup_{\alpha\in A,\, x\in K} [|T_\alpha|\, f_n](x) = 0$. By Definition 2.3 it follows that the family $\{T_\alpha : \alpha \in A\}$ is tight.

This completes the proof of Theorem 2.7. □

Theorem 2.8. *Let E_1 and E_2 be two Polish spaces, and let $U : C_b(E_1,\mathbb{R}) \to C_b(E_2,\mathbb{R})$ be a mapping with the following properties:*

(1) If f_1 and $f_2 \in C_b(E_1)$ are such that $f_1 \leqslant f_2$, then $U(f_1) \leqslant U(f_2)$. In other words the mapping $f \mapsto Uf$, $f \in C_b(E_1,\mathbb{R})$ is monotone.
(2) If f_1 and f_2 belong to $C_b(E_1,\mathbb{R})$, and if $\alpha \geqslant 0$, then $U(f_1+f_2) \leqslant U(f_1)+U(f_2)$, and $U(\alpha f_1) = \alpha U(f_1)$.
(3) U is unit preserving: $U(\mathbf{1}_{E_1}) = \mathbf{1}_{E_2}$.
(4) If $(f_n)_{n\in\mathbb{N}} \subset C_b(E_1,\mathbb{R})$ is a sequence which decreases pointwise to zero, then so does the sequence $(U(f_n))_{n\in\mathbb{N}}$.

Then for every $v \in H^+(E_2)$ there exists $u \in H^+(E_1)$ such that

$$\sup_{y\in E_2} v(y)U(\Re f)(y) \leqslant \sup_{x\in E_1} u(x)\Re f(x), \quad \text{for all } f \in C_b(E_1) \text{ and hence}$$

$$\sup_{y\in E_2} v(y)U|f|(y) \leqslant \sup_{x\in E_1} u(x)|f(x)|, \quad \text{for all } f \in C_b(E_1). \tag{2.35}$$

If the mapping U maps $C_b(E_1)$ to $L^\infty(E,\mathbb{R},\mathcal{E})$, then the conclusion about its continuity as described in (2.35) is still true provided it possesses the above properties (1), (2), (3), and (4) is replaced by

(4′) If $(f_n)_{n\in\mathbb{N}} \subset C_b(E_1,\mathbb{R})$ is a sequence which decreases pointwise to zero, then the sequence $(U(f_n))_{n\in\mathbb{N}}$ decreases to zero uniformly on compact subsets of E_2.

Proof. Put

$$M_{vU}^{\Re} = \Big\{\nu \in M^+(E_1) : \nu(E_1) = \sup_{y\in E_2} v(y),\ \Re\langle g,\nu\rangle \leqslant \sup_{y\in E_2} v(y)(U\Re g)(y) \quad \text{for all } g \in C_b(E_1)\Big\} \quad \text{and}$$

$$M_{vU}^{|\cdot|} = \Big\{\nu \in M^+(E_1) : \nu(E_1) = \sup_{y\in E_2} v(y),\ |\langle g,\nu\rangle| \leqslant \sup_{y\in E_2} v(y)(U|g|)(y) \quad \text{for all } g \in C_b(E_1)\Big\}. \tag{2.36}$$

A combination of Theorem 2.3 and its Corollary 2.2 shows that the collections $M_{vU}^{\Re}$ and $M_{vU}^{|\cdot|}$ are tight. Here we use hypothesis (4). We also observe that $M_{vU}^{\Re} = M_{vU}^{|\cdot|}$. This can be seen as follows. First suppose that $\nu \in M_{vU}^{|\cdot|}$ and choose $g \in C_b(E_1)$. Then we have

$$\langle \Re g + \|\Re g\|_\infty, \nu\rangle \leqslant \sup_{y\in E_2} v(y)(U|\Re g + \|g\|_\infty|)(y)$$

$$\begin{aligned}
&\leqslant \sup_{y\in E_2} \left(v(y)U\left(\Re g\right)(y)\right) + \sup_{y\in E_2} v(y)\left\|g\right\|_\infty \\
&= \sup_{y\in E_2} \left(v(y)U\left(\Re g\right)(y)\right) + \nu\left(E_1\right)\left\|g\right\|_\infty . \qquad (2.37)
\end{aligned}$$

From (2.37) we deduce $\Re\langle g,\nu\rangle \leqslant \sup_{y\in E_1}\left(v(y)U\left(\Re g\right)(y)\right)$, and hence $M_{vU}^{|\cdot|}\subset M_{vU}^{\Re}$. The reverse inclusion is shown by the following arguments:

$$\begin{aligned}
\left|\langle g,\nu\rangle\right| &= \sup_{\vartheta\in[-\pi,\pi]} \left\langle \Re\left(e^{i\vartheta}g\right),\nu\right\rangle \\
&\leqslant \sup_{\vartheta\in[-\pi,\pi]} \sup_{y\in E_2} v(y)U\left(\left|\Re\left(e^{i\vartheta}g\right)\right|\right)(y) \\
&\leqslant \sup_{\vartheta\in[-\pi,\pi]} \sup_{y\in E_2} v(y)U\left(|g|\right)(y) = \sup_{y\in E_2} v(y)U\left(|g|\right)(y). \qquad (2.38)
\end{aligned}$$

From (2.38) the inclusion $M_{vU}^{\Re}\subset M_{vU}^{|\cdot|}$ follows. So from now on we will write $M_{vU}=M_{vU}^{\Re}=M_{vU}^{|\cdot|}$. There exists a function $u\in H^+(E)$ such that for all $f\in C_b(E)$ and for all $\mu\in M$ the inequality $\Re\int f d\mu \leqslant \sup_{x\in E}\Re\left(u(x)f(x)\right)$ holds. The result in Theorem 2.8 is a consequence of the following equalities

$$\sup_{y\in E_2} v(y)U\Re f(y) = \sup\left\{\Re\langle f,\nu\rangle : \nu\in M_{vU}\right\}, \quad \text{and} \qquad (2.39)$$

$$\sup_{y\in E_2} v(y)U\left|f\right|(y) = \sup\left\{\left|\langle f,\nu\rangle\right| : \nu\in M_{vU}\right\}. \qquad (2.40)$$

The equality in (2.39) follows from the Theorem of Hahn-Banach. In the present situation it says that there exists a linear functional $\Lambda : C_b\left(E_1,\mathbb{R}\right)\to\mathbb{R}$ such that $\Lambda(f)\leqslant \sup_{y\in E_2} v(y)Uf(y)$, for all $f\in C_b\left(E_1,\mathbb{R}\right)$, and

$$\Lambda\left(\mathbf{1}_{E_1}\right) = \sup_{y\in E_2} v(y)U\left(\mathbf{1}_{E_1}\right)(y) = \sup_{y\in E_2} v(y)\mathbf{1}_{E_2}(y) = \sup_{y\in E_2} v(y). \qquad (2.41)$$

Let $f\in C_b\left(E_1,\mathbb{R}\right)$, $f\leqslant 0$. Then $\Lambda(f)\leqslant \sup_{y\in E_2} v(y)Uf(y)\leqslant 0$. Again using Hypothesis 4 shows that Λ can be identified with a positive Borel measure on E_1, which than belongs to M_{vU}. Consequently, the left-hand side of (2.39) is less than or equal to its right-hand side. Since the reverse inequality is trivial, the equality in (2.39) follows. The equality in (2.40) easily follows from (2.39).

The assertion about a sub-additive mapping U which sends functions in $C_b\left(E_1\right)$ to functions in $L^\infty\left(E,\mathbb{R},\mathcal{E}\right)$ can easily be adopted from the first part of the proof.

This concludes the proof of Theorem 2.8. □

The results in Proposition 2.3 below should be compared with Definition 4.3. We describe two operators to which the results of Theorem 2.8 are applicable. Let L be an operator with domain and range in $C_b(E)$, with the property that for all $\mu > 0$ and $f \in D(L)$ with $\mu f - Lf \geqslant 0$ implies $f \geqslant 0$. There is a close connection between this positivity property (i.e. positive resolvent property) and the maximum principle: see Definition 4.1 and inequality (4.46). In addition, suppose that the constant functions belong to $D(L)$, and that $L\mathbf{1} = 0$. Fix $\lambda > 0$, and define the operators $U_\lambda^j : C_b(E,\mathbb{R}) \to L^\infty(E,\mathbb{R},\mathcal{E})$, $j = 1, 2$, by the equalities ($f \in C_b(E,\mathbb{R})$):

$$U_\lambda^1 f = \sup_{K\in\mathcal{K}(E)} \inf_{g\in D(L)} \left\{ g \geqslant f\mathbf{1}_K : \lambda g - Lg \geqslant 0 \right\}, \quad \text{and} \tag{2.42}$$

$$U_\lambda^2 f = \inf_{g\in D(L)} \left\{ g \geqslant f : \lambda g - Lg \geqslant 0 \right\}. \tag{2.43}$$

Here the symbol $\mathcal{K}(E)$ stands for the collection of all compact subsets of E. Observe that, if $g \in D(L)$ is such that $\lambda g - Lg \geqslant 0$, then $g \geqslant 0$. This follows from the maximum principle.

Proposition 2.3. *Let the operator L be as above, and let the operators U_λ^1 and U_λ^2 be defined by (2.42) and (2.43) respectively. Then the following assertions hold true:*

(a) Suppose that the operator U_λ^1 has the additional property that for every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ which decreases pointwise to zero the sequence $\left(U_\lambda^1 f_n\right)_{n\in\mathbb{N}}$ does so uniformly on compact subsets of E. Then for every $u \in H^+(E)$ there exists a function $v \in H^+(E)$ such that

$$\sup_{x\in E} u(x) U_\lambda^1 f(x) \leqslant \sup_{x\in E} v(x) f(x), \quad \textit{and}$$
$$\sup_{x\in E} u(x) U_\lambda^1 |f|(x) \leqslant \sup_{x\in E} v(x) |f(x)| \quad \textit{for all } f \in C_b(E,\mathbb{R}). \tag{2.44}$$

(b) Suppose that the operator U_λ^2 has the additional property that for every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ which decreases pointwise to zero the sequence $\left(U_\lambda^2 f_n\right)_{n\in\mathbb{N}}$ does so uniformly on compact subsets of E. Then for every $u \in H^+(E)$ there exists a function $v \in H^+(E)$ such that the inequalities in (2.44) are satisfied with U_λ^2 instead of U_λ^1. Moreover, for $f \in D(L^n)$, $\mu \geqslant 0$, and $n \in \mathbb{N}$, the following inequalities hold:

$$\mu^n f \leqslant U_\lambda^2 \left(\left((\lambda+\mu) I - L \right)^n f \right), \quad \textit{and} \tag{2.45}$$
$$\mu^n \|uf\|_\infty \leqslant \left\| v \left((\lambda+\mu) I - L \right)^n f \right\|_\infty. \tag{2.46}$$

In (2.46) the functions u and v are the same as in (2.44) with U_λ^2 replacing U_λ^1.

The inequality in (2.46) could be used to say that the operator L is $\mathcal{T}_\beta$-dissipative: see inequality (4.14) in Definition 4.2. Also notice that $U_\lambda^1(f) \leqslant U_\lambda^2(f)$, $f \in C_b(E, \mathbb{R})$. It is not clear, under what conditions $U_\lambda^1(f) = U_\lambda^2(f)$. In Proposition 2.4 below we will return to this topic. The mapping U_λ^1 is heavily used in the proof of (iii) $\Longrightarrow$ (i) of Theorem 4.3. If the operator L in Proposition 2.3 satisfies the conditions spelled out in assertion (a), then it is called sequentially λ-dominant: see Definition 4.3.

Proof. The assertion in (a) and the first assertion in (b) is an immediate consequence of Theorem 2.8. Let $f \in D(L)$ be real-valued. The inequality (2.46) can be obtained by observing that

$$\begin{aligned}
&U_\lambda^2\left(\left(\lambda+\mu\right)I-L\right)f \\
&= \inf_{g\in D(L)}\left\{g \geqslant \left(\left(\lambda+\mu\right)I-L\right)f : \lambda g - Lg \geqslant 0\right\} \\
&= \inf_{g\in D(L)}\left\{g \geqslant \left(\left(\lambda+\mu\right)I-L\right)f : \right. \\
&\qquad\qquad \left.\left(\lambda+\mu\right)g - Lg \geqslant \mu g \geqslant \left(\left(\lambda+\mu\right)I-L\right)\left(\mu f\right)\right\} \\
&= \inf_{g\in D(L)}\left\{g \geqslant \left(\left(\lambda+\mu\right)I-L\right)f : \lambda g - Lg \geqslant 0,\ g \geqslant \mu f\right\} \geqslant \mu f. \qquad (2.47)
\end{aligned}$$

Repeating the arguments which led to (2.47) will show the inequality in (2.45). From (2.47) and (2.44) with U_λ^2 instead of U_λ^1 we obtain

$$\begin{aligned}
\sup_{x\in E} u(x)\left(\mu^n f\right)(x) &\leqslant \sup_{x\in E} U_\lambda^2\left(\left(\lambda+\mu\right)f - Lf\right)(x) \\
&\leqslant \sup_{x\in E} v(x)\left((\lambda+\mu)I-L\right)^n f(x), \qquad (2.48)
\end{aligned}$$

for $\mu \geqslant 0$ and $f \in D\left(L^n\right)$. The inequality in (2.46) is an easy consequence of (2.48). This concludes the proof of Proposition 2.3. □

The following proposition is used to show that the semigroup generated by the operator L is $\mathcal{T}_\beta$-equi-continuous: see Theorem 4.3.

Proposition 2.4. *Let the operator L with domain and range in $C_b(E)$ have the following properties:*

(1) For every $\lambda > 0$ the range of $\lambda I - L$ coincides with $C_b(E)$, and the inverse $R(\lambda) := (\lambda I - L)^{-1}$ exists as a positivity preserving bounded linear operator from $C_b(E)$ to $C_b(E)$. Moreover, $0 \leqslant f \leqslant \mathbf{1}$ implies $0 \leqslant \lambda R(\lambda) f \leqslant \mathbf{1}$.

(2) The equality $\lim_{\lambda\to\infty} \lambda R(\lambda) f(x) = f(x)$ holds for every $x \in E$, and $f \in C_b(E)$.

(3) If $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ is any sequence which decreases pointwise to zero, then for every $\lambda > 0$ the sequence $(\lambda R(\lambda) f_n)_{n\in\mathbb{N}}$ decreases to zero as well.

Fix $\lambda > 0$, and define the mappings U_λ^1 and U_λ^2 as in (2.42) and (2.43) respectively. Then the (in-)equalities

$$\sup\left\{(\mu R(\lambda+\mu))^k f;\, \mu > 0,\, k \in \mathbb{N}\right\} \leqslant U_\lambda^1(f) \leqslant U_\lambda^2(f) \tag{2.49}$$

hold for $f \in C_b(E,\mathbb{R})$. Suppose that $f \geqslant 0$. If the function in the left extremity of (2.49) belongs to $C_b(E)$, then the first two terms in (2.49) are equal. If it belongs to $D(L)$, then all three quantities in (2.49) are equal.

From the proof of Proposition 2.4, the following corollary is immediate: see see (2.50) below.

Corollary 2.4. *Let $\lambda_0 > 0$. Suppose that the family $\{\lambda R(\lambda) : \lambda \geqslant \lambda_0\}$ has the properties (2) and (3) of Proposition 2.4. Then the family of operators $\{\lambda R(\lambda) : \lambda \geqslant \lambda_0\}$ is $\mathcal{T}_\beta$-equi-continuous.*

Proof. [Proof of Proposition 2.4.] First we observe that for every $(\lambda, x) \in (0,\infty) \times E$ there exists a Borel measure $B \mapsto r(\lambda, x, B)$ such that $\lambda r(\lambda, x, E) \leqslant 1$, and $R(\lambda)f(x) = \int_E f(y) r(\lambda, x, dy)$, $f \in C_b(E)$. This result follows by considering the functional $\Lambda_{\lambda,x} : C_b(E) \to \mathbb{C}$, defined by $\Lambda_{\lambda,x}(f) = R(\lambda)f(x)$. In fact

$$r(\lambda, x, B) = \sup_{K\in\mathcal{K}(E),\, K\subset B} \inf\{R(\lambda)f(x) : f \geqslant \mathbf{1}_K\}, \quad B \in \mathcal{E}.$$

This result follows from Corollary 2.1. Often we write

$$R(\lambda)(f\mathbf{1}_B) = \int_B f(y) r(\lambda, x, dy), \quad B \in \mathcal{E},\ f \in C_b(E).$$

Observe that the mapping $B \mapsto R(\lambda)(f\mathbf{1}_B)$ is a positive Borel measure on E. Moreover, by Dini's lemma we see that

$$\lim_{n\to\infty} \sup_{\lambda\geqslant\lambda_0} \sup_{x\in K} \lambda R(\lambda) f_n(x) = 0, \quad \lambda_0 > 0, \tag{2.50}$$

whenever the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ decreases pointwise to zero, and K is a compact subset of E. From Theorem 2.7 and its Corollary 2.3 it then follows that the family of operators $\{\lambda R(\lambda) : \lambda \geqslant \lambda_0\}$ is equi-continuous for the strict topology $\mathcal{T}_\beta$, i.e. for every function $u \in H^+(E)$ there exists a function $v \in H^+(E)$ such that

$$\lambda \|uR(\lambda)f\|_\infty \leqslant \|vf\|_\infty \quad \text{for all } \lambda \geqslant \lambda_0 \text{ and all } f \in C_b(E). \tag{2.51}$$

Fix $f \in C_b(E, \mathbb{R})$ and $\lambda > 0$. Next we will prove the

$$U_\lambda^1(f) \geqslant \sup\left\{\left(\mu\left((\lambda+\mu)I-L\right)^{-1}\right)^k f : \mu > 0,\, k \in \mathbb{N}\right\}. \tag{2.52}$$

A version of this proof will be more or less retaken in (4.137) in the proof of the implication (iii) $\Longrightarrow$ (i) of Theorem 4.3 with $D_1 + L$ instead of L. First we observe that for $g \in D(L)$ we have

$$\lambda g(x) - Lg(x) = \lim_{\mu\to\infty} \mu\left(g(x) - \mu R(\lambda+\mu)\, g(x)\right), \quad x \in E. \tag{2.53}$$

If $g \in D(L)$ is such that $\lambda g - Lg \geqslant 0$, then $(\lambda+\mu)\,g - Lg \geqslant \mu g$, and hence $g \geqslant \mu R(\lambda+\mu)g$ for all $\mu > 0$. If $g \geqslant \mu R(\lambda+\mu)g$, then $\mu\left(g - \mu R(\lambda+\mu)g\right) \geqslant 0$, and by (2.53) we see $\lambda g - Lg \geqslant 0$. So that we have the following equality of subsets

$$\{g \in D(L) : \lambda g - Lg \geqslant 0\} = \{g \in D(L) : g \geqslant \mu R(\lambda+\mu)\, g \text{ for all } \mu > 0\}. \tag{2.54}$$

From (2.54) we infer

$$\{g \in D(L) : \lambda g - Lg \geqslant 0\} = \left\{g \in D(L) : g \geqslant \sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu R(\lambda+\mu)\right)^k g\right\}. \tag{2.55}$$

Let $g \in D(L)$ be such that $g \geqslant f\mathbf{1}_K$ and such that $\lambda g - Lg \geqslant 0$, then (2.55) implies $g \geqslant \sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu R(\lambda+\mu)\right)^k (f\mathbf{1}_K)$. Since the operators $\left(\mu R(\lambda+\mu)\right)^k$, $\mu > 0$, $k \in \mathbb{N}$, are integral operators, and bounded Borel measures are inner-regular (with respect to compact subsets), we obtain

$$g \geqslant \sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu R(\lambda+\mu)\right)^k f,$$

and hence

$$\sup_{K\in\mathcal{K}(E)} \inf_{g\in D(L)} \{g \geqslant f\mathbf{1}_K : \lambda g - Lg \geqslant 0\} \geqslant \sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu\left((\lambda+\mu)I-L\right)^{-1}\right)^k f. \tag{2.56}$$

The inequality in (2.56) implies (2.52) and hence, since the inequality $U_\lambda^1(f) \leqslant U_\lambda^2(f)$ is obvious, the inequalities in (2.49) follow. Here we employ the fact that $\lambda g - Lg \geqslant 0$ implies $g \geqslant 0$. Fix a compact subset K of E, and $f \geqslant 0$, $f \in C_b(E)$. If the function $g = \sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu\left((\lambda+\mu)I-L\right)^{-1}\right)^k f$ belongs to $C_b(E)$, then $g \geqslant f\mathbf{1}_K$, and $g \geqslant \mu R(\lambda+\mu)\, g$ for all $\mu > 0$. Hence it follows that

$$\sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu\left((\lambda+\mu)I-L\right)^{-1}\right)^k f$$

$$\geqslant \inf \{g \geqslant f\mathbf{1}_K : g \geqslant \mu R(\lambda+\mu) g, \, g \in C_b(E)\}. \tag{2.57}$$

Next we show that $\tau_\beta\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha) f = f$. From the assumptions (2) and (3), and from (2.51) it follows that $D(L) = R\left((\beta I - L)^{-1}\right)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$. Therefore let g be any function in $D(L)$, and let $u \in H^+(E)$. Consider, for $\alpha > \lambda_0$, the equalities

$$\begin{aligned} f - \alpha R(\alpha) f &= f - g - \alpha R(\alpha)(f-g) + g - \alpha R(\alpha) g \\ &= f - g - \alpha R(\alpha)(f-g) - R(\alpha)(Lg), \end{aligned} \tag{2.58}$$

and the corresponding inequalities

$$\begin{aligned} &\|u(f - \alpha R(\alpha) f)\|_\infty \\ &\leqslant \|u(f-g)\|_\infty + \|u\alpha R(\alpha)(f-g)\|_\infty + \|uR(\alpha)(Lg)\|_\infty \\ &\leqslant \|u(f-g)\|_\infty + \|v(f-g)\|_\infty + \frac{\|u\|_\infty}{\alpha}\|Lg\|_\infty. \end{aligned} \tag{2.59}$$

So that for given $\varepsilon > 0$ we first choose $g \in D(L)$ in such a way that

$$\|u(f-g)\|_\infty + \|v(f-g)\|_\infty \leqslant \frac{2}{3}\varepsilon. \tag{2.60}$$

Then we choose $\alpha_\varepsilon \geqslant \lambda_0$ so large that $\dfrac{\|u\|_\infty}{\alpha_\varepsilon}\|Lg\|_\infty \leqslant \dfrac{1}{3}\varepsilon$. From the latter, (2.59), and (2.60) we conclude:

$$\|u(f - \alpha R(\alpha) f)\|_\infty \leqslant \varepsilon, \quad \text{for } \alpha \geqslant \alpha_\varepsilon. \tag{2.61}$$

From (2.61) we see that $\mathcal{T}_\beta\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha) f = f$. So that the inequality in (2.57) implies:

$$\begin{aligned} &\sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu\left((\lambda+\mu) I - L\right)^{-1}\right)^k f \\ &\geqslant \inf \{g \geqslant f\mathbf{1}_K : g \geqslant \mu R(\lambda+\mu) g, \, g \in D((L)\}, \end{aligned} \tag{2.62}$$

and consequently $U^1_\lambda(f) \leqslant f^\lambda := \sup_{\mu>0,\, k\in\mathbb{N}} \left(\mu\left((\lambda+\mu) I - L\right)^{-1}\right)^k f$. It follows that $f^\lambda = U^1_\lambda(f)$ provided that f and f^λ both belong to $C_b(E)$. If $f^\lambda \in D(L)$, then $f^\lambda = U^1_\lambda(f)$ and $f^\lambda \geqslant \mu R(\lambda+\mu) f^\lambda$, and consequently $\lambda f^\lambda - L f^\lambda \geqslant 0$. The conclusion $U^2_\lambda(f) = f^\lambda$ is then obvious.

This finishes the proof of Proposition 2.4. □

In the following proposition we see that a multiplicative Borel measure is a point evaluation.

Proposition 2.5. *Let μ be a non-zero Borel measure with the property that $\int fg\,d\mu = \int f\,d\mu \int g\,d\mu$ for all functions f and $g \in C_b(E)$. Then there exists $x \in E$ such that $\int f\,d\mu = f(x)$ for $f \in C_b(E)$.*

Proof. Since $\mu \neq 0$ there exists $f \in C_b(E)$ such that $0 \neq \int f d\mu = \int f \mathbf{1} d\mu = \int f d\mu \int \mathbf{1} d\mu$, and hence $0 \neq \int \mathbf{1} d\mu = \left(\int \mathbf{1} d\mu\right)^2$. Consequently, $\int \mathbf{1} d\mu = 1$. Let f and g be functions in $C_b^+(E)$. Then we have

$$\begin{aligned}\int fgd\,|\mu| &= \sup\left\{\left|\int hd\mu\right| : |h| \leqslant fg,\, h \in C_b(E)\right\} \\ &= \sup\left\{\left|\int h_1h_2d\mu\right| : |h_1| \leqslant f,\, |h_2| \leqslant g,\; h_1,\, h_2 \in C_b(E)\right\} \\ &= \sup\left\{\left|\int h_1d\mu\right| : |h_1| \leqslant f,\, h_1 \in C_b(E)\right\} \\ &\qquad \times \sup\left\{\left|\int h_2d\mu\right| : |h_2| \leqslant g,\, h_2 \in C_b(E)\right\} \\ &= \int fd\,|\mu| \int gd\,|\mu|\,. \end{aligned} \tag{2.63}$$

From (2.63) it follows that the variation measure $|\mu|$ is multiplicative as well. Since E is a Polish space, the measure $|\mu|$ is inner-regular. So there exists a compact subset K of E such that $|\mu|\,(E\backslash K) \leqslant 1/2$, and hence $|\mu|\,(K) > 1/2$. Since $|\mu|$ is multiplicative it follows that $|\mu|\,(K) = 1 = |\mu|\,(E)$. It follows that the multiplicative measure $|\mu|$ is concentrated on the compact subset K, and hence it can be considered as a multiplicative measure on $C(K)$. But then there exists a point $x \in K$ such that $|\mu| = \delta_x$, the Dirac measure at x. So there exists a constant c_x such that $\mu = c_x\,|\mu| = c_x\delta_x$. Since $\mu(E) = \delta_x(E) = 1$ it follows that $c_x = 1$. This proves Proposition 2.5. □

2.2 Strong Markov processes and Feller evolutions

In the sequel E denotes a separable complete metrizable topological Hausdorff space. In other words E is a Polish space. The space $C_b(E)$ is the space of all complex valued bounded continuous functions. The space $C_b(E)$ is not only equipped with the uniform norm: $\|f\|_\infty := \sup_{x\in E}|f(x)|$, $f \in C_b(E)$, but also with the strict topology $\mathcal{T}_\beta$. It is considered as a subspace of the bounded Borel measurable functions $L^\infty(E)$, also endowed with the supremum norm.

Definition 2.4. A family $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ of operators defined on $L^\infty(E)$ is called a *Feller evolution* or a *Feller propagator* on $C_b(E)$ if it possesses the following properties:

(i) It leaves $C_b(E)$ invariant: $P(s,t)C_b(E) \subseteq C_b(E)$ for $0 \leqslant s \leqslant t \leqslant T$;

(ii) It is an evolution: $P(\tau,t) = P(\tau,s) \circ P(s,t)$ for all τ, s, t for which $0 \leqslant \tau \leqslant s \leqslant t$ and $P(t,t) = I$, $t \in [0,T]$;

(iii) It consists of contraction operators: $\|P(s,t)f\|_\infty \leqslant \|f\|_\infty$ for all $t \geqslant 0$ and for all $f \in C_b(E)$;

(iv) It is positivity preserving: $f \geqslant 0$, $f \in C_b(E)$, implies $P(s,t)f \geqslant 0$;

(v) For every $f \in C_b(E)$ the function $(s,t,x) \mapsto P(s,t)f(x)$ is continuous on the diagonal of the set $\{(s,t,x) \in [0,T] \times [0,T] \times E : 0 \leqslant s \leqslant t \leqslant T\}$ in the sense that for every element $(t,x) \in (0,T] \times E$ the equality $\lim_{s\uparrow t, y\to x} P(s,t)f(y) = f(x)$ holds, and for every element $(s,x) \in [0,T) \times E$ the equality $\lim_{t\downarrow s, y\to x} P(s,t)f(y) = f(x)$ holds.

(vi) For every $t \in [0,T]$ and $f \in C_b(E)$ the function $(s,x) \mapsto P(s,t)f(x)$ is Borel measurable and if $(s_n,x_n)_{n\in\mathbb{N}}$ is any sequence in $[0,t] \times E$ such that s_n decreases to $s \in [0,t]$, x_n converges to $x \in E$, and $\lim_{n\to\infty} P(s_n,t)\, g(x_n)$ exists in $\mathbb{C}$ for all $g \in C_b(E)$, then $\lim_{n\to\infty} P(s_n,t)\, f(x_n) = P(s,t)f(x)$.

(vii) For every $(t,x) \in (0,T] \times E$ and $f \in C_b(E)$ the following equality holds: $\lim_{s\uparrow t,\, s\geqslant\tau} P(\tau,s)\, f(x) = P(\tau,t)\, f(x)$, $\tau \in [0,t)$.

Remark 2.5. Since the space E is Polish, the continuity as described in (v) can also be described by sequences. So (v) is equivalent to the following condition: for all elements $(t,x) \in (0,T] \times E$ and $(s,x) \in [0,T) \times E$ the equalities

$$\lim_{n\to\infty} P(s_n,t)\, f(y_n) = f(x) \quad \text{and} \quad \lim_{n\to\infty} P(s,t_n)\, f(y_n) = f(x) \tag{2.64}$$

hold. Here $(s_n)_{n\in\mathbb{N}} \subset [0,t]$ is any sequence which increases to t, $(t_n)_{n\in\mathbb{N}} \subset [s,T]$ is any sequence which decreases to s, and $(y_n)_{n\in\mathbb{N}}$ is any sequence in E which converges to $x \in E$. If for all $f \in C_b(E)$ and $t \in [0,T]$ the function $(s,x) \mapsto P(s,t)f(x)$, $(s,x) \in [0,t] \times E$, is continuous, then (vi) and (vii) are satisfied. If the function $(s,t,x) \mapsto P(s,t)\, f(x)$ is continuous on the space $\{(s,t,x) \in [0,T] \times [0,T] \times E : s \leqslant t\}$, then the propagator $P(s,t)$ possesses properties (v) through (vii). In Proposition 2.6 we will single out a closely related property. Its proof is part of the proof of Theorem 2.10.

Definition 2.5. Let the family $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ of operators defined on $L^\infty(E)$ be a *Feller evolution* or a *Feller propagator*. It is called a strong Feller evolution if for every Borel measurable function $f \in L^\infty(E)$, the function $(\tau,t,x) \mapsto P(\tau,t)\, f(x)$, $0 \leqslant \tau < t \leqslant T$, $x \in E$, is continuous.

Proposition 2.6. *Let the family $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ possess the properties (i) through (iv) of Definition 2.4. Suppose that for every $f \in C_b(E)$ the function $(\tau,t,x) \mapsto P(\tau,t)f(x)$ is continuous on the space*

$$\{(\tau,t,x) \in [0,T] \times [0,T] \times E : \tau \leqslant t\}. \tag{2.65}$$

Then for every $f \in C_b([0,T] \times E)$ the function $(\tau,t,x) \mapsto P(\tau,t)f(t,\cdot)(x)$ is continuous on the space in (2.65).

It is noticed that assertions (iii) and (iv) together are equivalent to

(iii′) If $0 \leqslant f \leqslant 1$, $f \in C_b(E)$, then $0 \leqslant P(s,t)f \leqslant \mathbf{1}$, for $0 \leqslant s \leqslant t \leqslant T$.

In the presence of (iii), (ii) and (i), property (v) is equivalent to:

(v′) $\lim_{t\downarrow s} \|u(P(s,t)f - f)\|_\infty = 0$ and $\lim_{s\uparrow t} \|u(P(s,t)f - f)\|_\infty = 0$ for all $f \in C_b(E)$ and $u \in H(E)$. So that a Feller evolution is in fact $\mathcal{T}_\beta$-*strongly continuous* in the sense that, for every $f \in C_b(E)$ and $u \in H(E)$,

$$\lim_{\substack{(s,t)\to(s_0,t_0)\\ s\leqslant s_0\leqslant t_0\leqslant t}} \|u(P(s,t)f - P(s_0,t_0)f)\|_\infty = 0, \quad 0 \leqslant s_0 \leqslant t_0 \leqslant T. \tag{2.66}$$

Remark 2.6. Property (vi) is satisfied if for every $t \in (0,T]$ the function $(s,x) \mapsto P(s,x;t,E) = P(s,t)\mathbf{1}(x)$ is continuous on $[0,t] \times E$, and if for every sequence $(s_n,x_n)_{n\in\mathbb{N}} \subset [0,t] \times E$ for which s_n decreases to s and x_n converges to x, the inequality $\limsup_{n\to\infty} P(s_n,t)f(x_n) \geqslant P(s,t)f(x)$ holds for all $f \in C_b^+(E)$. Since functions of the form $x \mapsto P(s,t)f(x)$, $f \in C_b(E)$, belong to $C_b(E)$, it is also satisfied provided that for every $f \in C_b(E)$ we have

$$\lim_{n\to\infty} P(s_n,t)f = P(s,t)f, \quad \text{uniformly on compact subsets of } E.$$

This follows from the inequality:

$$\begin{aligned} &|P(s_n,t)f(x_n) - P(s,t)f(x)| \\ &\leqslant |P(s_n,t)f(x_n) - P(s,t)(x_n)| + |P(s,t)f(x_n) - P(s,t)f(x)| \end{aligned}$$

where $s_n \downarrow s$, $x_n \to x$ as $n \to \infty$, and $f \in C_b(E)$.

Proposition 2.7. *Let $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ be a family of operators having property (i) and (ii) of Definition 2.4. Then property (iii′) is equivalent to the properties (iii) and (iv) together.*

Moreover, if such a family $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ possesses property (i), (ii) and (iii), then it possesses property (v) if and only if it possesses (v′).

Proof. First suppose that the operator $P(s,t) : L^\infty(E) \to L^\infty(E)$ has the properties (iii) and (iv), and let $f \in C_b(E)$ be such that $0 \leqslant f \leqslant \mathbf{1}$. Then by (iii) and (iv) we have $0 \leqslant P(s,t)f(x) \leqslant \sup_{y\in E} f(y) \leqslant 1$, and hence (iii′) is satisfied. Conversely, let $f \in C_b(E)$ and $x \in E$. Then by (iii′) the operator $P(s,t)$ satisfies

$$\Re P(s,t)f(x) = \left[P(s,t)\Re f\right](x) \leqslant \sup_{y\in E} \Re f(y) \leqslant \left\|\Re f\right\|_\infty . \tag{2.67}$$

There exists $\vartheta \in [-\pi,\pi]$ such that by (2.67) we have

$$\begin{aligned}|P(s,t)f(x)| &= \Re\left[e^{i\vartheta}P(s,t)f\right](x)\\ &= \left[P(s,t)\Re\left(e^{i\vartheta}f\right)\right](x) \leqslant \left\|\Re\left(e^{i\vartheta}f\right)\right\|_\infty \leqslant \|f\|_\infty ,\end{aligned}$$

from which (iii) easily follows. Property (iv) easily follows from (iii′).

Next, suppose that the family $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ possesses property (v′). Then, by taking $s_0 = t_0$, it clearly has property (v). Fix $(s_0,t_0) \in [0,T] \times [0,T]$ in such a way that $s_0 \leqslant t_0$. For the converse implication we employ Theorem 2.7 with the families of operators

$$\{P(s_m,s_0) : 0 \leqslant s_m \leqslant s_{m+1} \leqslant s_0\} \text{ and } \{P(t_0,t_m) : t_0 \leqslant t_{m+1} \leqslant t_m \leqslant T\} \tag{2.68}$$

respectively. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence functions in $C_b^+(E)$ which decreases pointwise to zero. Then by Dini's lemma and assumption (v) we know that

$$\lim_{n\to\infty} \sup_{m\in\mathbb{N}} \sup_{x\in K} P(s_m,s_0) f_n(x) = \lim_{n\to\infty} \sup_{m\in\mathbb{N}} \sup_{x\in K} P(t_0,t_m) f_n(x) = 0 \tag{2.69}$$

for all compact subsets K of E. From (2.69) we see that the sequences of operators in (2.68) are tight. By Theorem 2.7 it follows that they are equi-continuous. If the pair (s,t) belongs to $[0,s_0] \times [t_0,T]$, then we write

$$P(s,t) f - P(s_0,t_0) f = P(s,t_0)\left(P(t_0,t) - I\right) f + \left(P(s,s_0) - I\right) P(s_0,t_0) f. \tag{2.70}$$

Let u be a function in $H(E)$. Since the first sequence in (2.68) is equi-continuous and by invoking (2.70) there exists a function $v \in H(E)$ such that the following inequality holds for all $m \in \mathbb{N}$ and all $f \in C_b(E)$:

$$\begin{aligned}&\left\|u\left(P(s_m,t_m) f - P(s_0,t_0) f\right)\right\|_\infty\\ &\leqslant \left\|v\left(P(t_0,t_m) - I\right) f\right\|_\infty + \left\|u\left(P(s_m,s_0) - I\right) P(s_0,t_0) f\right\| .\end{aligned} \tag{2.71}$$

In order to prove the equality in (2.66) it suffices to show that the right-hand side of (2.71) tends to zero if $m \to \infty$. By the properties of the functions u and v it suffices to prove that

$$\lim_{m\to\infty} \left\|\mathbf{1}_K\left(P(s_m,s_0) f - f\right)\right\|_\infty = \lim_{m\to\infty} \left\|\mathbf{1}_K\left(P(t_0,t_m) f - f\right)\right\|_\infty = 0 \tag{2.72}$$

for every compact subset K of E and for every function $f \in C_b(E)$. The equalities in (2.72) follow from the sequential compactness of K and (v) which imply that

$$\lim_{m\to\infty} P\left(s_m, s_0\right) f\left(x_m\right) = f\left(x_0\right) = \lim_{m\to\infty} P\left(t_0, t_m\right) f\left(x_m\right)$$

whenever s_m increases to s_0, t_m decreases to t_0 and x_m converges to x_0.

This completes the proof of Proposition 2.7. □

2.2.1 *The operators* $\vee_t$*,* $\wedge_t$ *and* ϑ_t

Before we introduce the definition of time-inhomogeneous Markov process we introduce the operators $\vee_t$ and $\wedge_t$, and ϑ_t relative to a stochastic process $s \mapsto X(s) \in E$, $s \in [0,T]$. These operators are called respectively maximum time operator, minimum time operator, and time translation operator. Let $Y : s \mapsto (s, X(s))$ be the corresponding space-time process. These are operators from the sample-path space $[0,T]\times\Omega$ to itself. Their defining property is given by $Y \circ \vee_t(s) = (s \vee t, X(s \vee t))$, $Y \circ \wedge_t(s) = (s \wedge t, X(s \wedge t))$, and $Y \circ \vartheta_t(s) = ((s+t) \wedge T, X((s+t) \wedge T))$, s, $t \in [0,T]$. This is perhaps the right place to explain the compositions $F \circ \vee_t$, $F \circ \wedge_t$, and $F \circ \vartheta_t$, if $F : \Omega \to \mathbb{C}$ is $\mathcal{F}_T^0$-measurable, and if $t \in [0,T]$. Such functions F are called random variable. If F is of the form $F = \prod_{j=1}^n f_j\left(t_j, X\left(t_j\right)\right)$, where the functions f_j, $1 \leqslant j \leqslant n$, are bounded Borel functions, defined on $[0,T]\times E$, then, by definition,

$$F \circ \vee_t = \prod_{j=1}^n f_j\left(t_j \vee t, X\left(t_j \vee t\right)\right), \quad F \circ \wedge_t = \prod_{j=1}^n f_j\left(t_j \wedge t, X\left(t_j \wedge t\right)\right),$$

and

$$F \circ \vartheta_t = \prod_{j=1}^n f_j\left(\left(t_j + t\right) \wedge T, X\left(\left(t_j + t\right) \wedge T\right)\right). \tag{2.73}$$

If t in (2.73) is an $\left(\mathcal{F}_s^0\right)_{s\in[0,T]}$-stopping time, then a similar definition is applied. By the Monotone Class Theorem, the definitions in (2.73) extend to all $\mathcal{F}_T^0$ measurable variables F, i.e. to all random variables. For a discussion on the Monotone Class Theorem see Subsection 2.4.2.

Definition 2.6. Let for every $(\tau, x) \in [0,T] \times E$, a probability measure $\mathbb{P}_{\tau,x}$ on $\mathcal{F}_T^\tau$ be given. Suppose that for every bounded random variable $Y : \Omega \to \mathbb{R}$ the equality

$$\mathbb{E}_{\tau,x}\left(Y \circ \vee_t \,\middle|\, \mathcal{F}_t^\tau\right) = \mathbb{E}_{t,X(t)}\left[Y \circ \vee_t\right]$$

holds $P_{\tau,x}$-almost surely for all $(\tau, x) \in [0,T] \times E$ and for all $t \in [\tau, T]$. Then the process

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), (E, \mathcal{E})\right\} \tag{2.74}$$

is called a *Markov process*. If the fixed time $t \in [\tau, T]$ may be replaced with a stopping time S attaining values in $[\tau, T]$, then the process in (2.74) is called a *strong Markov process*. By definition $\mathbb{P}_{\tau,\triangle}(A) = \mathbf{1}_A(\omega_\triangle) = \delta_{\omega_\triangle}(A)$. Here A belongs to $\mathcal{F}$, and $\omega_\triangle(s) = \triangle$ for all $s \in [0,T]$. If

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), (E, \mathcal{E})\right\}$$

is a Markov process, then we write

$$P(\tau, x; t, B) = \mathbb{P}_{\tau,x}(X(t) \in B), \quad B \in \mathcal{E}, \quad x \in E,\ \tau \leqslant t \leqslant T, \tag{2.75}$$

for the corresponding *transition function*. The operator family (of evolutions, propagators)

$$\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$$

is defined by

$$[P(s,t)f](x) = \mathbb{E}_{s,x}\left[f(X(t))\right] = \int f(y) P\left(s, x; t, dy\right),\ \ f \in C_b(E),\ s \leqslant t \leqslant T.$$

Let $S : \Omega \to [\tau, T]$ be an $(\mathcal{F}_t^\tau)_{t\in[\tau,T]}$-stopping time. Then the σ-field $\mathcal{F}_S^\tau$ is defined by

$$\mathcal{F}_S^\tau = \bigcap_{t\in[\tau,T]} \left\{A \in \mathcal{F}_T^\tau : A \bigcap \{S \leqslant t\} \in \mathcal{F}_t^\tau\right\}.$$

Of course, a random variable $S : \Omega \to [\tau, T]$ is called an $(\mathcal{F}_t^\tau)_{t\in[\tau,T]}$-stopping time, provided that for every $t \in [\tau, T]$ the event $\{S \leqslant t\}$ belongs to $\mathcal{F}_t^\tau$.

2.2.2 *Generators of Markov processes and maximum principles*

We begin with the definition of of the generator of a time-dependent Feller evolution.

Definition 2.7. A family of operators $L(t)$, $0 \leqslant t \leqslant T$, is said to be the (infinitesimal) *generator* of a Feller evolution $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$, if $L(s)f = \mathcal{T}_\beta\text{-}\lim\limits_{t\downarrow s} \dfrac{P(s,t)f - f}{t-s}$, $0 \leqslant s \leqslant T$. This means that a function f belongs to $D(L(s))$ whenever $L(s)f := \lim\limits_{t\downarrow s} \dfrac{P(s,t)f - f}{t-s}$ exists in $C_b(E)$, equipped with the strict topology. It is the same as saying that the function $L(s)f$ belongs to $C_b(E)$, that the family of functions $\left\{\frac{P(s,t)f-f}{t-s} : t \in (s,T)\right\}$ is uniformly bounded and that convergence takes place uniformly on compact subsets of E.

Such a family of operators is considered as an operator L with domain in the space $C_b([0,T]\times E)$. A function $f \in C_b([0,T]\times E)$ is said to belong to $D(L)$ if for every $s \in [0,T]$ the function $x \mapsto f(s,x)$ is a member of $D(L(s))$ and if the function $(s,x) \mapsto L(s)f(s,\cdot)(x)$ belongs to $C_b([0,T]\times E)$. Instead of $L(s)f(s,\cdot)(x)$ we often write $L(s)f(s,x)$. If a function $f \in D(L)$ is such that the function $s \mapsto f(s,x)$ is continuously differentiable, then we say that f belongs to $D^{(1)}(L)$. The time derivative operator $\frac{\partial}{\partial s}$ is often written as D_1. Its domain is denoted by $D(D_1)$, and hence $D^{(1)}(L) = D(D_1)\bigcap D(L)$.

Definition 2.8. The family of operators $L(s)$, $0 \leqslant s \leqslant T$, is said to generate a time-inhomogeneous Markov process

$$\left\{\left(\Omega, \mathcal{F}_T^{\tau}, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant \tau\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), (E, \mathcal{E})\right\} \tag{2.76}$$

if for all functions $u \in D(L)$, for all $x \in E$, and for all pairs (τ, s) with $0 \leqslant \tau \leqslant s \leqslant T$ the following equality holds:

$$\frac{d}{ds}\mathbb{E}_{\tau,x}\left[u(s,X(s))\right] = \mathbb{E}_{\tau,x}\left[\frac{\partial u}{\partial s}(s,X(s)) + L(s)u(s,\cdot)(X(s))\right]. \tag{2.77}$$

Here it is assumed that the derivatives are interpreted as limits from the right which converge uniformly on compact subsets of E, and that the differential quotients are uniformly bounded.

So these derivatives are $\mathcal{T}_\beta$-derivatives.

Definition 2.9. By definition the Skorohod space $D([0,T],E)$ consists of all functions from $[0,T]$ to E which possess left limits in E and are right-continuous. The Skorohod space $D\left([0,T],E^{\triangle}\right)$ consists of all functions from $[0,T]$ to $E^{\triangle}$ which possess left limits in $E^{\triangle}$ and are right-continuous. More precisely, a path (or function) $\omega : [0,T] \to E^{\triangle}$ belongs to $D\left([0,T],E^{\triangle}\right)$ if it possesses the following properties:

(a) if $\omega(t) \in E$, and $s \in [0,t]$, then there exists $\varepsilon > 0$ such that $X(\rho) \in E$ for $\rho \in [0, t+\varepsilon]$, and $\omega(s) = \lim_{\rho\downarrow s}\omega(\rho)$ and $\omega(s-) := \lim_{\rho\uparrow s}\omega(\rho)$ belong to E.

(b) if $\omega(t) = \triangle$ and $s \in [t, T]$, then $\omega(s) = \triangle$. In other words $\triangle$ is an absorbing state.

Observe that the range of $\omega \in D([0, T], E)$ is contained in a totally bounded subset of E. Such sets are relatively compact. Also observe that the range of $\omega \in D\left([0, T], E^{\triangle}\right)$ restricted to an interval of the form $[0, t]$ is also totally bounded provided that $\omega(t) \in E$. It follows that paths $\omega \in D\left([0, T], E^{\triangle}\right)$ restricted to intervals of the form $[0, t]$ have relatively compact range as long as they have not reached the absorption state $\triangle$, i.e. as long as $\omega(t) \in E$.

Theorems 2.9, 2.10, 2.11, and 2.12 in §2.3 are well known in case E is a locally compact second countable Hausdorff space. In fact the sample space Ω should depend on τ. This is taken care of by assuming that the measure $\mathbb{P}_{\tau,x}$ is defined on the σ-field $\mathcal{F}_T^{\tau}$.

Let L be a linear operator with domain $D(L)$ and range $R(L)$ in $C_b(E)$. The following definition should be compared with Definition 4.1, and with assertion (b) in Proposition 4.3 in Chapter 4.

Definition 2.10. Let E_0 be subset of E. The operator L satisfies the maximum principle on E_0, provided

$$\sup_{x \in E_0} \Re\left(\lambda f(x) - Lf(x)\right) \geqslant \lambda \sup_{x \in E_0} \Re f(x), \text{ for all } \lambda > 0, \text{ and for all } f \in D(L). \tag{2.78}$$

If L satisfies (2.78) on $E_0 = E$, then the operator L satisfies the maximum principle of Definition 4.1.

The next definition is the same as the one in Definition 4.5 in Chapter 4.

Definition 2.11. Let E_0 be a subset of E. Suppose that the operator L has the property that for every $\lambda > 0$ and for every $x_0 \in E_0$ the number $\Re h(x_0) \geqslant 0$, whenever $h \in D(L)$ is such that $\Re(\lambda I - L)h \geqslant 0$ on E_0. Then the operator L is said to satisfy the weak maximum principle on E_0.

The following proposition says that the concepts in the definitions 2.10 and 2.11 coincide, provided $\mathbf{1} \in D(L)$ and $L\mathbf{1} = 0$.

Proposition 2.8. *If the operator L satisfies the maximum principle on E_0, then L satisfies the weak maximum principle on E_0. Suppose that the constant functions belong to $D(L)$, and that $L\mathbf{1} = 0$. If L satisfies the weak maximum principle on E_0, then it satisfies the maximum principle on E_0.*

Proof. First we observe that (2.78) is equivalent to

$$\inf_{x \in E_0} \Re\left(\lambda f(x) - Lf(x)\right) \leqslant \lambda \inf_{x \in E_0} \Re f(x), \text{ for all } \lambda > 0, \text{ and for all } f \in D(L). \tag{2.79}$$

Hence, if $\lambda f - Lf \geqslant 0$ on E_0, then (2.79) implies that $\Re f(x_0) \geqslant 0$ for all $x_0 \in E_0$.

Conversely, suppose that $\mathbf{1} \in D(L)$ and that $L\mathbf{1} = 0$. Let $f \in D(L)$, put $m = \inf\{\Re f(y) : y \in E_0\}$, and assume that

$$\inf_{x\in E_0} \Re(\lambda f - Lf)(x) > \lambda \inf_{y\in E_0} \Re f(y) = \lambda m. \tag{2.80}$$

Then there exists $\varepsilon > 0$ such that $\inf_{x\in E_0} \Re(\lambda f - Lf)(x) \geqslant \lambda(m+\varepsilon)$. Hence, since $L\mathbf{1} = 0$, $\inf_{x\in E} \Re(\lambda I - L)(f - m - \varepsilon)(x) \geqslant 0$. Since the operator L satisfies the weak maximum principle, we see $\Re(f - m - \varepsilon) \geqslant 0$ on E_0. Since this is equivalent to $\Re f \geqslant m+\varepsilon$ on E_0, which contradicts the definition of m. Hence, our assumption in (2.80) is false, and consequently,

$$\inf_{x\in E_0} \Re(\lambda f - Lf)(x) \leqslant \lambda \inf_{y\in E_0} \Re f(y). \tag{2.81}$$

Since (2.81) is equivalent to (2.78) this concludes the proof of Proposition 2.8. □

Definition 2.12. Let an operator L, with domain and range in $C_b(E)$, satisfy the maximum principle. Then L is said to possess the *global Korovkin* property, if there exists $\lambda_0 > 0$ such that fore every $x_0 \in E$, the subspace $S(\lambda_0, x_0)$, defined by

$$\begin{aligned} S(\lambda_0, x_0) = \Big\{ g \in C_b(E) : \quad & \text{for every } \varepsilon > 0 \text{ the inequality} \\ & \sup\{h_1(x_0) : (\lambda_0 I - L)h_1 \leqslant \Re\, g + \varepsilon,\ h_1 \in D(L)\} \\ & \geqslant \inf\{h_2(x_0) : (\lambda_0 I - L)h_2 \geqslant \Re\, g - \varepsilon,\ h_2 \in D(L)\} \text{ is valid} \Big\}, \end{aligned} \tag{2.82}$$

coincides with $C_b(E)$.

Remark 2.7. Let D be a subspace of $C_b(E)$ with the property that for every $x_0 \in E$ the space $S(x_0)$, defined by

$$\begin{aligned} S(x_0) = \{ g \in C_b(E) : \quad & \text{for every } \varepsilon > 0 \text{ the inequality} \\ & \sup\{h_1(x_0) : h_1 \leqslant \Re\, g + \varepsilon, \quad h_1 \in D\} \\ & \geqslant \inf\{h_2(x_0) : h_2 \geqslant \Re\, g - \varepsilon, \quad h_2 \in D\} \quad \text{holds}\}, \end{aligned} \tag{2.83}$$

coincides with $C_b(E)$. Then such a subspace D could be called a *global Korovkin* subspace of $C_b(E)$. In fact the inequality in (2.83) is pretty much the same as the one in (2.82) in case $L = 0$.

Any countable union of compact subsets of E is called σ-compact subset. In what follows the symbol $\mathcal{K}_\sigma(E)$ denotes the collection of σ-compact subsets of E. In practical situations the set E_0 in the following definition is a member of $\mathcal{K}_\sigma(E)$ or a Polish (for instance an open) subset of E.

Definition 2.13. Let E_0 be subset of E. Let an operator L, with domain and range in $C_b(E)$, satisfy the maximum principle on E_0. Then L is said to possess the Korovkin property on E_0, if there exists $\lambda_0 > 0$ such that for every $x_0 \in K$, the subspace $S_{\text{loc}}(\lambda_0, x_0, E_0)$, defined by

$$S_{\text{loc}}(\lambda_0, x_0, E_0) = \Big\{ g \in C_b(E) : \quad \text{for every } \varepsilon > 0 \text{ the inequality} \qquad (2.84)$$
$$\sup_{h_1 \in D(L)} \{ h_1(x_0) : (\lambda_0 I - L)\, h_1 \leqslant \Re\, g + \varepsilon, \text{ on } E_0 \}$$
$$\geqslant \inf_{h_2 \in D(L)} \{ h_2(x_0) : (\lambda_0 I - L)\, h_2 \geqslant \Re\, g - \varepsilon, \text{ on } E_0 \} \Big\},$$

coincides with $C_b(E)$.

2.3 Strong Markov processes: Main results

The following theorems 2.9 through 2.13 contain the basic results about strong Markov processes on Polish spaces, their sample paths, and their generators. Theorem 2.9 says that a Feller evolution (or propagator) can be considered as the one-dimensional distributions, or marginals, of a strong Markov process. Theorem 2.10 describes the reverse situation: with certain Markov processes we may associate Feller propagators. In Theorem 2.11 the intimate link between unique solutions to the martingale problem and the strong Markov property is established. Theorem 2.12 contains a converse result: Markov processes can be considered as solutions to the martingale problem. Finally, in Theorem 2.13 operators which possess unique linear extensions which generate Feller evolutions, and for which the martingale problem is uniquely solvable, are described. For such operators the martingale problem is said to be well-posed. A Hunt process is a strong Markov process which is quasi-left continuous with respect to the minimum completed admissible filtration $\{\mathcal{F}_t^\tau\}_{\tau \leqslant t \leqslant T}$: see item (4) in Theorem 2.9 and Definition 2.15. For Theorem 2.9 in the locally compact setting and a time-homogeneous Feller evolution (i.e. a Feller-Dynkin semigroup) the reader may e.g. consult [Blumenthal and Getoor (1968)]. It will be convenient to insert some definitions before formulating the main results of Part 2 of this book. The following definition should be compared with the definitions given in (3.24), (3.25), (3.26), and (3.27) in §3.1.

Definition 2.14. Let $\{\mathcal{G}_t^\tau : 0 \leqslant \tau \leqslant t \leqslant T\}$ be family of σ-fields. This family is called a double filtration if $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant t \leqslant T$ implies $\mathcal{G}_t^{\tau_2} \subset \mathcal{G}_t^{\tau_1}$, and $\tau \leqslant t_1 \leqslant t_2 \leqslant T$ implies $\mathcal{G}_{t_1}^{\tau} \subset \mathcal{G}_{t_2}^{\tau}$. Unless specified otherwise a family of the form $\{\mathcal{G}_t^{\tau} : 0 \leqslant \tau \leqslant t \leqslant T\}$ always denotes a double filtration. A random variable $S : \Omega \to [\tau, T]$ is called a $(\mathcal{G}_t^{\tau})_{t\in[\tau,T]}$-stopping time whenever $\{S \leqslant t\} \in \mathcal{G}_t^{\tau}$ for all $t \in [\tau, T]$. The corresponding σ-field $\mathcal{G}_S^{\tau}$ is defined by

$$\mathcal{G}_S^{\tau} = \bigcap_{t\in[\tau,T]} \left\{A \in \mathcal{G}_T^{\tau} : A \bigcap [\tau, t] \in \mathcal{G}_t^{\tau}\right\}.$$

The right closure $\mathcal{G}_{t+}^{\tau}$ of the σ-field $\mathcal{G}_t^{\tau}$ is defined by $\mathcal{G}_{t+}^{\tau} = \bigcap_{s\in(t,T]} \mathcal{G}_s^{\tau}$. If $S : \Omega \to [\tau, T]$ is a $(\mathcal{G}_t^{\tau})_{t\in[\tau,T]}$-stopping time, then the σ-field $\mathcal{G}_{S+}^{\tau}$ is defined by

$$\mathcal{G}_{S+}^{\tau} = \bigcap_{t\in(\tau,T]} \left\{A \in \mathcal{G}_T^{\tau} : A \bigcap [\tau, t] \in \mathcal{G}_{t+}^{\tau}\right\}.$$

A random variable $S : \Omega \to [\tau, T]$ is called a terminal $(\mathcal{G}_t^{\tau})_{t\in[\tau,T]}$-stopping time provided that $\{t_1 < S \leqslant t_2\} \in \mathcal{G}_{t_2}^{t_1}$ for all $\tau \leqslant t_1 \leqslant t_2 \leqslant T$. If S_1 and $S_2 : \Omega \to [\tau, T]$ are terminal $(\mathcal{G}_t^{\tau})_{t\in[\tau,T]}$-stopping times such that $S_1 \leqslant S_2$, then the σ-field $\mathcal{G}_{S_2}^{S_1}$ is defined by

$$\begin{aligned}
\mathcal{G}_{S_2}^{S_1} &= \bigcap_{\tau\leqslant\rho<T} \left\{A \in \mathcal{G}_{S_2}^{\tau} : A \bigcap \{S_1 > \rho\} \in \mathcal{G}_{S_2}^{\rho}\right\} \\
&= \bigcap_{\tau\leqslant\rho<T} \bigcap_{\tau<t\leqslant T} \left\{A \in \mathcal{G}_{S_2}^{\tau} : A \bigcap \{S_1 > \rho\} \bigcap \{S_2 \leqslant t\} \in \mathcal{G}_t^{\rho}\right\} \\
&= \bigcap_{\tau<t\leqslant T} \bigcap_{t\leqslant\rho<T} \left\{A \in \mathcal{G}_{S_2}^{\tau} : A \bigcap \{S_1 > \rho\} \bigcap \{S_2 \leqslant t\} \in \mathcal{G}_t^{\rho}\right\} \\
&= \bigcap_{\tau<t\leqslant T} \left\{A \in \mathcal{G}_{S_2}^{\tau} : A \bigcap \{S_2 \leqslant t\} \in \mathcal{G}_t^{S_1\wedge t}\right\}
\end{aligned}$$

where

$$\mathcal{G}_t^{S_1\wedge t} = \bigcap_{\tau\leqslant\rho<t} \left\{A \in \mathcal{G}_t^{\tau} : A \bigcap \{S_1 > \rho\} \in \mathcal{G}_t^{\rho}\right\}.$$

The right closure $\mathcal{G}_{S_2+}^{S_1}$ of the σ-field $\mathcal{G}_{S_2}^{S_1}$ is defined by

$$\mathcal{G}_{S_2+}^{S_1} = \bigcap_{\tau\leqslant\rho<T} \bigcap_{\tau<t\leqslant T} \left\{A \in \mathcal{G}_T^{\tau} : A \bigcap \{S_1 > \rho\} \bigcap \{S_2 \leqslant t\} \in \mathcal{G}_{t+}^{\rho}\right\}. \tag{2.85}$$

Let $\left(\Omega, \mathcal{G}_T^{\tau}, \left(\mathcal{G}_{t_2}^{t_1}\right)_{\tau\leqslant t_1\leqslant t_2\leqslant T}, \mathbb{P}_{\tau,x}\right)$ be a probability space with a double filtration. Fix $\tau \leqslant t_1 < t_2 \leqslant T$. Then the $\mathbb{P}_{\tau,x}$-closure $\overline{\mathcal{G}_{t_2}^{t_1}}^{\mathbb{P}_{\tau,x}}$ of the σ-field $\mathcal{G}_{t_2}^{t_1}$ in $\mathcal{G}_T^{\tau}$ is defined by

$$\begin{aligned}
\overline{\mathcal{G}_{t_2}^{t_1}}^{\mathbb{P}_{\tau,x}} = \big\{A \in \mathcal{G}_T^{\tau} &: \text{there exist } A_1,\, A_2 \in \mathcal{G}_{t_2}^{t_1} \text{ such that} \\
&A_1 \subset A \subset A_2 \text{ and } \mathbb{P}_{\tau,x}[A_2 \backslash A_1] = 0\big\}.
\end{aligned} \tag{2.86}$$

If in (2.86) the σ-field $\mathcal{G}_T^\tau$ is replaced with the power set of Ω, then we obtain the $\mathbb{P}_{\tau,x}$-completion of the σ-field $\mathcal{G}_{t_2}^{t_1}$. Similar conventions are employed for $\mathbb{P}_{\tau,\mu}$-closures and $\mathbb{P}_{\tau,\mu}$-completions; here $\mathbb{P}_{\tau,\mu}[A] = \int_E \mathbb{P}_{\tau,x}[A]\, d\mu(x)$, $A \in \mathcal{G}_T^\tau$. Occasionally we need the following σ-field:

$$\mathcal{G}_{S_2}^{S_1,\vee} = \bigcap_{0 \leqslant \rho \leqslant T} \left\{ A \in \mathcal{G}_T^\tau : \vee_\rho^{-1} A \in \mathcal{G}_{S_2 \vee \rho}^{S_1} \right\}.$$

Here the operator $\vee_\rho : \Omega \to \Omega$ is $\mathcal{G}_{t_2 \vee \rho}^{t_1 \vee \rho}$-$\mathcal{G}_{t_2}^{t_1}$-measurable, $\tau \leqslant t_1 \leqslant t_2 \leqslant T$, $\rho \in [0, T]$. Suppose that $X(s) \circ \vee_\rho = X(s \vee \rho)$ for all $s \in [\tau, T]$. If the σ-fields $\mathcal{G}_{t_2}^{t_1}$ are generated by the state variables $(X(s) : t_1 \leqslant s \leqslant t_2)$, $\tau \leqslant t_1 \leqslant t_2 \leqslant T$, then the maximum operators $\vee_\rho$, $\tau \leqslant \rho \leqslant T$, possess such measurability properties.

In subsection 2.2.1 the reader finds some information on the operators $\vee_t$, $\wedge_t$ and ϑ_t, $t \geqslant 0$. Notice that in the definition of $\mathcal{G}_{S_2}^{S_1}$ we need the fact that the stopping times S_1 and S_2 are terminal and satisfy $\tau \leqslant S_1 \leqslant S_2 \leqslant T$, because we want to be sure that events of the form $\{S_2 \leqslant t\}$ belong to this σ-field. Such an event belongs to $\mathcal{G}_{S_2}^{S_1}$ provided that for every ρ, $\rho' \in [\tau, t]$, $\rho' \leqslant \rho$, the event

$$\{S_2 \leqslant t\} \bigcap \{S_2 \leqslant \rho_2\} \bigcap \{S_1 > \rho'\} = \{\rho' < S_1 \leqslant t \wedge \rho\} \bigcap \{\rho' < S_2 \leqslant t \wedge \rho\}$$

belongs to $\mathcal{G}_\rho^{\rho'}$. The latter follows from the inequality $S_1 \leqslant S_2$ together with the assumption that the stopping times S_1 and S_2 are terminal. Also note that the right closure of $\mathcal{G}_{S_2+}^{S_1}$ is given by

$$\mathcal{G}_{S_2+}^{S_1} = \bigcap_{h>0} \mathcal{G}_{(S_2+h)\wedge T}^{S_1}. \tag{2.87}$$

The notion of strong Markov process relative to the minimal double filtration $\{\mathcal{F}_t^\tau : 0 \leqslant \tau \leqslant t \leqslant T\}$ is explained in Definition 2.6. The same definition can be used if a more general double filtration $\{\mathcal{F}_t^\tau : 0 \leqslant \tau \leqslant t \leqslant T\}$ is employed. In the following definition we collect some notions related to continuity of our Markov process.

Definition 2.15. Let

$$\left\{ \left(\Omega, \mathcal{G}_T^\tau, (\mathcal{G}_t^\tau)_{t \in [\tau,T]}, \mathbb{P}_{\tau,x}\right), (X(t), \tau \leqslant t \leqslant T), (\vee_t : \tau \leqslant t \leqslant T), (E, \mathcal{E}) \right\}, \tag{2.88}$$

be a Markov process. It is called normal if $\mathbb{P}_{\tau,x}[X(\tau) = x] = 1$ for all $(\tau, x) \in [0, T] \times E$. It is called right-continuous if $\lim_{t \downarrow s} X(t) = X(s)$, $\mathbb{P}_{\tau,x}$-almost surely for $\tau \leqslant s \leqslant T$, possesses left limits in E on its life time (i.e. $\lim_{t\uparrow s} X(t)$ exists in E, whenever $\zeta > s$), and is quasi-left continuous (i.e. if $(\tau_n : n \in \mathbb{N})$ is an increasing sequence of $\left(\mathcal{F}^\tau_{t+}\right)$-stopping times, $X(\tau_n)$ converges $\mathbb{P}_{\tau,x}$-almost surely to $X(\tau_\infty)$ on the event $\{\tau_\infty < \zeta\}$, where $\tau_\infty = \sup_{n\in\mathbb{N}} \tau_n$). Here ζ is the life time of the process $t \mapsto X(t)$: $\zeta = \inf\{s > 0 : X(s) = \triangle\}$, when $X(s) = \triangle$ for some $s \in [0,T]$, and elsewhere $\zeta = T$. If in (2.88) the minimal σ-fields $\mathcal{F}^\tau_t = \sigma(X(\rho) : \tau \leqslant \rho \leqslant t)$ are taken instead of $\mathcal{G}^\tau_t$, and if the Markov process in (2.88) has all these properties, then it is called a Hunt process.

In the following theorem we see that with a Feller evolution

$$\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\} \tag{2.89}$$

a strong Markov process can be associated in such a way that the one-dimensional distributions or marginals are determined by the operators $f \mapsto P(\tau,t) f$, $f \in C_b(E)$. In fact every operator $P(\tau,t)$ can be written as

$$P(\tau,t) f(x) = \int P(\tau,x;t,dy) f(y), \quad f \in C_b(E),$$

where the mapping

$$(\tau,x,t,B) \mapsto P(\tau,x;t,B), \quad (\tau,x,t,B) \in [0,T] \times E \times [0,T] \times \mathcal{E},\ \tau \leqslant t,$$

is a sub-probability transition function.

Definition 2.16. If the Feller evolution in (2.89) is strong Feller, then the corresponding Markov process in (2.90) below is said to have the strong Feller property, or to be strong Feller.

The proof of the following theorem can be found in Chapter 3 subsection 3.1.1.

Theorem 2.9. *Let $\{P(\tau,t) : \tau \leqslant t \leqslant T\}$ be a Feller evolution in $C_b(E)$. Then there exists a strong Markov process (in fact a Hunt process)*

$$\{(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}), (X(t), \tau \leqslant t \leqslant T), (\vee_t : \tau \leqslant t \leqslant T), (E,\mathcal{E})\}, \tag{2.90}$$

such that $[P(\tau,t)f](x) = \mathbb{E}_{\tau,x}[f(X(t))]$, $f \in C_b(E)$, $t \geqslant 0$. Moreover this Markov process possesses the following properties:

(1) it is normal, i.e. $\mathbb{P}_{\tau,x}[X(\tau) = x] = 1$;
(2) it is right continuous, i.e. $\lim_{t\downarrow s} X(t) = X(s)$, $\mathbb{P}_{\tau,x}$-almost surely for $\tau \leqslant s \leqslant T$;
(3) it possesses left limits in E on its life time, i.e. $\lim_{t\uparrow s} X(t)$ exists in E, whenever $\zeta > s$;

(4) *it is quasi-left continuous: i.e. if* $(\tau_n : n \in \mathbb{N})$ *is an increasing sequence of* $\left(\mathcal{F}_{t+}^{\tau}\right)$*-stopping times,* $X(\tau_n)$ *converges* $\mathbb{P}_{\tau,x}$*-almost surely to* $X(\tau_\infty)$ *on the event* $\{\tau_\infty < \zeta\}$*, where* $\tau_\infty = \sup_{n\in\mathbb{N}} \tau_n$*).*

Here ζ *is the life time of the process* $t \mapsto X(t)$*:* $\zeta = \inf\{s > 0 : X(s) = \triangle\}$*, when* $X(s) = \triangle$ *for some* $s \in [0,T]$*, and elsewhere* $\zeta = T$*. Put*

$$\mathcal{F}_{t+}^{\tau} = \bigcap_{s\in(t,T]} \mathcal{F}_{s}^{\tau} = \bigcap_{s\in(t,T]} \sigma\left(X(\rho) : \tau \leqslant \rho \leqslant s\right). \tag{2.91}$$

Let $F : \Omega \to \mathbb{C}$ *be a bounded* $\mathcal{F}_T^s$*-measurable random variable. Then*

$$\mathbb{E}_{s,X(s)}\left[F\right] = \mathbb{E}_{\tau,x}\left[F \mid \mathcal{F}_s^{\tau}\right] = \mathbb{E}_{\tau,x}\left[F \mid \mathcal{F}_{s+}^{\tau}\right] \tag{2.92}$$

$\mathbb{P}_{\tau,x}$*-almost surely for all* $\tau \leqslant s$ *and* $x \in E$*. Consequently, the process defined in (2.90) is in fact a Markov process with respect to the right closed filtrations:* $\left(\mathcal{F}_{t+}^{\tau}\right)_{t\in[\tau,T]}$*,* $\tau \in [0,T]$*. Moreover, the events* $\{X(t) \in E\}$ *and* $\{X(t) \in E,\, \zeta \geqslant t\}$ *coincide* $\mathbb{P}_{\tau,x}$*-almost surely for* $\tau \leqslant t \leqslant T$ *and* $x \in E$*. Even more is true, the process defined in (2.90) is strong Markov with respect to the filtrations* $\left(\mathcal{F}_{t+}^{\tau}\right)_{t\in[\tau,T]}$*,* $\tau \in [0,T]$*, in the sense that*

$$\mathbb{E}_{S,X(S)}\left[F \circ \vee_S\right] = \mathbb{E}_{\tau,x}\left[F \circ \vee_S \mid \mathcal{F}_{S+}^{\tau}\right] \tag{2.93}$$

for all bounded $\mathcal{F}_T^0$*-measurable random variables* $F : \Omega \to \mathbb{C}$ *and for all* $\left(\mathcal{F}_{t+}^{\tau}\right)_{t\in[\tau,T]}$*-stopping times* $S : \Omega \to [\tau,T]$*. The σ-field* $\mathcal{F}_{S+}^{\tau}$ *is defined in Definition 2.14 equality (2.85): see (2.87), and (2.97) in Remark 2.8 as well. As* Ω *the Skorohod space* $D([0,T],E)$*, if* $P(\tau,x : t,E) = 1$ *for all* $0 \leqslant \tau \leqslant t \leqslant T$*, or* $D\left([0,T],E^{\triangle}\right)$*, otherwise, may be chosen.*

The following theorem contains kind of a converse statement to Theorem 2.9. Its proof can be found in Chapter 3 subsection 3.1.2.

Theorem 2.10. *Conversely, let*

$$\left\{\left(\Omega,\mathcal{F}_T^{\tau},\mathbb{P}_{\tau,x}\right),\left(X(t),\tau \leqslant t \leqslant T\right),\left(\vee_t : \tau \leqslant t \leqslant T\right),\left(E,\mathcal{E}\right)\right\} \tag{2.94}$$

be a strong Markov process which is normal, right continuous, and possesses left limits in E *on its life time. Put, for* $x \in E$ *and* $0 \leqslant \tau \leqslant t \leqslant T$*, and* $f \in L^\infty([0,T]\times E,\mathcal{E})$*,*

$$\left[P(\tau,t)f(t,\cdot)\right](x) = \mathbb{E}_{\tau,x}\left[f(t,X(t))\right] = \int P(\tau,x;t,dy)\, f(t,y)\,, \tag{2.95}$$

where $P(\tau,x;t,B) = \mathbb{P}_{\tau,x}\left[X(t) \in B\right]$*,* $B \in \mathcal{E}$*. Suppose that the function* $(s,t,x) \mapsto P(s,t)f(x)$ *is continuous on the set*

$$\{(s,t,x) \in [0,T]\times[0,T]\times E : s \leqslant t\}$$

for all functions f belonging to $C_b(E)$, $0 \leqslant s \leqslant t \leqslant T$. Then the family $\{P(s,t) : T \geqslant t \geqslant s \geqslant 0\}$ is a Feller evolution. Moreover, functions of the form $(s,t,x) \mapsto P(s,t)f(t,\cdot)(x)$, $f \in C_b([0,T] \times E)$, are continuous on the same space. The maximum operators $\vee_t : \Omega \to \Omega$, $t \in [\tau, T]$, have the property that for all $(\tau, x) \in [0,T] \times E$ the equality $X(s) \circ \vee_t = X(s \vee t)$ holds $\mathbb{P}_{\tau,x}$-almost surely for all $t \in [\tau, T]$.

The following theorem shows that for generators of Feller evolutions the martingale problem is uniquely solvable. Its proof is to be found in Chapter 3 subsection 3.1.3.

Theorem 2.11. *Let the family $L = \{L(s) : 0 \leqslant s \leqslant T\}$ be the generator of a Feller evolution in $C_b(E)$ and let the process in (2.90) be the corresponding Markov process. For every $f \in D^{(1)}(L) = D(D_1) \bigcap D(L)$ and for every $(\tau, x) \in [0,T] \times E$, the process*

$$t \mapsto f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t \left(\frac{\partial}{\partial s} + L(s) \right) f(s, X(s))\, ds \tag{2.96}$$

is a $\mathbb{P}_{\tau,x}$-martingale for the filtration $(\mathcal{F}_t^\tau)_{T \geqslant t \geqslant \tau}$, where each σ-field $\mathcal{F}_t^\tau$, $T \geqslant t \geqslant \tau \geqslant 0$, is the $\mathbb{P}_{\tau,x}$-completion of $\sigma(X(u) : \tau \leqslant u \leqslant t)$. In fact the σ-field $\mathcal{F}_t^\tau$ may be taken as the $\mathbb{P}_{\tau,x}$-completion of the right closure $\mathcal{F}_t^\tau = \bigcap_{s>t} \sigma(X(\rho) : \tau \leqslant \rho \leqslant s)$. It is also possible to complete $\mathcal{F}_t^\tau$ with respect to $\mathbb{P}_{\tau,\mu}$, given by $\mathbb{P}_{\tau,\mu}(A) = \int \mathbb{P}_{\tau,x}(A) d\mu(x)$. For $\mathcal{F}_t^\tau$ the following σ-field may be chosen:

$$\mathcal{F}_t^\tau = \bigcap_{\mu \in P(E)} \bigcup_{T \geqslant s > t} \{\mathbb{P}_{\tau,\mu}\text{-completion of } \sigma(X(u) : \tau \leqslant u \leqslant s)\}.$$

The following theorem makes it clear that there is a converse to the statement in Theorem 2.11. For a proof the reader may consult subsection 3.1.4 in Chapter 3.

Theorem 2.12. *Conversely, let $L = \{L(s) : 0 \leqslant s \leqslant T\}$ be a family of $\mathcal{T}_\beta$-densely defined linear operators with domain $D(L(s))$ and range $R(L(s))$ in $C_b(E)$, such that $D^{(1)}(L)$ is $\mathcal{T}_\beta$-dense in $C_b([0,T] \times E)$. Let*

$$((\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}) : (\tau, x) \in [0,T] \times E)$$

be the unique family of probability spaces with state variables $(X(t) : t \in [0,T])$ defined on the filtered space $\left(\Omega, (\mathcal{F}_t^\tau)_{\tau \leqslant t \leqslant T}\right)$ with values in the state space $(E, \mathcal{E})$ possessing the following properties: for all pairs

$0 \leqslant \tau \leqslant t \leqslant T$ the state variable $X(t)$ is $\mathcal{F}_t^\tau$-$\mathcal{E}$-measurable, for all pairs $(\tau, x) \in [0,T] \times E$, $\mathbb{P}_{\tau,x}[X(\tau) = x] = 1$, and for all $f \in D^{(1)}(L)$ the process

$$t \mapsto f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t \left(\frac{\partial}{\partial s} + L(s)\right) f(s, X(s))\, ds$$

is a $\mathbb{P}_{\tau,x}$-martingale with respect to the filtration $(\mathcal{F}_t^\tau)_{\tau \leqslant t \leqslant T}$. Then the family of operators $L = \{L(s) : 0 \leqslant s \leqslant T\}$ possesses a unique extension

$$L_0 = \{L_0(s) : 0 \leqslant s \leqslant T\},$$

which generates a Feller evolution in $C_b(E)$. It is required that the operator $D_1 + L$ is sequentially λ-dominant in the sense of Definition 4.3; i.e. for every sequence of functions $(\psi_m)_{m\in\mathbb{N}} \subset C_b([0,T] \times E)$ which decreases pointwise to zero the sequence $\{\psi_n^\lambda : n \in \mathbb{N}\}$, defined by

$$\psi_n^\lambda = \sup_{K \in \mathcal{K}([0,T]\times E)} \inf \{g \geqslant \psi_n \mathbf{1}_K : g \in D(D_1 + L),\ (\lambda I - D_1 - L) g \geqslant 0\},$$

decreases uniformly on compact subsets of $[0,T] \times E$ to zero as well. In addition, the sample space Ω is supposed to be the Skorohod space $D([0,T], E^\triangle)$; in particular $X(t) \in E$, $\tau \leqslant s < t$, implies $X(s) \in E$.

The following theorem gives Korovkin type conditions in order that a family of operators possesses a unique extension which generates a Feller evolution. For the proof the reader is referred to Chapter 3 subsection 3.1.5.

Theorem 2.13. *(Unique Markov extensions) Suppose that the $\mathcal{T}_\beta$-densely defined linear operator*

$$D_1 + L = \left\{\frac{\partial}{\partial s} + L(s) : 0 \leqslant s \leqslant T\right\},$$

with domain and range in $C_b([0,T] \times E)$, possesses the global Korovkin property and satisfies the maximum principle, as exhibited in Definition 2.10. Also suppose that L assigns real functions to real functions. Then the family $L = \{L(s) : 0 \leqslant s \leqslant T\}$ extends to a unique generator $L_0 = \{L_0(s) : 0 \leqslant s \leqslant T\}$ of a Feller evolution, and the martingale problem is well posed for the family of operators $\{L(s) : 0 \leqslant s \leqslant T\}$. Moreover, the Markov process associated with $\{L_0(s) : 0 \leqslant s \leqslant T\}$ solves the martingale problem uniquely for the family $L = \{L(s) : 0 \leqslant s \leqslant T\}$.

Let E_0' be a subset of E which is Polish for the relative topology. Put $E_0 = [0,T] \times E_0'$. The same conclusion is true with E_0' instead of E if the operator $D_1 + L$ possesses the following properties:

(1) If $f \in D^{(1)}(L)$ vanishes on E_0, then $D_1 f + Lf$ vanishes on E_0 as well.

(2) The operator $D_1 + L$ satisfies the maximum principle on E_0.
(3) The operator $D_1 + L$ is positive $\mathcal{T}_\beta$-dissipative on E_0.
(4) The operator $D_1 + L$ is sequentially λ-dominant on E_0 for some $\lambda > 0$.
(5) The operator $D_1 + L$ has the Korovkin property on E_0.

The notion of maximum principle on E_0 is explained in Definitions 2.11 and 2.10: see Proposition 2.8 as well. The concept of Korovkin property on a subset E_0 can be found in Definition 2.13 with $D_1 + L$ instead of L. Let $(D_1 + L)\restriction_{E_0}$ be the operator defined by $D\left((D_1 + L)\restriction_{E_0}\right) = \left\{f\restriction_{E_0} : f \in D^{(1)}(L)\right\}$, and $(D_1 + L)\restriction_{E_0}(f\restriction_{E_0}) = D_1 f + Lf\restriction_{E_0}$, $f \in D(L)$. Then the operator $L\restriction_{E_0}$ possesses a unique linear extension to the generator L_0 of a Feller semigroup on $C_b(E_0)$.

For the notion of $\mathcal{T}_\beta$-dissipativity the reader is referred to inequality (4.14) in Definition 4.2, and for the notion of sequentially λ-dominant operator see Definition 4.3. In Proposition 2.3, and in (4.16) of Definition 4.3 the function ψ_n^λ in Theorem 2.12 is denoted by $U_\lambda^1(\psi_n)$. The sequential λ-dominance will guarantee that the semigroup which can be constructed starting from the other hypotheses in Theorems 2.12 and 2.13 is a Feller semigroup indeed: see Theorem 4.3.

Remark 2.8. Notice that in (2.93) we cannot necessarily write

$$\mathbb{E}_{S,X(S)}\left[F \circ \vee_S\right] = \mathbb{E}_{\tau,x}\left[F \circ \vee_S \,\middle|\, \mathcal{F}_S^\tau\right],$$

because events of the form $\{S \leqslant t\}$ may not be $\mathcal{F}_t^\tau$-measurable, and hence the σ-field $\mathcal{F}_S^\tau$ is not well-defined. In (2.93) the σ-field $\mathcal{F}_{S+}^\tau$ is defined by

$$\mathcal{F}_{S+}^\tau = \bigcap_{t \geqslant 0} \left\{A \in \mathcal{F}_T^\tau : A \bigcap [S \leqslant t] \in \mathcal{F}_{t+}^\tau\right\}. \tag{2.97}$$

Remark 2.9. Let $d : E \times E \to [0,1]$ be a metric on E which turns E into a complete metrizable space, and let $\triangle$ be an isolated point of $E^\triangle = E \bigcup \{\triangle\}$. The metric $d_\triangle : E^\triangle \times E^\triangle \to [0,1]$ defined by

$$d_\triangle(x,y) = d(x,y)\,\mathbf{1}_E(x)\mathbf{1}_E(y) + \left|\mathbf{1}_{\{\triangle\}}(x) - \mathbf{1}_{\{\triangle\}}(y)\right|$$

turns $E^\triangle$ into a complete metrizable space. Moreover, if (E,d) is separable, then so is $\left(E^\triangle, d_\triangle\right)$. We also notice that the function $x \mapsto \mathbf{1}_E(x)$, $x \in E^\triangle$, belongs to $C_b\left(E^\triangle\right)$.

Remark 2.10. Let $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ be an evolution family on $C_b(E)$. Suppose that for any sequence of functions $(f_n)_{n\in\mathbb{N}}$ which decreases pointwise to zero $\lim_{n\to\infty} P(\tau,t)\,f_n(x) = 0$, $0 \leqslant \tau \leqslant t \leqslant T$. Then there exists a family of Borel measures $\{B \mapsto P(\tau,x;t,B) : 0 \leqslant \tau \leqslant t \leqslant T\}$ such that

$$P(\tau,t)\,f(x) = \int f(y)P(\tau,x;t,dy)\,, \quad f \in C_b(E). \tag{2.98}$$

This is a consequence of Corollary 2.1. In addition the family

$$\{B \mapsto P(\tau,x;t,B) : 0 \leqslant \tau \leqslant t \leqslant T\}$$

satisfies the equation of Chapman-Kolmogorov:

$$\int P(\tau,x;s,dz)\,P(s,z;t,B) = P(\tau,x;t,B)\,, \quad 0 \leqslant \tau \leqslant s \leqslant t \leqslant T, \quad B \in \mathcal{E}. \tag{2.99}$$

Next, for $B \in \mathcal{E}^{\triangle}$, and $0 \leqslant \tau \leqslant t \leqslant T$ we put

$$N(\tau,x;t,B) = P\left(\tau,x;t,B\bigcap E\right) + (1 - P(\tau,x;t,E))\,\mathbf{1}_B(\triangle), \ x \in E, \text{ and}$$
$$N(\tau,\triangle;t,B) = \mathbf{1}_B(\triangle)\,. \tag{2.100}$$

Then the family

$$\{B \mapsto N(\tau,x;t,B) : 0 \leqslant \tau \leqslant t \leqslant T\}$$

satisfies the Chapman-Kolmogorov equation on $E^{\triangle}$, $N\left(\tau,x;t,E^{\triangle}\right) = 1$, and $N(\tau,\triangle;t,E) = 0$. So that if $B \mapsto P(\tau,x;t,B)$ is a sub-probability on $\mathcal{E}$, then $B \mapsto N(\tau,x;t,B)$ is a probability measure on $\mathcal{E}^{\triangle}$, the Borel field of $E^{\triangle}$.

Remark 2.11. Besides the family of (maximum) time operators $\{\vee_t : t \in [0,T]\}$ we have the following more or less natural families: $\{\wedge_t : t \in [0,T]\}$ (minimum time operators), and the time translation or time shift operators $\left\{\vartheta_t^T : t \in [0,T]\right\}$. Instead of ϑ_t^T we usually write ϑ_t. The operators $\wedge_t : \Omega \to \Omega$ have the basic properties: $\wedge_s \circ \wedge_t = \wedge_{s\wedge t}$, $s,\ t \in [0,T]$, and $X(s) \circ \wedge_t = X(s \wedge t)$, $s,\ t \in [0,T]$. The operators $\vartheta_t : \Omega \to \Omega$, $t \in [0,T]$, have the following basic properties: $\vartheta_s \circ \vartheta_t = \vartheta_{s+t}$, $s, t \in [0,T]$, and $X(s) \circ \vartheta_t = X((s+t) \wedge T) = X(\vartheta_{s+t}(0))$. Compare with subsection 2.2.1.

It is clear that if a diffusion process, i.e. a $\mathcal{P}_{\tau,x}$-almost surely continuous Markov process $(X(t), \Omega, \mathcal{F}_t^{\tau}, \mathbb{P}_{\tau,x})$ generated by the family of operators $L(\tau)$, $\tau \in [0,T]$, exists, then for every pair $(\tau,x) \in [0,T] \times \mathbb{R}^d$, the measure $\mathbb{P}_{\tau,x}$ solves the martingale problem $\pi(\tau,x)$. Conversely, if the family $L(\tau)$, $\tau \in [0,T]$, is given, we can try to solve the martingale problem for all $(\tau,x) \in [0,T] \times \mathbb{R}^d$, find the measures $\mathbb{P}_{\tau,x}$, and then try to prove that

$X(t)$ is a Markov process with respect to the family of measures $\mathbb{P}_{\tau,x}$. For instance, if we know that for every pair $(\tau, x) \in [0, T] \times \mathbb{R}^d$ the martingale problem $\pi(\tau, x)$ is uniquely solvable, then the Markov property holds, provided that there exists operators $\vee_s : \Omega \to \Omega$, $0 \leqslant s \leqslant T$ such that $X_t \circ \vee_s = X_{t \vee s}$, $\mathbb{P}_{\tau,x}$-almost surely for $\tau \leqslant t \leqslant T$, and $\tau \leqslant s \leqslant T$. For the time-homogeneous case see, e.g., [Ethier and Kurtz (1986)] or [Ikeda and Watanabe (1998)]. The martingale problem goes back to Stroock and Varadhan (see [Stroock and Varadhan (1979)]). It found numerous applications in various fields of Mathematics. We refer the reader to [Liggett (2005)], [Kolokoltsov (2004b)], and [Kolokoltsov (2004a)] for more information about and applications of the martingale problem. In [Eberle (1999)] the reader may find singular diffusion equations which possess or which do not possess unique solutions. Consequently, for (singular) diffusion equations without unique solutions the martingale problem is not uniquely solvable. Another important example is given by Nadirashvili [Nadirashvili (1997)].

Remark 2.12. Examples of (Feller) semigroups can be manufactured by taking a continuous function $\varphi : [0, \infty) \times E \to E$ with the property that

$$\varphi(s + t, x) = \varphi(t, \varphi(s, x)),$$

for all s, $t \geqslant 0$ and $x \in E$. Then the mappings $f \mapsto P(t)f$, with $P(t)f(x) = f(\varphi(t, x))$ defines a semigroup. It is a Feller semigroup if $\lim_{x \to \triangle} \varphi(t, x) = \triangle$. An explicit example of such a function, which does not provide a Feller-Dynkin semigroup on $C_0(\mathbb{R})$ is given by $\varphi(t, x) = \dfrac{x}{\sqrt{1 + \frac{1}{2}tx^2}}$ (example due to V. Kolokoltsov). Put $u(t, x) = P(t)f(x) = f(\varphi(t, x))$. Then $\dfrac{\partial u}{\partial t}(t, x) = -x^3 \dfrac{\partial u}{\partial x}(t, x)$. In fact this (counter-)example shows that solutions to the martingale problem do not necessarily give rise to Feller-Dynkin semigroups. These are semigroups which preserve not only the continuity, but also the fact that functions which tend to zero at $\triangle$ are mapped to functions with the same property. However, for Feller semigroups we only require that continuous functions with values in $[0, 1]$ are mapped to continuous functions with the same properties. Therefore, it is not needed to include a hypothesis like (2.101) below in Theorem 2.12. Here (2.101) reads as follows: for every $(\tau, s, t, x) \in [0, T]^3 \times E$, $\tau < s < t$, the equality

$$\mathbb{P}_{\tau,x}[X(t) \in E] = \mathbb{P}_{\tau,x}[X(t) \in E,\ X(s) \in E] \tag{2.101}$$

holds. On the other hand this hypothesis is implicitly assumed, because as sample path space we take the Skorohod space $D\left([0, T], E^{\triangle}\right)$. If $X(t) \in E$, then $0 \leqslant s < t$ implies $X(s) \in E$.

In fact the result as stated is correct, but in case E happens to be locally compact, then the resulting semigroup need not be a Feller-Dynkin semigroup. This means that the corresponding family of operators assigns bounded continuous functions to functions in $C_0(E)$, but they need not vanish at $\triangle$. This means that the main result, Theorem 2.5, as stated in [van Casteren (1992)] is not correct. That is solutions to the martingale problem can, after having visited $\triangle$, still be alive.

Nadirashvili [Nadirashvili (1997)] constructs an elliptic operator in a bounded open domain $U \subset \mathbb{R}^d$ with a regular boundary such that the martingale problem is not uniquely solvable. More precisely the result reads as follows. Consider an elliptic operator $L = \sum_{j,k=1}^{d} a_{j,k}^2 \frac{\partial^2}{\partial x_j \partial x_k}$, where $a_{j,k} = aj,k$ are measurable functions on $\mathbb{R}^d$ such that

$$c^{-1} |\xi|^2 \leqslant \sum_{j,k=1}^{d} a_{j,k} \xi_j \xi_k \leqslant c |\xi|^2, \quad \xi \in \mathbb{R}^d,$$

for some ellipticity constant $c \geqslant 1$. There exists a diffusion process $(X(t), \mathbb{P}_x)$ corresponding to the operator L which can be defined as a solution to the martingale problem, i.e. $\mathbb{P}[X(0) = x]$, and the process

$$t \mapsto f(X(t)) - f(X(0)) - \int_0^t Lf(X(s))\, ds,\ t \geqslant 0,$$

is a local $\mathbb{P}_x$-martingale for all $f \in C^2(\mathbb{R}^d)$. For more details on diffusion processes see the comments after Remark 2.11. Nadirashvili is interested in nonuniqueness in the above martingale problem and in nonuniqueness of solutions to the Dirichlet problem $Lu = 0$ in Ω, the unit ball in $\mathbb{R}^d$, $u = g$ on $\partial\Omega$, where $\Omega \subset \mathbb{R}^d$ is a bounded domain with smooth boundary and $g \in C^2(\partial\Omega)$. In particular, so-called good solutions u to the Dirichlet problem are investigated. These functions u are the limit of a subsequence of solutions u_n, $n \in \mathbb{N}$, to $L^n u_n = \sum_{j,k=1}^{d} a_{j,k}^n \frac{\partial^2 u^n}{\partial x_j \partial x_k} = 0$ in Ω, $u^n = g$ on $\partial\Omega$, where the operators L^n are elliptic with smooth coefficients $a_{j,k}^n$ and a common ellipticity constant c such that $a_{j,k}^n \to a_{j,k}$ almost everywhere in Ω as $n \to \infty$. The main result is the following theorem: There exists an elliptic operator L of the above form defined in the unit ball $B_1 \subset \mathbb{R}^d$, $d \geqslant 3$, and there is a function $g \in C^2(\partial B_1)$ such that the formulated Dirichlet problem has at least two good solutions. An immediate consequence is nonuniqueness in the corresponding martingale problem.

In the case of a non-compact space the metric without the Lévy part is not adequate enough. That is why we have added the Lévy term. The problem is that the limits of the finite-dimensional distributions, given in (3.108) below, on its own need not be a measure, and so there is no way of applying Kolmogorov's extension theorem.

For applications of the martingale problem in relation to partially observed systems and hidden Markov processes see e.g. a forthcoming book [Kurtz and Nappo (2010)], which goes back to [Kurtz and Ocone (1988)], and to [Kurtz (1998)].

2.3.1 *Some historical remarks and references*

In [Dorroh and Neuberger (1993)] the authors also use the strict topology to describe the behavior of semigroups acting on the space of bounded continuous functions on a Polish space. In fact the author of the present book was at least partially motivated by their work to establish a general theory for Markov processes on Polish spaces. Another motivation is provided by results on bi-topological spaces as established by e.g. Kühnemund in [Kühnemund (2003)]. Other authors have used this concept as well, e.g. Es-Sarhir and Farkas in [Es-Sarhir and Farkas (2005)]. The notion of "strict topology" plays a dominant role in Hirschfeld [Hirschfeld (1974)]. As already mentioned Buck [Buck (1958)] was the first author who introduced the notion of strict topology (in the locally compact setting). He denoted it by β in §3 of [Buck (1958)]. There are several other authors who used it and proved convergence and approximation properties involving the strict topology: Buck [Buck (1974)], Prolla [Prolla (1993)], Prolla and Navarro [Prolla and Navarro (1997)], Katsaras [Katsaras (1983)], Ruess [Ruess (1977)], Giles [Giles (1971)], Todd [Todd (1965)], Wells [Wells (1965)]. This list is not exhaustive: the reader is also referred to Prolla [Prolla (1977)], and the literature cited there. The strict topology is also called the mixed topology: see e.g. Goldys and van Neerven [Goldys and van Neerven (2003)], Wiweger [Wiweger (1961)], Sentilles [Sentilles (1972)], and Wheeler [Wheeler (1983)]. In [Cerrai (2001)] and [Cerrai (1994)] Cerrai calls the corresponding convergence the $\mathcal{K}$-convergence: see Definition B.1.1 in [Cerrai (2001)].

In [Varadhan (2007)] Varadhan describes a metric on the space $D([0,1],\mathbb{R})$ which turns it into a complete metrizable separable space; i.e. the Skorohod topology turns $D([0,1],\mathbb{R})$ into a Polish space. On the other hand it is by no means necessary that the Skorohod topology is the most natural topology to be used on the space $D\left([0,1],\mathbb{R}^d\right)$. For example in [Jakubowski (1997)] Jakubowski employs a quite different topology on this space. In [Jakubowski (2000)] Jakubowski elaborates on Skorohod's ideas about sequential convergence of distributions of stochastic processes. After that the S-topology, as introduced by Jakubowski, has been used by several others as well: see the references in [Boufoussi and van Casteren (2004a)] as well. Definition 2.17 below also appears in [Boufoussi and van Casteren (2004a)]. Although the definition is confined to $\mathbb{R}$-valued paths, the S-topology also extends easily to the finite dimensional Euclidean space $\mathbb{R}^d$. By $\mathcal{V}_+ \subset D\left([0,T],\mathbb{R}\right)$ we denote the space of nonnegative and nondecreasing functions $V : [0,T] \to [0,\infty)$ and $\mathcal{V} = \mathcal{V}_+ - \mathcal{V}_+$. We know that any element $V \in \mathcal{V}_+$ determines a unique positive measure dV on $[0,T]$ and $\mathcal{V}$ can be equipped with the topology of weak convergence of measures; i.e. the equality $\lim_{n\to\infty}\int_0^T \varphi(s)dV_n(s) = \int_0^T \varphi(s)dV(s)$ for all functions $\varphi \in C\left([0,T],\mathbb{R}\right)$ describes the weak convergence of the sequence $\left(V_n\right)_{n\in\mathbb{N}} \subset \mathcal{V}$ to $V \in \mathcal{V}$. Without loss of generality we may assume that the functions $V \in \mathcal{V}$ are right-continuous and possess left limits in $\mathbb{R}$.

Definition 2.17. Let $\left(Y^n\right)_{1\leqslant n\leqslant\infty} \subset D\left([0,T],\mathbb{R}\right)$. The sequence $\left(Y^n\right)_{n\in\mathbb{N}}$ is said to converges to Y^∞ with respect to the S-topology, if for every $\varepsilon > 0$ there exist elements $\left(V^{n,\varepsilon}\right)_{1\leqslant n\leqslant\infty} \subset \mathcal{V}$ such that $\left\|V^{n,\varepsilon} - Y^n\right\|_\infty \leqslant \varepsilon$, $n = 1,\dots,\infty$, and $\displaystyle\lim_{n\to\infty}\int_0^T \varphi(s)\,dV^{n,\varepsilon}(s) = \int_0^T \varphi(s)\,dV^{\infty,\varepsilon}(s)$, for all $\varphi \in C\left([0,T],\mathbb{R}\right)$.

2.4 Dini's lemma, Scheffé's theorem, and the monotone class theorem

The contents of this section is taken from Appendix E in [Demuth and van Casteren (2000)]. In this section we formulate and discuss these three theorems.

2.4.1 *Dini's lemma and Scheffé's theorem*

The contents of this subsection is devoted to Dini's lemma and Scheffé's theorem. Another proof of Dini's lemma can be found in [Stroock (1999)], Lemma 7.1.23, p. 146.

Lemma 2.2. *(Dini) Let $\left(f_n : n \in \mathbb{N}\right)$ be a sequence of continuous functions on the locally compact Hausdorff space E. Suppose that $f_n(x) \geqslant f_{n+1}(x) \geqslant$*

0 for all $n \in \mathbb{N}$ and for all $x \in E$. If $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in E$, then, for all compact subsets K of E, $\lim_{n\to\infty} \sup_{x\in K} f_n(x) = 0$. If the function f_1 belongs to $C_0(E)$, then $\lim_{n\to\infty} \sup_{x\in E} f_n(x) = 0$.

Proof. We only prove the second assertion. Fix $\eta > 0$ and consider the subset

$$\bigcap_{n\in\mathbb{N}} \{x \in E : f_n(x) \geqslant \eta\} .$$

Since, by assumption, the function f_1 belongs to $C_0(E)$, and $\lim\limits_{n\to\infty} f_n(x) = 0$, $x \in E$, it follows that the intersection

$$\bigcap_{n\in\mathbb{N}} \{x \in E : f_n(x) \geqslant \eta\}$$

is void. As a consequence $E = \bigcup_{n\in\mathbb{N}} \{f_n < \eta\}$. Let $\varepsilon > 0$ and put $K = \{f_1 \geqslant \varepsilon\}$. The subset K is compact. By the preceding argument there exist $n_\varepsilon \in \mathbb{N}$ for which $K \subseteq \{f_{n_\varepsilon} < \varepsilon\}$. For $n \geqslant n_\varepsilon$, we have $0 \leqslant f_n(x) \leqslant \varepsilon$ for all $x \in E$.

This completes the proof of Lemma 2.2. □

In Definition 2.18 and in Theorem 2.14 of this subsection $(E, \mathcal{E}, m)$ may be any measure space with $m(B) \geqslant 0$ for $B \in \mathcal{E}$.

Definition 2.18. A collection of functions $\{f_j : j \in J\}$ in $L^1(E, \mathcal{E}, m)$ is *uniformly L^1-integrable* if for every $\varepsilon > 0$ there exists $g \in L^1(E, \mathcal{E}, \mu)$, $g \geqslant 0$, for which

$$\sup_{j\in J} \int_{\{|f_j| \geqslant g\}} |f_j| \, dm \leqslant \varepsilon.$$

Remark 2.13. If the collection $\{f_j : j \in J\}$ is uniformly L^1-integrable, and if $\{g_j : j \in J\}$ is a collection for which $|g_j| \leqslant |f_j|$, m-almost everywhere, for all $j \in J$, then the collection $\{g_j : j \in J\}$ is uniformly L^1-integrable as well.

Remark 2.14. Cauchy sequences in $L^1(E, \mathcal{E}, m)$ are uniformly L^1-integrable.

Remark 2.15. Let $f \geqslant 0$ be a function in $L^1(\mathbb{R}^\nu, \mathcal{B}, m)$, where m is the Lebesgue measure. Suppose $\int f(x) dm(x) = 1$ and $\lim_{n\to\infty} n^\nu f(nx) = 0$ for all $x \neq 0$. Put $f_n(x) = n^\nu f(nx)$, $n \in \mathbb{N}$. Then the sequence is not uniformly L^1-integrable. This will follow from Theorem 2.14 below.

A version of Scheffé's theorem reads as follows. Our proof uses the arguments in the proof of Theorem 3.3.5 (Lieb's version of Fatou's lemma) in Stroock [Stroock (1999)], p. 54. Another proof can be found in Bauer [Bauer (1981)], Theorem 2.12.4, p. 103.

Theorem 2.14. *(Scheffé) Let* $(f_n : n \in \mathbb{N})$ *be a sequence in* $L^1(E, \mathcal{E}, m)$. *If* $\lim_{n\to\infty} f_n(x) = f(x)$, *m-almost everywhere, then the sequence* $(f_n : n \in \mathbb{N})$ *is uniformly* L^1*-integrable if and and only*

$$\lim_{n\to\infty} \int |f_n(x)|\, dm(x) = \int |f(x)|\, dm(x).$$

Proof. Consider the m-almost everywhere pointwise inequality

$$0 \leqslant |f_n - f| + |f| - |f_n| \leqslant 2\,|f|\,. \tag{2.102}$$

First suppose that the sequence $\{f_n : n \in \mathbb{N}\}$ is uniformly L^1-integrable. Then, by Fatou's lemma,

$$\int |f(x)|\, dm(x) = \int \liminf |f_n(x)|\, dm(x) \leqslant \liminf \int |f_n(x)|\, dm(x)$$

(choose $g \in L^1(E, m)$ such that $\int_{\{|f_n|\geqslant g\}} |f_n(x)|\, dm(x) \leqslant 1$)

$$\begin{aligned} &\leqslant \liminf \int_{\{|f_n|\geqslant g\}} |f_n(x)|\, dm(x) + \int_{\{|f_n|\leqslant g\}} |f_n(x)|\, dm(x) \\ &\leqslant 1 + \int g(x) dm(x). \end{aligned} \tag{2.103}$$

From (2.103) we see that the function f belongs to $L^1(E, m)$. From Lebesgue's dominated convergence theorem in conjunction with (2.103) we infer

$$\lim_{n\to\infty} \int \left(|f_n - f| + |f| - |f_n|\right) dm = 0. \tag{2.104}$$

Since the sequence $\{f_n : n \in \mathbb{N}\}$ is uniformly L^1-integrable, and since for m-almost all x, $\lim_{n\to\infty} f_n(x) = f(x)$, we see that $\lim_{n\to\infty} \int |f_n - f|\, dm = 0$. So from (2.104) we get

$$\lim_{n\to\infty} \int |f_n|\, dm = \int |f|\ dm < \infty. \tag{2.105}$$

Conversely, suppose (2.105) holds. Then f belongs to $L^1(E, m)$. Again we may invoke Lebesgue's dominated convergence theorem to conclude (2.104) from (2.102). Again using (2.105) implies $\lim_{n\to\infty} \int |f_n - f|\, dm = 0$. An appeal to Remark 2.14 yields the desired result in Theorem 2.14. $\square$

2.4.2 *Monotone class theorem*

Our presentation of the monotone class theorems is taken from [Blumenthal and Getoor (1968)], pp. 5–7. For other versions of this theorem see e.g. [Sharpe (1988)], pp. 364–366. Theorems 2.15, 2.16, and Propositions 2.9, 2.10 we give closely related versions of this theorem.

Definition 2.19. Let Ω be a set and let $\mathcal{S}$ be a collection of subsets of Ω. Then $\mathcal{S}$ is a *Dynkin* system if it has the following properties:

(a) $\Omega \in \mathcal{S}$;
(b) if A and B belong to $\mathcal{S}$ and if $A \supseteq B$, then $A \backslash B$ belongs to $\mathcal{S}$;
(c) if $(A_n : n \in \mathbb{N})$ is an increasing sequence of elements of $\mathcal{S}$, then the union $\bigcup_{n=1}^{\infty} A_n$ belongs to $\mathcal{S}$.

The following result on Dynkin systems is well-known.

Theorem 2.15. *Let $\mathcal{M}$ be a collection of subsets of of Ω, which is stable under finite intersections. The Dynkin system generated by $\mathcal{M}$ coincides with the σ-field generated by $\mathcal{M}$.*

Remark 2.16. A collection of subsets of Ω which is closed under finite intersections is also called a π-system. A collection of subsets $\mathcal{L}$ of Ω is called a λ-system if it has the following properties: (1) $\Omega \in \mathcal{L}$; (2) if A belongs to $\mathcal{L}$, then its complement A^c also is a member of $\mathcal{L}$; (3) if $(A_j)_{j\in\mathbb{N}}$ is a mutually disjoint sequence in $\mathcal{L}$, then its union $\bigcup_j A_j$ belongs to $\mathcal{L}$. If a λ-system $\mathcal{L}$ is at the same time a π-system, then it is a σ-field. The π-λ theorem says that the smallest λ-system containing a given π-system $\mathcal{P}$ coincides with the σ-field generated by $\mathcal{P}$. The π-λ theorem is closely related to Theorem 2.15. For more details see e.g. Vestrup [Vestrup (2003)].

Theorem 2.16. *Let Ω be a set and let $\mathcal{M}$ be a collection of subsets of of Ω, which is stable (or closed) under finite intersections. Let $\mathcal{H}$ be a vector space of real valued functions on Ω satisfying:*

(i) The constant function $\mathbf{1}$ belongs to $\mathcal{H}$ and $\mathbf{1}_A$ belongs to $\mathcal{H}$ for all $A \in \mathcal{M}$;
(ii) if $(f_n : n \in \mathbb{N})$ is an increasing sequence of non-negative functions in $\mathcal{H}$ such that $f = \sup_{n\in\mathbb{N}} f_n$ is finite (bounded), then f belongs to $\mathcal{H}$.

Then $\mathcal{H}$ contains all real valued functions (bounded) functions on Ω, that are $\sigma(\mathcal{M})$ measurable.

Proof. Put $\mathcal{D} = \{A \subseteq \Omega : \mathbf{1}_A \in \mathcal{H}\}$. Then by (i) Ω belongs to $\mathcal{D}$ and $\mathcal{D} \supseteq \mathcal{M}$. If A and B are in $\mathcal{D}$ and if $B \supseteq A$, then $B \backslash A$ belongs to $\mathcal{D}$. If $(A_n : n \in \mathbb{N})$ is an increasing sequence in $\mathcal{D}$, then $\mathbf{1}_{\bigcup A_n} = \sup_n \mathbf{1}_{A_n}$ belongs to $\mathcal{D}$ by (ii). Hence $\mathcal{D}$ is a Dynkin system, that contains $\mathcal{M}$. Since $\mathcal{M}$ is closed under finite intersection, it follows by Theorem 2.15 that $\mathcal{D} \supseteq \sigma(\mathcal{M})$. If $f \geqslant 0$ is measurable with respect to $\sigma(\mathcal{M})$, then

$$f = \sup_n \frac{1}{2^n} \sum\nolimits_{j=1}^{n2^n} \mathbf{1}_{\{f \geqslant j2^{-n}\}} = \sup_n \frac{1}{2^n} \left\lfloor 2^n \min(f, n) \right\rfloor . \tag{2.106}$$

Since the functions $\mathbf{1}_{\{f \geqslant j2^{-n}\}}$, j, $n \in \mathbb{N}$, are $\sigma(\mathcal{M})$-measurable, we see that f belongs to $\mathcal{H}$. Here we employed the fact that $\sigma(\mathcal{M}) \subseteq \mathcal{D}$. If f is $\sigma(\mathcal{M})$-measurable, then we write f as a difference of two non-negative $\sigma(\mathcal{M})$-measurable functions. This establishes Theorem 2.16. □

The previous theorems, i.e. Theorems 2.15 and 2.16, are used in the following form. Let Ω be a set and let $(E_i, \mathcal{E}_i)_{i \in I}$ be a family of measurable spaces, indexed by an arbitrary set I. For each $i \in I$, let $\mathcal{S}_i$ denote a collection of subsets of of E_i, closed under finite intersection, which generates the σ-field $\mathcal{E}_i$, and let $f_i : \Omega \to E_i$ be a map from Ω to E_i. In our presentation of the Markov property the space E_i are all the same, and the maps f_i, $i \in I$, are the state variables $X(t)$, $t \geqslant 0$. in this context the following two propositions follow.

Proposition 2.9. *Let $\mathcal{M}$ be the collection of all sets of the form $\bigcap_{i \in J} f_i^{-1}(A_i)$, $A_i \in \mathcal{S}_i$, $i \in J$, $J \subseteq I$, J finite. Then $\mathcal{M}$ is a collection of subsets of Ω which is stable under finite intersection and $\sigma(\mathcal{M}) = \sigma(f_i : i \in I)$.*

Proposition 2.10. *Let $\mathcal{H}$ be a vector space of real-valued functions on Ω such that:*

(i) the constant function $\mathbf{1}$ belongs to $\mathcal{H}$;
(ii) if $(h_n : n \in \mathbb{N})$ is an increasing sequence of non-negative functions in $\mathcal{H}$ such that $h = \sup_n h_n$ is finite (bounded), then h belongs to $\mathcal{H}$;
(iii) $\mathcal{H}$ contains all products of the form $\prod_{i \in J} \mathbf{1}_{A_i} \circ f_i$, $J \subseteq I$, J finite, and $A_i \in \mathcal{S}_i$, $i \in J$.

Under these assumptions $\mathcal{H}$ contains all real-valued functions (bounded) functions in $\sigma(f_i : i \in I)$.

Definition 2.20. Theorems 2.15 and 2.16, and Propositions 2.9 and 2.10 are called the *monotone class theorems*.

Other theorems and results on integration theory, not explained in the book, can be found in any textbook on the subject. In particular this is true for Fatou's lemma and Fubini's theorem on the interchange of the order of integration. Proofs of these results can be found in [Bauer (1981)] and [Stroock (1999)]. The same references contain proofs of the Radon-Nikodym theorem. This theorem may be phrased as follows.

Theorem 2.17. *(Radon-Nikodym) If a finite measure μ on some σ-finite measure space $(E, \mathcal{E}, m)$ is absolutely continuous with respect to m, then there exists a function $f \in L^1(E, \mathcal{E}, m)$ such that $\mu(A) = \int_A f(x)dm(x)$ for all subsets $A \in \mathcal{E}$.*

The measure μ is said to be absolutely continuous with respect to m if $m(A) = 0$ implies $\mu(A) = 0$, and the measure m is said to be σ-finite if there exists an increasing sequence $(E_n : n \in \mathbb{N})$ in $\mathcal{E}$ such that $E = \bigcup_{n\in\mathbb{N}} E_n$ and for which $m(E_n) < \infty$, $n \in \mathbb{N}$. A very important application is the existence of *conditional expectations*. This can be seen as follows.

Corollary 2.5. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{F}_0$ be a sub-field of $\mathcal{F}$, and let $Y : \Omega \to [0, \infty]$ be a $\mathcal{F}$-measurable function (random variable) in $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists a function $G \in L^1(E, \mathcal{F}_0, m)$ such that $E[Y\mathbf{1}_A] = \mu(A) = \mathbb{E}[G\mathbf{1}_A]$ for all $A \in \mathcal{F}_0$.*

By convention the random variable G is written as $G = \mathbb{E}\left[Y \mid \mathcal{F}_0\right]$. It is called the conditional expectation on the σ-field $\mathcal{F}_0$.

Proof. Put $m(A) = \mathbb{E}[Y\mathbf{1}_A]$, $A \in \mathcal{F}$, and let μ be the restriction of m to $\mathcal{F}_0$. If for some $A \in \mathcal{F}_0$, $m(A) = 0$, then $\mu(A) = 0$. The Radon-Nikodym theorem yields the existence of a function $G \in L^1(E, \mathcal{F}_0, m)$ such that $E[Y\mathbf{1}_A] = \mu(A) = \mathbb{E}[G\mathbf{1}_A]$ for all $A \in \mathcal{F}_0$. □

2.4.3 *Some additional information*

The reader may find additional material about strong Markov process theory in [Ethier and Kurtz (1986)], [Gillespie (1992)], [Sharpe (1988)]. Material about infinite-dimensional stochastic processes and calculus can be found in e.g. [Da Prato and Zabczyk (1992a, 1996); Cerrai (2001); Hairer *et al.* (2004); Hairer (2009); Seidler (1997); Sanz-Solé (2005)]. In the following references the reader may find topics on or related to Malliavin calculus: [Bell (2006); Malliavin (1978); Norris (1986); Nualart (1998, 2006, 2009); Üstünel and Zakai (2000b); Kusuoka and Stroock (1984, 1985, 1987)]. The following references discuss some more general stochastic processes including a number of concrete examples: [Kallenberg (2002)], [Shanbhag and Rao (2001)]. In the following references the authors apply Malliavin calculus to models in financial mathematics: [Di Nunno *et al.* (2009)], [Malliavin and Thalmaier (2006)]. Notice that Malliavin calculus or stochastic variation calculus is "by nature" an infinite-dimensional calculus. For recent results on stochastic partial differential equations see e.g. [Zhang (2010); Kotelenez and Kurtz (2010); Holden *et al.* (2010)].

This page is intentionally left blank

Chapter 3

Strong Markov processes: Proof of main results

The present chapter is completely devoted to a proof of the main results of Part 2. Except for some results from Chapter 4 all proofs can be found in the present chapter. Together with the results in Chapter 4 all proofs are self-contained. Unfortunately there is quite a bit of technicality involved in all these proofs. This technicality is due to fact that we are working in a Polish space which is not necessarily locally compact, that the Markov processes involved are time-dependent, may have finite life time, and may have jumps.

3.1 Proof of the main results: Theorems 2.9 through 2.13

In the present chapter we will prove Theorems 2.9, 2.10, 2.11, 2.12 and 2.13, which form the main results of Part 2 of this book. We will need a number of auxiliary results which can be found in the current section or occasionally in the sections 4.1 and 4.2. In particular the latter is true for Proposition 4.1, Proposition 4.4 and its Corollary 4.2, Corollary 4.3 to Proposition 4.5, and Theorem 4.4. We will always give the relevant references. We need the following definition.

Definition 3.1. Let $\{X(t)\}_{t\in[0,T]}$, and $\{Y(t)\}_{t\in[0,T]}$ be stochastic processes on $\left(\Omega, \mathcal{F}_T^0, \mathbb{P}\right)$. The process $\{X(t)\}_{t\in[0,T]}$ is a modification of $\{Y(t)\}_{t\in[0,T]}$ if $P\left[X(t) = Y(t)\right] = 1$ for all $t \in [0,T]$.

3.1.1 *Proof of Theorem 2.9*

This subsection contains the proof of Theorem 2.9. It employs the Kolmogorov's extension theorem and it uses the Polish nature of the state space

E in an essential way. For more details on the Kolmogorov extension (or existence) theorem see e.g. [Aliprantis and Border (1994)], [Bhattacharya and Waymire (2007)], [Neveu (1965)], and [Dudley (2002)]. In subsection 3.1.7 the reader will find some information; in particular Theorem 3.1 is essential in this aspect. One of the main difficulties is to prove that orbits of the form $\left\{\widetilde{X}(s) : \tau \leqslant s \leqslant t,\, \widetilde{X}(t) \in E\right\}$ are $\mathbb{P}_{\tau,x}$-almost surely contained in (sequentially) relatively compact subsets of E: for details see Proposition 3.2 below.

Proof. [Proof of Theorem 2.9.] We begin with the proof of the existence of a Markov process (2.90), starting from a Feller evolution: see Definition 2.4. First we assume $P(\tau,t)\mathbf{1} = \mathbf{1}$. Remark 2.10 will be used to prove Theorem 2.9 in case $P(\tau,t)\mathbf{1} < \mathbf{1}$. Temporarily we write $\Omega = E^{[0,T]}$ endowed with the product topology, and product σ-field (also called product σ-algebra), which is the smallest σ-field on Ω which renders all coordinate mappings, or state variables, measurable. The state variables $X(t) : \Omega \to E$ are defined by $X(t,\omega) = X(t)(\omega) = \omega(t)$, $\omega \in \Omega$, and the maximum mappings $\vee_s : \Omega \to \Omega$, $s \in [0,T]$, are defined by $\vee_s(\omega)(t) = \omega(s \vee t)$. Let the family of Borel measures on

$$\{B \mapsto P(\tau,x;t,B) : B \in \mathcal{E},\, (\tau,x) \in [0,T] \times E,\, t \in [\tau,T]\} \tag{3.1}$$

be determined by the equalities:

$$P(\tau,t) f(x) = \int f(y) P(\tau,x;t,dy), \quad f \in C_b(E). \tag{3.2}$$

By Kolmogorov's extension theorem (see Theorem 3.1 below) there exists a family of probability spaces

$$(\Omega, \mathcal{F}_T^\tau, \mathcal{P}_{\tau,x}), \quad (\tau,x) \in [0,T] \times E,$$

such that

$$\begin{aligned}&\mathbb{E}_{\tau,x}\left[f\left(X(t_1),\dots,X(t_n)\right)\right]\\ &= \underbrace{\int\dots\int}_{n\times} f(y_1,\dots,y_n)\, P(\tau,x;t_1,dy_1)\dots P(t_{n-1},y_{n-1};t_n,dy_n)\end{aligned} \tag{3.3}$$

where $\tau \leqslant t_1 < \dots < t_n \leqslant T$, and $f \in L^\infty\left(E^n, \mathcal{E}^{\otimes^n}\right)$. Notice that a family of probability spaces together with a process $t \mapsto X(t)$ such that (3.1), (3.2) and (3.3) are satisfied is Markov process in the sense that for $(\tau,x) \in [0,T] \times E$ and $s \in [\tau,T]$ the following equality holds $\mathbb{P}_{\tau,x}$-almost surely:

$$\mathbb{E}_{\tau,x}\left[f\left(X(t_1),\dots,X(t_n)\right) \,\middle|\, \mathcal{F}_s^\tau\right] = \mathbb{E}_{s,X(s)}\left[f\left(X(t_1),\dots,X(t_n)\right)\right] \tag{3.4}$$

for all bounded Borel measurable functions f on E^n, and for all finite subsets $\tau \leqslant s \leqslant t_1 < \cdots < t_n \leqslant T$. In order to prove (3.4) the propagator property, i.e. $P(\rho_0, \rho) P(\rho, \rho_1) = P(\rho_0, \rho_1)$, $\rho_0 \leqslant \rho \leqslant \rho_1 \leqslant T$, is used several times. For $f \in C_b([0,T] \times E)$, $0 \leqslant f$, and $\alpha > 0$ given we introduce the following processes:

$$t \mapsto \alpha R(\alpha) f(t, X(t)) = \alpha \int_t^\infty e^{-\alpha(\rho - t)} P(t, \rho \wedge T) f(\rho \wedge T, \cdot)(X(t))\, d\rho$$

$$= \alpha \int_t^\infty e^{-\alpha(\rho - t)} \mathbb{E}_{t, X(t)}\left[f(\rho \wedge T, X(\rho \wedge T))\right], \quad t \in [0,T], \quad \text{and} \tag{3.5}$$

$$s \mapsto P(s,t) f(t, \cdot)(X(s)) = \mathbb{E}_{s, X(s)}\left[f(t, X(t))\right], \quad s \in [0,t],\ t \in [0,T]. \tag{3.6}$$

The processes in (3.5) and (3.6) could have been more or less unified by considering the process:

$$(s,t) \mapsto \alpha \int_t^\infty e^{-\alpha(\rho - t)} P(s, \rho \wedge T) f(\rho \wedge T, \cdot)(X(s))\, d\rho$$

$$= \alpha P(s,t) R(\alpha) f(t, \cdot)(X(s)), \quad 0 \leqslant s \leqslant t \leqslant T. \tag{3.7}$$

Observe that $\lim_{\alpha \to \infty} \alpha R(\alpha) f(t, X(t)) = f(t, X(t))$, $t \in [0,T]$. Here we use the continuity of the function $\rho \mapsto P(t,\rho) f(\rho, \cdot)(X(t))$ at $\rho = t$. In addition, for $(\tau, x) \in [0,T] \times E$ fixed, we have that the family of functionals $f \mapsto \alpha R(\alpha) f(t, \cdot)(X(t))$, $\alpha \geqslant 1$, $t \in [\tau, T]$, is $\mathbb{P}_{\tau,x}$-almost surely equi-continuous for the strict topology: see Corollary 2.4.

Our first task will be to prove that for every $(\tau, x) \in [0,T] \times E$ the orbit $\{(t, X(t)) : t \in [\tau, T]\}$ is a $\mathbb{P}_{\tau,x}$-almost surely sequentially compact. Therefore we choose an infinite sequence $(\rho_n, X(\rho_n))_{n \in \mathbb{N}}$ where $\rho_n \in [\tau, T]$, $n \in \mathbb{N}$. This sequence contains an infinite subsequence $(s_n, X(s_n))_{n \in \mathbb{N}}$ such that $s_n < s_{n+1}$, $n \in \mathbb{N}$, or an infinite subsequence $(t_n, X(t_n))_{n \in \mathbb{N}}$ such that $t_n > t_{n+1}$, $n \in \mathbb{N}$. In the first case we put $s = \sup_{n \in \mathbb{N}} s_n$, and in the second case we write $t = \inf_{n \in \mathbb{N}} t_n$. In either case we shall prove that there exists a subsequence which is $\mathbb{P}_{\tau,x}$-almost surely a Cauchy sequence in $[\tau, T] \times E$ for a compatible uniformly bounded metric. First we deal with the case that t_n decreases to $t \geqslant \tau$. Then we consider the stochastic process in (3.6) given by $\rho \mapsto \mathbb{E}_{t, X(t)}\left[f(\rho, X(\rho))\right]$ where f is an arbitrary function in $C_b([0,T] \times E)$. By hypothesis on the transition function $P(\tau, x; t, B)$ we have

$$\lim_{n \to \infty} \mathbb{E}_{t, X(t)}\left[f(t_n, X(t_n))\right]$$

$$= \lim_{n \to \infty} \int P(t, X(t); t_n, dy) f(t_n, y) = f(t, X(t)). \tag{3.8}$$

By applying the argument in (3.8) to the process $\rho \mapsto \mathbb{E}_{t,X(t)}[|f(\rho, X(\rho))|^2]$, $\rho \in [t, T]$, the Markov property implies

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right] - f\left(\rho, X(\rho)\right)\right|^2\right] \\
&= \mathbb{E}_{\tau,x}\left[\left|f\left(\rho, X(\rho)\right)\right|^2\right] + \mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right]\right|^2\right] \\
&\quad - 2\Re\mathbb{E}_{\tau,x}\left[\overline{f\left(\rho, X(\rho)\right)}\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right]\right]
\end{aligned}$$

(Markov property: $\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right] = \mathbb{E}_{\tau,x}\left[f\left(\rho, X(\rho)\right) \;\middle|\; \mathcal{F}_t^\tau\right]$ $\mathbb{P}_{\tau,x}$-almost surely)

$$\begin{aligned}
&= \mathbb{E}_{\tau,x}\left[\left|f\left(\rho, X(\rho)\right)\right|^2\right] + \mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right]\right|^2\right] \\
&\quad - 2\Re\mathbb{E}_{\tau,x}\left[\overline{\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right]}\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right]\right] \\
&= \mathbb{E}_{\tau,x}\left[\left|f\left(\rho, X(\rho)\right)\right|^2\right] - \mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{t,X(t)}\left[f\left(\rho, X(\rho)\right)\right]\right|^2\right]. \qquad (3.9)
\end{aligned}$$

Applying the argument in (3.8) to the process $\rho \mapsto \mathbb{E}_{t,X(t)}\left[\left|f\left(\rho, X(\rho)\right)\right|^2\right]$, $\rho \in [t, T]$, and employing (3.9) we obtain:

$$\lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{t,X(t)}\left[f\left(t_n, X\left(t_n\right)\right)\right] - f\left(\rho, X\left(\rho\right)\right)\right|^2\right] = 0. \qquad (3.10)$$

Again using (3.8) and invoking (3.10) we see that

$$\lim_{n\to\infty} f\left(t_n, X\left(t_n\right)\right) = f\left(t, X(t)\right)$$

in the space $L^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$. Hence there exists a subsequence denote by $\left(f\left(t_{n_k}, X\left(t_{n_k}\right)\right)\right)_{k\in\mathbb{N}}$ which converges $\mathbb{P}_{\tau,x}$-almost surely to $f\left(t, X(t)\right)$. Let $d : E \times E \to [0, 1]$ be a metric on E which turns it into a Polish space, and let $\left(x_j\right)_{j\in\mathbb{N}}$ be a countable dense sequence in E. The previous arguments are applied to the function $f : [0, T] \times E \to \mathbb{R}$ defined by

$$f(\rho, x) = \sum_{j=1}^{\infty} 2^{-j}\left(d\left(x_j, x\right) + \left|\rho_j - \rho\right|\right), \qquad (3.11)$$

where the sequence $\left(\rho_j\right)_{j\in\mathbb{N}}$ is a dense sequence in $[0, T]$. Like in the earlier reasoning there exists a subsequence $\left(t_{n_k}, X\left(t_{n_k}\right)\right)_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} f\left(t_{n_k}, X\left(t_{n_k}\right)\right) = f\left(t, X(t)\right), \quad \mathbb{P}_{\tau,x}\text{-almost surely.} \qquad (3.12)$$

It follows that $\lim_{k\to\infty} t_{n_k} = t$. From (3.12) we also infer that

$$\lim_{k\to\infty} d\left(x_j, X\left(t_{n_k}\right)\right) = d\left(x_j, X(t)\right), \quad \mathbb{P}_{\tau,x}\text{-almost surely for all } j \in \mathbb{N}. \qquad (3.13)$$

Since the sequence $(x_j)_{j\in\mathbb{N}}$ is dense in E we see that

$$\lim_{k\to\infty} d\left(y, X\left(t_{n_k}\right)\right) = d\left(y, X(t)\right), \quad \mathbb{P}_{\tau,x}\text{-almost surely for all } y \in E. \tag{3.14}$$

The substitution $y = X(t)$ in (3.14) shows that

$$\lim_{k\to\infty} \left(t_{n_k}, X\left(t_{n_k}\right)\right) = \left(t, X(t)\right), \quad \mathbb{P}_{\tau,x}\text{-almost surely}. \tag{3.15}$$

Again let $f \in C_b\left([0,T] \times E\right)$ be given. Next we consider the situation where we have an infinite subsequence $\left(s_n, X\left(s_n\right)\right)_{n\in\mathbb{N}}$ such that $s_n < s_{n+1}$, $n \in \mathbb{N}$. Put $s = \sup_{n\in\mathbb{N}} s_n$, and consider the process

$$\rho \mapsto \mathbb{E}_{\rho,X(\rho)}\left[f\left(s, X(s)\right)\right] = \int P\left(\rho, X(\rho); s, dy\right) f(s,y) = P\left(\rho, s\right) f\left(s, \cdot\right)\left(X(s)\right) \tag{3.16}$$

which is $\mathbb{P}_{\tau,x}$-martingale with respect to the filtration $\left(\mathcal{F}_t^\tau\right)_{\tau\leqslant\rho\leqslant s}$. Since the process in (3.16) is a martingale we know that the limit

$$\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f\left(s, X(s)\right)\right]$$

exists. We also have

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f\left(s, X(s)\right)\right]\right] = \lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s_n,X(s_n)}\left[f\left(s, X(s)\right)\right]\right] \\
&= \lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[\mathbb{E}_{\tau,X(\tau)}\left[f\left(s, X(s)\right) \,\middle|\, \mathcal{F}_{s_n}^\tau\right]\right] = \mathbb{E}_{\tau,x}\left[f\left(s, X(s)\right)\right].
\end{aligned} \tag{3.17}$$

Like in (3.9) we write

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{\rho,X(\rho)}\left[f\left(s, X(s)\right)\right] - f\left(\rho, X(\rho)\right)\right|^2\right] \\
&= \mathbb{E}_{\tau,x}\left[\left|\mathbb{E}_{\rho,X(\rho)}\left[f\left(s, X(s)\right)\right]\right|^2\right] + \mathbb{E}_{\tau,x}\left[\left|f\left(\rho, X(\rho)\right)\right|^2\right] \\
&\quad - 2\Re\mathbb{E}_{\tau,x}\left[\overline{f\left(\rho, X(\rho)\right)}\mathbb{E}_{\rho,X(\rho)}\left[f\left(s, X(s)\right)\right]\right].
\end{aligned} \tag{3.18}$$

The expression in (3.18) converges to 0 as $\rho \uparrow s$. Here we used the following identity:

$$\begin{aligned}
&\lim_{\rho\uparrow s} \mathbb{E}_{\tau,x}\left[g\left(\rho, X(\rho)\right)\right] = \lim_{\rho\uparrow s} P\left(\tau, \rho\right) g\left(\rho, \cdot\right)(x) \\
&= P\left(\tau, s\right) g\left(s, \cdot\right)(x) = \mathbb{E}_{\tau,x}\left[g\left(s, X(s)\right)\right].
\end{aligned} \tag{3.19}$$

Consequently, the $\mathbb{P}_{\tau,x}$-martingale $\left(\mathbb{E}_{s_n,X(s_n)}\left[f\left(s, X(s)\right)\right]\right)_{n\in\mathbb{N}}$ converges $\mathbb{P}_{\tau,x}$-almost surely and in the space $L^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$ to the random variable $f\left(s, X(s)\right)$. In addition, the sequence $\left(f\left(s_n, X\left(s_n\right)\right)\right)_{n\in\mathbb{N}}$ converges in the space $L^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$ to the same random variable $f\left(s, X(s)\right)$.

Then there exists a subsequence $\left(f\left(s_{n_k}, X\left(s_{n_k}\right)\right)\right)_{n\in\mathbb{N}}$ which converges $\mathbb{P}_{\tau,x}$-almost surely to $f(s, X(s))$. Again we employ the function in (3.11) to prove that

$$\lim_{k\to\infty} X\left(s_{n_k}\right) = X(s), \quad \mathbb{P}_{\tau,x}\text{-almost surely.} \tag{3.20}$$

The equalities (3.15) and (3.20) show that the orbit $\{(\rho, X(\rho)) : \rho \in [\tau, T]\}$ is $\mathbb{P}_{\tau,x}$ almost surely a sequentially compact subset of E. Since the space E is complete metrizable we infer that this orbit is $\mathbb{P}_{\tau,x}$-almost surely a compact subset of E. We still have to show that there exists a modification $\left\{\widetilde{X}(s) : s \in [0,T]\right\}$ of the process $\{X(s) : s \in [0,T]\}$ which possesses left limits, is right-continuous $\mathbb{P}_{\tau,x}$-almost surely, and is such that

$$P(\tau, t) f(x) = \mathbb{E}_{\tau,x}\left[f(X(t))\right] = \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(t)\right)\right], \quad f \in C_b(E). \tag{3.21}$$

For the notion of modification see Definition 3.1. In order to achieve this we begin by using a modified version of the process in (3.5):

$$t \mapsto e^{-\alpha t} R(\alpha) f(t, X(t)) = \int_t^\infty e^{-\alpha\rho} P(t, \rho\wedge T) f(\rho\wedge T, \cdot)(X(t))\, d\rho, \tag{3.22}$$

for $t \in [0, T]$. The process in (3.22) is a $\mathbb{P}_{\tau,x}$-supermartingale with respect to the filtration $\left(\mathcal{F}_\rho^\tau\right)_{\tau\leqslant\rho\leqslant T}$. Since the process in (3.22) is a $\mathbb{P}_{\tau,x}$-supermartingale on the interval $[\tau, T]$ we deduce that for t varying over countable subsets its left and right limits exist $\mathbb{P}_{\tau,x}$-almost surely. Then the process in (3.5) shares this property as well. For a detailed argument which substantiates this claim see the propositions 3.3 and 3.4 below. Since the orbit $\{(\rho, X(\rho)) : \rho \in [\tau, T]\}$ is $\mathbb{P}_{\tau,x}$-almost surely relatively compact, and since the function f belongs to $C_b([0,T]\times E)$ we infer that for sequences the process $t \mapsto f(t, X(t))$ possesses $\mathbb{P}_{\tau,x}$-almost surely left and right limits in E. Again an appeal to the function f in (3.11) shows that the limits $\lim_{s\uparrow t,\, s\in D} X(s)$ and $\lim_{t\downarrow s,\, t\in D} X(t)$ exist $\mathbb{P}_{\tau,x}$-almost surely for $t \in (\tau, T]$ and $s \in [\tau, T]$. Here we wrote $D = \{k2^{-n} : k \in \mathbb{N}, n \in \mathbb{N}\}$ for the collection of non-negative dyadic numbers. A redefinition (modification) $\widetilde{X}(\rho)$ of the process $X(\rho)$, $\rho \in [0, T]$, reads as follows:

$$\widetilde{X}(\rho) = \lim_{t\downarrow\rho,\, t\in D\bigcap(\rho,T],\, t>\rho} X(t),\ \rho \in [0,T), \quad \widetilde{X}(T) = X(T). \tag{3.23}$$

The proof of Theorem 2.9 will be continued after inserting an important intermediate result, which we obtained thus far. □

This intermediate important result reads as follows.

Proposition 3.1. *The process* $\left\{\widetilde{X}(\rho) : \rho \in [0,T]\right\}$ *is continuous from the right and has left limits in* E $\mathbb{P}_{\tau,x}$*-almost surely. Moreover, its* $\mathbb{P}_{\tau,x}$*-distribution coincides with that of the process* $\{X(\rho) : \rho \in [0,T]\}$*. Fix* $(\tau, x) \in [0,T] \times E$ *and* $t \in [\tau, T]$*. On the event* $\left\{\widetilde{X}(t) \in E\right\}$ *the orbits* $\left\{\left(s, \widetilde{X}(s)\right) : s \in [\tau, t]\right\}$ *are* $\mathbb{P}_{\tau,x}$*-almost surely relatively compact subsets of* $[\tau, T] \times E$*.*

Fix $0 \leqslant \tau \leqslant t \leqslant T$, and let S, S_1 and S_2 be $\left(\widetilde{\mathcal{F}}_t^\tau\right)_{t\in[\tau,T]}$-stopping times. In what follows we will make use of the following σ-fields:

$$\begin{aligned}
\widetilde{\mathcal{F}}_t^\tau &= \sigma\left(\widetilde{X}(\rho) : \tau \leqslant \rho \leqslant t\right); \\
\widetilde{\mathcal{F}}_{t+}^\tau &= \bigcap_{0<\varepsilon\leqslant T-t} \sigma\left(\widetilde{X}(\rho) : \tau \leqslant \rho \leqslant t+\varepsilon\right) = \bigcap_{0<\varepsilon\leqslant T-t} \widetilde{\mathcal{F}}_{t+\varepsilon}^\tau; & (3.24) \\
\widetilde{\mathcal{F}}_T^{S,\vee} &= \sigma\left(\left(\rho \vee S, \widetilde{X}(\rho\vee S)\right) : 0 \leqslant \rho \leqslant T\right); & (3.25) \\
\widetilde{\mathcal{F}}_{S_2}^{S_1,\vee} &= \bigcap_{s\in[0,T]} \left\{A \in \widetilde{\mathcal{F}}_T^{S_1,\vee} : A \bigcap \{S_2 \leqslant s\} \in \widetilde{\mathcal{F}}_s^0\right\}; & (3.26) \\
\widetilde{\mathcal{F}}_{S_2+}^{S_1,\vee} &= \bigcap_{s\in[0,T]} \left\{A \in \widetilde{\mathcal{F}}_T^{S_1,\vee} : A \bigcap \{S_2 < s\} \in \widetilde{\mathcal{F}}_s^0\right\} \\
&= \bigcap_{0<\varepsilon\leqslant T} \bigcap_{s\in[0,T-\varepsilon]} \left\{A \in \widetilde{\mathcal{F}}_T^{S_1,\vee} : A \bigcap \{S_2 \leqslant s\} \in \widetilde{\mathcal{F}}_{s+\varepsilon}^0\right\} \\
&= \bigcap_{\varepsilon>0} \widetilde{\mathcal{F}}_{(S_2+\varepsilon)\wedge T}^{S_1,\vee}. & (3.27)
\end{aligned}$$

The σ-field in (3.24) is called the right closure of $\widetilde{\mathcal{F}}_t^\tau$, the σ-field in (3.25) is called the σ-field after time S, the σ-field in (3.26) is called the σ-field between time S_1 and S_2, and finally the one in (3.27) is called the right closure of the one in (3.26).

Proof. [Continuation of the proof Theorem 2.9.] Our most important aim is to prove that the process

$$\left\{\left(\Omega, \widetilde{\mathcal{F}}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(\widetilde{X}(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), (E, \mathcal{E})\right\} \tag{3.28}$$

is a strong Markov process. We begin by proving the following $\mathbb{P}_{\tau,x}$-almost sure equalities:

$$\begin{aligned}
\mathbb{E}_{s,\widetilde{X}(s)}\left[F \circ \vee_s\right] &= \mathbb{E}_{\tau,x}\left[F \circ \vee_s \mid \mathcal{F}_s^\tau\right] & (3.29) \\
&= \mathbb{E}_{\tau,x}\left[F \circ \vee_s \mid \mathcal{F}_{s+}^\tau\right] = \mathbb{E}_{\tau,x}\left[F \circ \vee_s \mid \widetilde{\mathcal{F}}_{s+}^\tau\right]. & (3.30)
\end{aligned}$$

First we take F of the form $F = f\left(\widetilde{X}(s)\right)$ where $f \in C_b(E)$. By an approximation argument it then follows that (3.29) and (3.30) also hold for $F = f\left(\widetilde{X}(s)\right)$ with $f \in L^{\infty}(E, \mathcal{E})$. So let $f \in C_b(E)$. Since $\mathbb{P}_{s,y}\left[\widetilde{X}(s) = y\right] = 1$ and $f\left(\widetilde{X}(s)\right) \circ \vee_s = f\left(\widetilde{X}(s)\right)$ we see

$$\mathbb{E}_{s,\widetilde{X}(s)}\left[f\left(\widetilde{X}(s)\right) \circ \vee_s\right] = \mathbb{E}_{s,\widetilde{X}(s)}\left[f\left(\widetilde{X}(s)\right)\right] = f\left(\widetilde{X}(s)\right). \tag{3.31}$$

Since the random variable $f\left(\widetilde{X}(s)\right)$ is measurable with respect to the σ-field $\mathcal{F}_{s+}^{\tau}$ by (3.31) we also have the $\mathbb{P}_{\tau,x}$-almost sure equalities:

$$\mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(s)\right) \circ \vee_s \;\middle|\; \mathcal{F}_{s+}^{\tau}\right] = \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(s)\right) \;\middle|\; \mathcal{F}_{s+}^{\tau}\right] = f\left(\widetilde{X}(s)\right). \tag{3.32}$$

Next we calculate, while using the Markov property of the process $t \mapsto X(t)$ and right-continuity of the function $t \mapsto P(s,t) f(y)$, $s \in [\tau, T]$, $y \in E$,

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(s)\right) \circ \vee_s \;\middle|\; \mathcal{F}_{s+}^{\tau}\right] = \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(s)\right) \;\middle|\; \mathcal{F}_{s}^{\tau}\right] \\
&= \lim_{\varepsilon \downarrow 0} \mathbb{E}_{\tau,x}\left[f\left(X(s+\varepsilon)\right) \;\middle|\; \mathcal{F}_{s}^{\tau}\right] = \lim_{\varepsilon \downarrow 0} \mathbb{E}_{s,X(s)}\left[f\left(X(s+\varepsilon)\right)\right] \\
&= \lim_{\varepsilon \downarrow 0} P(s, s+\varepsilon) f(\cdot))\left(X(s)\right) = f\left(X(s)\right).
\end{aligned} \tag{3.33}$$

In order to complete the arguments for the proof of (3.29) and (3.30) for F of the form $F = f\left(\widetilde{X}(s)\right)$, $f \in C_b(E)$, we have to show the equality $f\left(\widetilde{X}(s)\right) = f\left(X(s)\right)$ $\mathbb{P}_{\tau,x}$-almost surely. This will be accomplished by the following identities:

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\left|f\left(\widetilde{X}(s)\right) - f\left(X(s)\right)\right|^2\right] \\
&= \lim_{t \downarrow s} \mathbb{E}_{\tau,x}\left[|f(X(t))|^2\right] - 2 \lim_{t \downarrow s} \Re \mathbb{E}_{\tau,x}\left[\overline{f(X(s))} f(X(t))\right] + \mathbb{E}_{\tau,x}\left[|f(X(s))|^2\right]
\end{aligned}$$

(Markov property for the process $t \mapsto X(t)$)

$$\begin{aligned}
&= \lim_{t \downarrow s} \mathbb{E}_{\tau,x}\left[\left|f\left(X(t)\right)\right|^2\right] - 2 \lim_{t \downarrow s} \Re \mathbb{E}_{\tau,x}\left[\overline{f\left(X(s)\right)} \mathbb{E}_{s,X(s)}\left[f\left(X(t)\right)\right]\right] \\
&\qquad + \mathbb{E}_{\tau,x}\left[\left|f\left(X(s)\right)\right|^2\right]
\end{aligned}$$

(relationship between Feller propagator and Markov property of X)

$$\begin{aligned}
&= \lim_{t \downarrow s} P(\tau, t) \left|f(\cdot)\right|^2 (x) - 2 \lim_{t \downarrow s} \Re\left[P(\tau, s) \overline{f(\cdot)} P(s,t) f(\cdot)\right](x) \\
&\qquad + P(\tau, s) \left|f(\cdot)\right|^2 (x) \\
&= P(\tau, s) \left|f(\cdot)\right|^2 (x) - 2\left[P(\tau, s) \overline{f(\cdot)} f(\cdot)\right](x) + P(\tau, s) \left|f(\cdot)\right|^2 (x) \\
&= 0.
\end{aligned} \tag{3.34}$$

From (3.34) we infer that $f\left(\widetilde{X}(s)\right) = f\left(X(s)\right)$ $\mathbb{P}_{\tau,x}$-almost surely. From (3.32), (3.33), and (3.34) we deduce the equalities in (3.29) and (3.30) for a variable F of the form $F = f\left(\widetilde{X}(s)\right)$, $f \in C_b(E)$. An approximation arguments then yields (3.29) and (3.30) for $f \in L^{\infty}\left(E, \mathcal{E}\right)$.

In order to prove (3.29) in full generality it suffices by the Monotone Class Theorem and an approximation argument to prove the equalities in (3.30) for random variables F of the form $F = \prod_{j=0}^{n} f_j\left(\widetilde{X}\left(s_j\right)\right)$, where the functions f_j, $0 \leqslant j \leqslant n$, belong to $C_b(E)$ and where $s = s_0 < s_1 < s_2 < \cdots < s_n \leqslant T$. Since the equality $f\left(\widetilde{X}(s)\right) = f\left(X(s)\right)$ holds $\mathbb{P}_{\tau,x}$-almost surely, it is easy to see that by using the equalities (3.32), (3.33), and (3.34), it suffices to take the variable F of the form $F = \prod_{j=1}^{n+1} f_j\left(\widetilde{X}\left(s_j\right)\right)$ where as above the functions f_j, $1 \leqslant j \leqslant n+1$, belong to $C_b(E)$ and where $s < s_1 < s_2 < \cdots < s_n < s_{n+1} \leqslant T$. For $n = 0$ we have $\mathbb{P}_{\tau,x}$-almost surely

$$\mathbb{E}_{\tau,x}\left[f_1\left(\widetilde{X}\left(s_1\right)\right) \bigm| \mathcal{F}_s^{\tau}\right] = \lim_{\varepsilon\downarrow 0} \mathbb{E}_{\tau,x}\left[f_1\left(X\left(s_1+\varepsilon\right)\right) \bigm| \mathcal{F}_s^{\tau}\right]$$

(Markov property of the process X)

$$\begin{aligned}
&= \lim_{\varepsilon\downarrow 0} \mathbb{E}_{s,X(s)}\left[f_1\left(X\left(s_1+\varepsilon\right)\right)\right] = \lim_{\varepsilon\downarrow 0} P\left(s, s_1+\varepsilon\right) f\left(X(s)\right) \\
&= \lim_{\varepsilon\downarrow 0} P\left(s, s_1\right) P\left(s_1, s_1+\varepsilon\right) f\left(X(s)\right) = P\left(s, s_1\right) f\left(X(s)\right) \\
&= P\left(s, s_1\right) f\left(\widetilde{X}(s)\right) = \mathbb{E}_{s,\widetilde{X}(s)}\left[f\left(\widetilde{X}\left(s_1\right)\right)\right].
\end{aligned} \tag{3.35}$$

The equalities in (3.35) imply (3.29) with $F = f_1\left(\widetilde{X}\left(s_1\right)\right)$ where $f_1 \in C_b(E)$ and $s < s_1 \leqslant T$. Then we apply induction with respect to n to obtain (3.29) for F of the form $F = \prod_{j=1}^{n+1} f_j\left(\widetilde{X}\left(s_j\right)\right)$ where as above the functions f_j, $1 \leqslant j \leqslant n+1$, belong to $C_b(E)$ and where $s < s_1 < s_2 < \cdots < s_n < s_{n+1} \leqslant T$. In fact using the measurability of $\widetilde{X}\left(s_j\right)$ with respect to the σ-field $\mathcal{F}_{s_n+}^{\tau}$, $1 \leqslant j \leqslant n$, and the tower property of conditional expectation we get $\mathbb{P}_{\tau,x}$-almost surely:

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\prod_{j=1}^{n+1} f_j\left(\widetilde{X}\left(s_j\right)\right) \bigm| \mathcal{F}_s^{\tau}\right] \\
&= \mathbb{E}_{\tau,x}\left[\prod_{j=1}^{n} f_j\left(\widetilde{X}\left(s_j\right)\right) \mathbb{E}_{\tau,x}\left[f_{n+1}\left(\widetilde{X}\left(s_{n+1}\right)\right) \bigm| \mathcal{F}_{s_n}^{\tau}\right] \mathcal{F}_s^{\tau}\right]
\end{aligned}$$

(Markov property for $n=1$)

$$= \mathbb{E}_{\tau,x}\left[\prod_{j=1}^{n} f_j\left(\widetilde{X}(s_j)\right)\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f_{n+1}\left(\widetilde{X}(s_{n+1})\right)\right]\mathcal{F}_s^{\tau}\right]$$

(induction hypothesis)

$$\begin{aligned}
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\prod_{j=1}^{n} f_j\left(\widetilde{X}(s_j)\right)\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f_{n+1}\left(\widetilde{X}(s_{n+1})\right)\right]\right]\\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\prod_{j=1}^{n} f_j\left(\widetilde{X}(s_j)\right)\mathbb{E}_{s,\widetilde{X}(s)}\left[f_{n+1}\left(\widetilde{X}(s_{n+1})\right)\,\Big|\,\mathcal{F}_{s_n}^{s}\right]\right]\\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\prod_{j=1}^{n} f_j\left(\widetilde{X}(s_j)\right) f_{n+1}\left(\widetilde{X}(s_{n+1})\right)\right]\\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\prod_{j=1}^{n+1} f_j\left(\widetilde{X}(s_j)\right)\right].
\end{aligned} \tag{3.36}$$

So that (3.36) proves (3.29) for $F=\prod_{j=1}^{n+1} f_j\left(\widetilde{X}(s_j)\right)$ where the functions f_j, $1\leqslant n+1$, belong to $C_b(E)$, and $s<s_1<\cdots<s_{n+1}$. As remarked above from (3.32), (3.33), and (3.34) the equality in (3.29) then also follows for all random variables of the form $F=\prod_{j=0}^{n} f_j\left(\widetilde{X}(s_j)\right)$ with $f_j\in C_b(E)$ for $0\leqslant j\leqslant n$ and $0=s_0<s_1<\cdots<s_n\leqslant T$. By the Monotone Class Theorem and approximation arguments it then follows that (3.29) is true for all bounded $\mathcal{F}_T^{\tau}$ random variables F.

Next we proceed with a proof of the equalities in (3.30). Since $\widetilde{\mathcal{F}}_{s+}^{\tau}\subset\mathcal{F}_{s+}^{\tau}$, and the variable $\mathbb{E}_{s,\widetilde{X}(s)}\left[F\circ\vee_s\right]$ is $\mathcal{F}_{s+}^{\tau}$-measurable, it suffices to prove the first equality in (3.30), to wit

$$\mathbb{E}_{\tau,x}\left[F\circ\vee_s\,\Big|\,\mathcal{F}_{s+}^{\tau}\right]=\mathbb{E}_{s,\widetilde{X}(s)}\left[F\circ\vee_s\right] \tag{3.37}$$

for any bounded $\mathcal{F}_T^{\tau}$-measurable random variable F. We will not prove the equality in (3.37) directly, but we will show the following ones instead:

$$\mathbb{E}_{\tau,x}\left[F\circ\vee_s\,\Big|\,\mathcal{F}_{s+}^{\tau}\right]=\mathbb{E}_{s,\widetilde{X}(s)}\left[F\circ\vee_s\,\Big|\,\mathcal{F}_{s+}^{s}\right]=\mathbb{E}_{s,\widetilde{X}(s)}\left[F\circ\vee_s\,\Big|\,\widetilde{\mathcal{F}}_{s+}^{s}\right], \tag{3.38}$$

under the condition that the function $(s,x)\mapsto P(s,t)f(x)$ is Borel measurable on $[\tau,t]\times E$ for $f\in C_b(E)$, which is part of (vi) in Definition 2.4. In order to prove the equalities in (3.38) it suffices by the Monotone Class Theorem to take F of the form $F=\prod_{j=0}^{n} f_j\left(\widetilde{X}(s_j)\right)$ with

$s = s_0 < s_1 < \cdots < s_n \leqslant T$ and where de functions f_j, $0 \leqslant j \leqslant n$, are bounded Borel measurable functions. By another approximation argument we may assume that the functions f_j, $0 \leqslant j \leqslant n$, belong to $C_b(E)$. An induction argument shows that it suffices to prove (3.38) for $F = f_0\left(\widetilde{X}\left(s_0\right)\right) f_1\left(\widetilde{X}\left(s_1\right)\right)$ where $s = s_0 < s_1 \leqslant T$, and the functions f_0 and f_1 are members of $C_b(E)$. The case $f_1 = \mathbf{1}$ was taken care of in the equalities (3.31) and (3.32). Since the variable $f_0\left(\widetilde{X}(s)\right)$ is $\mathcal{F}_{s+}^{s}$-measurable the proof of the equalities in (3.38) reduces to the case where $F = f\left(\widetilde{X}(t)\right)$ where $\tau < s < t \leqslant T$ and $f \in C_b(E)$. The following equalities show the first equality in (3.38). With $s < s_{n+1} < s_n < t$ and $\lim_{n\to\infty} s_n = s$ we have

$$\begin{aligned}
\mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(t)\right) \mid \mathcal{F}_{s+}^{\tau}\right] &= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(t)\right) \mid \mathcal{F}_{s_n}^{\tau}\right] \mid \mathcal{F}_{s+}^{\tau}\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s_n,X(s_n)}\left[f\left(\widetilde{X}(t)\right)\right] \mid \mathcal{F}_{s+}^{\tau}\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f\left(\widetilde{X}(t)\right)\right] \mid \mathcal{F}_{s+}^{\tau}\right] \\
&= \mathbb{E}_{\tau,x}\left[\lim_{n\to\infty}\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f\left(\widetilde{X}(t)\right)\right] \mid \mathcal{F}_{s+}^{\tau}\right] \\
&= \lim_{n\to\infty}\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f\left(\widetilde{X}(t)\right)\right] \qquad (3.39)\\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\lim_{n\to\infty}\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f\left(\widetilde{X}(t)\right)\right] \mid \mathcal{F}_{s+}^{s}\right] \\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\mathbb{E}_{s_n,\widetilde{X}(s_n)}\left[f\left(\widetilde{X}(t)\right)\right] \mid \mathcal{F}_{s+}^{s}\right] \\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[\mathbb{E}_{s,\widetilde{X}(s)}\left[f\left(\widetilde{X}(t)\right) \mid \mathcal{F}_{s_n}^{s}\right] \mid \mathcal{F}_{s+}^{s}\right] \\
&= \mathbb{E}_{s,\widetilde{X}(s)}\left[f\left(\widetilde{X}(t)\right) \mid \mathcal{F}_{s+}^{s}\right]. \qquad (3.40)
\end{aligned}$$

In these equalities we used the fact that the process $\rho \mapsto \mathbb{E}_{\rho,\widetilde{X}(\rho)}\left[f\left(\widetilde{X}(t)\right)\right]$, $s < \rho \leqslant t$ is $\mathbb{P}_{s,y}$-martingale for $(s,y) \in [0,t) \times E$. The equality in (3.40) implies the first equality in (3.38). The second one can be obtained by repeating the four final steps in the proof of (3.40) with $\widetilde{\mathcal{F}}_{s+}^{s}$ instead of $\mathcal{F}_{s+}^{s}$. Here we use that the random variable in (3.39) is measurable with respect to the σ-field $\widetilde{\mathcal{F}}_{s+}^{s}$, which is smaller than $\mathcal{F}_{s+}^{s}$.

In order to deduce (3.37) from (3.38) we will need the full strength of property (vi) in Definition 2.4. In fact using the representation in (3.39) and using the continuity property in (vi) shows (3.37) for $F = f\left(\widetilde{X}(t)\right)$, $f \in C_b(E)$. By the previous arguments the full assertion in (3.30) follows. In fact Proposition 3.3 gives a detailed proof of the equalities in (3.74) below. The equalities in (3.39) then follow from the Monotone Class Theorem.

Next we want to prove that the process $t \mapsto \widetilde{X}(t)$ possesses the strong Markov property. This means that for any given $\left(\widetilde{\mathcal{F}}_{t+}^{\tau}\right)_{t\in[\tau,T]}$-stopping time $S : \Omega \to [\tau, T]$ we have to prove an equality of the form (see (2.93))

$$\mathbb{E}_{S,\widetilde{X}(S)}\left[F \circ \vee_S\right] = \mathbb{E}_{\tau,x}\left[F \circ \vee_S \;\middle|\; \widetilde{\mathcal{F}}_{S+}^{\tau}\right], \tag{3.41}$$

and this for all bounded $\mathcal{F}_T^{\tau}$-measurable random variables F. By the Monotone Class Theorem it follows that it suffices to prove (3.41) for bounded random variables F of the form $F = \prod_{j=0}^{n} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right)$ where the functions f_j, $0 \leqslant j \leqslant n$, are bounded Borel functions on $[\tau, T] \times E$, and $\tau = s_0 < s_1 < \cdots < s_n \leqslant T$. By another approximation argument it suffices to replace the bounded Borel functions f_j, $0 \leqslant j \leqslant n$, by bounded continuous functions on $[\tau, T] \times E$. By definition the stopping time S is $\widetilde{\mathcal{F}}_{S+}^{\tau}$-measurable. Let us show that $\widetilde{X}(S)$ is $\mathcal{F}_{S+}^{\tau}$-measurable. Therefore we approximate the stopping time S from above by stopping times S_n, $n \in \mathbb{N}$, of the form

$$S_n = \tau + \frac{T-\tau}{2^n}\left\lceil \frac{2^n\left(S-\tau\right)}{T-\tau}\right\rceil. \tag{3.42}$$

If $t \in [\tau, T]$, then

$$\{S_n \leqslant t\} = \bigcup_{k=0}^{\left\lfloor 2^n \frac{t-\tau}{T-\tau}\right\rfloor} \left\{\frac{k-1}{2^n}(T-\tau) + \tau < S \leqslant \frac{k}{2^n}(T-\tau) + \tau\right\}, \tag{3.43}$$

and hence S_n is $\left(\mathcal{F}_{t+}^{\tau}\right)_{t\in[\tau,T]}$-stopping time. Moreover, on the event

$$\left\{\frac{k-1}{2^n}(T-\tau) + \tau < S \leqslant \frac{k}{2^n}(T-\tau) + \tau\right\}$$

the stopping time S_n takes the value $S_n = t_{k,n}$, where $t_{k,n} = \tau + \dfrac{k(T-\tau)}{2^n}$. Consequently, we have the following equality of events:

$$\left\{S_n = \tau + \frac{k(T-\tau)}{2^n} = t_{k,n}\right\} = \left\{\frac{k-1}{2^n}(T-\tau) + \tau < S \leqslant \frac{k}{2^n}(T-\tau) + \tau\right\},$$

so that for $k \leqslant \dfrac{2^n(t-\tau)}{T-\tau}$, which is equivalent to $t_{k,n} \leqslant t$, the event $\left\{S_n = \tau + \dfrac{k(T-\tau)}{2^n}\right\}$ is $\widetilde{\mathcal{F}}_{t+}^{\tau}$-measurable, and on this event the state variable $\widetilde{X}\left(S_n\right) = \widetilde{X}\left(t_{k,n}\right)$ is $\widetilde{\mathcal{F}}_{t_{k,n}+}^{\tau}$-measurable. As a consequence we see that on the event $\{S_n \leqslant t\}$ the state variable $\widetilde{X}(S_n)$ is $\widetilde{\mathcal{F}}^{\tau}_{t+}$-measurable. Then the space-time variable $\left(S_n, \widetilde{X}(S_n)\right)$ is measurable with respect to the σ-field $\widetilde{\mathcal{F}}^{\tau}_{S+}$. In addition, we have

$$S \leqslant S_{n+1} \leqslant S_n \leqslant S + \frac{T-\tau}{2^n}, \tag{3.44}$$

and hence the space-time variable $\left(S, \widetilde{X}(S)\right)$ is $\widetilde{\mathcal{F}}^{\tau}_{S+}$-measurable as well. This proves the equality in (3.41) in case $F = f\left(\tau \vee S, \widetilde{X}(\tau \vee S)\right)$ where $f \in C_b([\tau, T] \times E)$. As a preparation for the case $F = \prod_{j=0}^{n} f_j\left(s_j \vee S, \widetilde{X}(s_j \vee S)\right)$ where the functions f_j, $0 \leqslant j \leqslant n$, are bounded Borel functions on $[\tau, T] \times E$, and $\tau = s_0 < s_1 < \cdots < s_n \leqslant T$, we first consider the case $(\tau < t \leqslant T)$

$$F = f\left(t \vee S, \widetilde{X}(t \vee S)\right) \mathbf{1}_{\{S \leqslant t\}} = f\left(t, \widetilde{X}(t)\right) \mathbf{1}_{\{S \leqslant t\}} \tag{3.45}$$

where $f \in C_b([\tau, T] \times E)$. On the event $\{S \leqslant t\}$ we approximate the stopping time S from above by stopping times S_n, $n \in \mathbb{N}$, of the form

$$S_n(t) = \tau + \frac{t-\tau}{2^n}\left\lceil \frac{2^n(S-\tau)}{t-\tau} \right\rceil. \tag{3.46}$$

Then on the event $\{S \leqslant t\}$ we have the following inclusions of σ-fields:

$$\begin{aligned} &\widetilde{\mathcal{F}}^{\tau}_{S+} \bigcap \{S \leqslant t\} \\ &= \widetilde{\mathcal{F}}^{\tau}_{S \wedge t+} \bigcap \{S \leqslant t\} \subset \widetilde{\mathcal{F}}^{\tau}_{S_{n+1}(t)+} \bigcap \{S \leqslant t\} \subset \widetilde{\mathcal{F}}^{\tau}_{S_n(t)+} \bigcap \{S \leqslant t\} \end{aligned} \tag{3.47}$$

and

$$\bigcap_{n=1}^{\infty} \widetilde{\mathcal{F}}^{\tau}_{S_n(t)+} \bigcap \{S \leqslant t\} = \widetilde{\mathcal{F}}^{\tau}_{S+} \bigcap \{S \leqslant t\}. \tag{3.48}$$

Here we wrote $\mathcal{F} \bigcap A_0 = \{A \bigcap A_0 : A \in \mathcal{F}\}$ when $\mathcal{F}$ is any σ-field on Ω and $A_0 \subset \Omega$. Then we have

$$\begin{aligned} &\mathbb{E}_{\tau,x}\left[f\left(t \vee S, \widetilde{X}(t \vee S)\right) \mathbf{1}_{\{S \leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right] \\ &= \mathbb{E}_{\tau,x}\left[f\left(t, \widetilde{X}(t)\right) \mathbf{1}_{\{S \leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right] \\ &= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{\tau,x}\left[f\left(t, \widetilde{X}(t)\right) \mathbf{1}_{\{S \leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S_n(t)+}\right] \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right] \\ &= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{S_n(t), \widetilde{X}(S_n(t))}\left[f\left(t, \widetilde{X}(t)\right)\right] \mathbf{1}_{\{S \leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right] \\ &= \lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[\mathbb{E}_{S_n(t), \widetilde{X}(S_n(t))}\left[f\left(t, \widetilde{X}(t)\right)\right] \mathbf{1}_{\{S \leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right] \end{aligned}$$

$$= \mathbb{E}_{\tau,x}\left[\lim_{n\to\infty}\mathbb{E}_{S_n(t),\widetilde{X}(S_n(t))}\left[f\left(t,\widetilde{X}(t)\right)\right]\mathbf{1}_{\{S\leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right]$$

(employ (3.48) and the arguments leading to equality (3.38))

$$= \lim_{n\to\infty}\mathbb{E}_{S_n(t),\widetilde{X}(S_n(t))}\left[f\left(t,\widetilde{X}(t)\right)\right]\mathbf{1}_{\{S\leqslant t\}}$$
$$= \mathbb{E}_{S,\widetilde{X}(S)}\left[f\left(t,\widetilde{X}(t)\right) \mid \widetilde{\mathcal{F}}^{S,\vee}_{S+}\right]\mathbf{1}_{\{S\leqslant t\}}$$

(appeal to (3.37) which relies on property (vi) of Definition 2.4)

$$= \mathbb{E}_{S,\widetilde{X}(S)}\left[f\left(t,\widetilde{X}(t)\right)\right]\mathbf{1}_{\{S\leqslant t\}}. \tag{3.49}$$

From (3.49) and the $\widetilde{\mathcal{F}}^{S,\vee}_{S+}$-measurability of the stochastic state variable $\left(S,\widetilde{X}(S)\right)$ we infer

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[f\left(t\vee S,\widetilde{X}(t\vee S)\right) \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right]\\
&= \mathbb{E}_{\tau,x}\left[f\left(t\vee S,\widetilde{X}(t\vee S)\right)\mathbf{1}_{\{S\leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right]\\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(t\vee S,\widetilde{X}(t\vee S)\right)\mathbf{1}_{\{S>t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right]\\
&= \mathbb{E}_{\tau,x}\left[f\left(t,\widetilde{X}(t)\right)\mathbf{1}_{\{S\leqslant t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right]\\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(S,\widetilde{X}(S)\right)\mathbf{1}_{\{S>t\}} \mid \widetilde{\mathcal{F}}^{\tau}_{S+}\right]\\
&= \mathbb{E}_{S,\widetilde{X}(S)}\left[f\left(t,\widetilde{X}(t)\right)\mathbf{1}_{\{S\leqslant t\}}\right] + f\left(S,\widetilde{X}(S)\right)\mathbf{1}_{\{S>t\}}\\
&= \mathbb{E}_{S,\widetilde{X}(S)}\left[f\left(t\vee S,\widetilde{X}(t\vee S)\right)\right].
\end{aligned} \tag{3.50}$$

Next we consider the case $F = \prod_{j=0}^{n+1} f_j\left(s_j\vee S,\widetilde{X}(s_j\vee S)\right)$ where the functions f_j, $0\leqslant j\leqslant n+1$, are bounded Borel functions on $[\tau,T]\times E$, and $\tau = s_0 < s_1 < \cdots < s_{n+1} \leqslant T$. From (3.50) and the $\widetilde{\mathcal{F}}^{S,\vee}_{S+}$-measurability of the stochastic state variable $\left(S,\widetilde{X}(S)\right)$ we obtain (3.41) in case $F = f_0\left(\tau,\widetilde{X}(\tau)\right) f_1\left(s_1,\widetilde{X}(s_1)\right)$, and thus

$$F\circ\vee_S = f_0\left(\tau\vee S,\widetilde{X}(\tau\vee S)\right) f_1\left(s_1\vee S,\widetilde{X}(s_1\vee S)\right).$$

So that the cases $n=0$ and $n=1$ have been taken care of. The remaining part of the proof uses induction. From (3.50) with the maximum operator $s_n\vee S$ replacing S together with the induction hypothesis we get

$$\mathbb{E}_{\tau,x}\left[\prod_{j=0}^{n+1} f_j\left(s_j\vee S,\widetilde{X}(s_j\vee S)\right) \mid \mathcal{F}^{\tau}_{S+}\right]$$

$$
\begin{aligned}
&= \mathbb{E}_{\tau,x}\left[\prod_{j=0}^{n} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right)\right. \\
&\qquad \left. \times\, \mathbb{E}_{\tau,x}\left[f_{n+1}\left(s_{n+1} \vee S, \widetilde{X}\left(s_{n+1} \vee S\right)\right) \mid \mathcal{F}^{\tau}_{s_n \vee S+}\right] \mid \mathcal{F}^{\tau}_{S+}\right] \\
&= \mathbb{E}_{\tau,x}\left[\prod_{j=0}^{n} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right)\right. \\
&\qquad \left. \times\, \mathbb{E}_{s_n \vee S, \widetilde{X}(s_n \vee S)}\left[f_{n+1}\left(s_{n+1} \vee S, \widetilde{X}\left(s_{n+1} \vee S\right)\right)\right] \mid \mathcal{F}^{\tau}_{S+}\right]
\end{aligned}
$$

(induction hypothesis)

$$
\begin{aligned}
&= \mathbb{E}_{S,\widetilde{X}(S)}\left[\prod_{j=0}^{n} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right)\right. \\
&\qquad \left. \times\, \mathbb{E}_{s_n \vee S, \widetilde{X}(s_n \vee S)}\left[f_{n+1}\left(s_{n+1} \vee S, \widetilde{X}\left(s_{n+1} \vee S\right)\right)\right]\right] \\
&= \mathbb{E}_{S,\widetilde{X}(S)}\left[\prod_{j=0}^{n} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right)\right. \\
&\qquad \left. \times\, \mathbb{E}_{S,\widetilde{X}(S)}\left[f_{n+1}\left(s_{n+1} \vee S, \widetilde{X}\left(s_{n+1} \vee S\right)\right) \mid \mathcal{F}^{S,\vee}_{s_n \vee S+}\right]\right] \\
&= \mathbb{E}_{S,\widetilde{X}(S)}\left[\mathbb{E}_{S,\widetilde{X}(S)}\left[\prod_{j=0}^{n+1} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right) \mid \mathcal{F}^{S,\vee}_{s_n \vee S+}\right]\right] \\
&= \mathbb{E}_{S,\widetilde{X}(S)}\left[\prod_{j=0}^{n+1} f_j\left(s_j \vee S, \widetilde{X}\left(s_j \vee S\right)\right)\right]. \qquad (3.51)
\end{aligned}
$$

The strong Markov property of the process $\widetilde{X}$ follows from (3.51), an approximation argument and the Monotone Class Theorem.

We still need to redefine our process and probability measures $\mathbb{P}_{\tau,x}$ on the Skorohod space $D([0,T],E)$, $(\tau,x) \in [0,T] \times E$ in such a way that the distribution of the process $\widetilde{X}$ is preserved. This can be done replacing (3.28) with the collection

$$
\left\{\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}^{\tau}_{T}, \widetilde{\mathbb{P}}_{\tau,x}\right), \left(\widetilde{X}(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), (E,\mathcal{E})\right\} \qquad (3.52)
$$

where $\widetilde{\Omega} = D([0,T],E)$, and $\widetilde{\mathbb{P}}_{\tau,x}$ is determined by the equality $\widetilde{\mathbb{E}}_{\tau,x}[F] = \mathbb{E}_{\tau,x}[F \circ \pi]$. Here $F : \widetilde{\Omega} \to \mathbb{C}$ is a bounded variable which is measurable with respect to the σ-field generated by the coordinate variables: $\widetilde{X}(t) : \widetilde{\omega} \mapsto \widetilde{\omega}(t)$, $t \in [\tau, T]$, $\widetilde{\omega} \in \widetilde{\Omega}$. Recall that $\Omega = E^{[0,T]}$. Notice that the restriction of $\widetilde{X}(t)$ to $\widetilde{\Omega}$ is evaluation of $\widetilde{\omega} \in \widetilde{\Omega}$ at t. The mapping $\pi : \Omega \to \widetilde{\Omega}$ is defined by $\pi(\omega)(t) = \widetilde{X}(t,\omega)$, $t \in [0,T]$, $\omega \in \Omega'$. Here $\Omega' \subset \Omega$ has the property that for all $(\tau, x) \in [0,T] \times E$ its complement in Ω is $\mathbb{P}_{\tau,x}$-negligible. We will describe the space Ω'. Let D be the collection of positive dyadic numbers. For Ω' we may choose the space:

$$\begin{aligned}\Omega' := &\left\{\omega \in \Omega : t \mapsto \omega(t),\, t \in D \bigcap [0,T] \text{ has left and right limits in } E\right\} \\ &\bigcap \left\{\omega \in \Omega : \text{the range } \{\omega(t) : t \in D \bigcap [0,T]\} \text{ is totally bounded in } E\right\}.\end{aligned} \tag{3.53}$$

Let $(x_j)_{j\in\mathbb{N}}$ be a sequence in E which is dense, and let d be a metric on $E \times E$ which turns E into a Polish space. Put $B(x,\varepsilon) = \{y \in E : d(y,x) < \varepsilon\}$. Define, for any finite subset of $[0,T]$ with an even number of members $U = \{t_1, \ldots, t_{2n}\}$ say, and $\varepsilon > 0$, the random variable $H_\varepsilon(U)$ by

$$H_\varepsilon(U)(\omega) = \sum_{j=1}^{n} \mathbf{1}_{\left\{d\left(\widetilde{X}(t_{2j-1}), \widetilde{X}(t_{2j})\right) \geq \varepsilon\right\}}(\omega).$$

We also put

$$\begin{aligned}&H_\varepsilon\left(D \bigcap [0,T]\right) \\ &= \sup\left\{H_\varepsilon(U) : U \subset D \bigcap [0,T],\, U \text{ contains an even number of elements}\right\}.\end{aligned}$$

Then the subset Ω' of $\Omega = E^{[0,T]}$ can be described as follows:

$$\begin{aligned}\Omega' = &\bigcap_{n=1}^{\infty} \left\{\omega \in \Omega : H_{1/n}\left(D \bigcap [0,T]\right)(\omega) < \infty\right\} \\ &\bigcap \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \left\{\omega \in \Omega : \left(\widetilde{X}(s)(\omega)\right)_{s \in D \bigcap [0,T]} \subset \bigcup_{j=1}^{n} B(x_j, 1/m)\right\}.\end{aligned} \tag{3.54}$$

The description in (3.54) shows that the subset Ω' is a measurable subset of Ω. In addition we have $\mathbb{P}_{\tau,x}(\Omega') := \mathbb{P}_{\tau,x}(\Omega'_\tau) = 1$ for all $(\tau, x) \in [0,T] \times E$. Here

$$\Omega'_\tau = \left\{\omega \in \Omega' : \omega(\rho) = \omega(\tau),\, \rho \in D \bigcap [0,\tau]\right\}, \tag{3.55}$$

which may be identified with $\left\{\omega\restriction_{[\tau,T]} : \omega \in \Omega'\right\}$ which is a measurable subset of $\Omega_\tau = E^{[\tau,T]}$. In order to complete the construction and the proof of Theorem 2.9 we need to prove the quasi-left continuity of the process $\widetilde{X}$.

So let $(\tau_n)_{n\in\mathbb{N}}$ be an increasing sequence of $\left(\widetilde{\mathcal{F}}_t^\tau\right)_{t\in[\tau,T]}$-stopping times with values in $[\tau, T]$. Put $\tau_\infty = \sup_{n\in\mathbb{N}} \tau_n$. Let f and g be functions $C_b^+(E)$, and let $h > 0$. Then by the strong Markov property we have for $m \leqslant n$

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right] \\
&= \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right]\right] \\
&= \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h \geqslant \tau_\infty\right] \\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h < \tau_\infty\right]
\end{aligned}$$

(the process $\rho \mapsto \mathbb{E}_{\rho,\widetilde{X}(\rho)}\left[g\left(\widetilde{X}(s)\right)\right]$ is a right-continuous $\mathbb{P}_{\tau,x}$-martingale on $[\tau, s]$)

$$\begin{aligned}
&= \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h \geqslant \tau_\infty\right] \\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h < \tau_\infty\right] \\
&= \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) P(\tau_n, (\tau_m+h)\wedge T)\, g\left(\widetilde{X}(\tau_n)\right), \tau_m+h \geqslant \tau_\infty\right] \\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h < \tau_\infty\right].
\end{aligned} \tag{3.56}$$

Put $L = \lim_{n\to\infty} \widetilde{X}(\tau_n)$. Upon taking limits, as $n \to \infty$, and employing the fact that the propagator $P(\tau, t)$ is continuous from the left on the diagonal in (3.56) we obtain:

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right] \\
&= \lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) P(\tau_n, \tau_\infty)\, P(\tau_\infty, (\tau_m+h)\wedge T)\, g\left(\widetilde{X}(\tau_n)\right),\right. \\
&\qquad \left.\tau_m+h \geqslant \tau_\infty\right] \\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h < \tau_\infty\right] \\
&= \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) P(\tau_\infty, (\tau_m+h)\wedge T)\, g(L), \tau_m+h \geqslant \tau_\infty\right] \\
&\quad + \mathbb{E}_{\tau,x}\left[f\left(\widetilde{X}(\tau_m)\right) \mathbb{E}_{\tau_m,\widetilde{X}(\tau_m)}\left[g\left(\widetilde{X}((\tau_m+h)\wedge T)\right)\right], \tau_m+h < \tau_\infty\right].
\end{aligned} \tag{3.57}$$

Next we let $m \to \infty$ in (3.57) to get

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[f(L)\, g\left(\widetilde{X}((\tau_\infty+h)\wedge T-)\right)\right] \\
&= \mathbb{E}_{\tau,x}\left[f(L)\, P(\tau_\infty, (\tau_\infty+h)\wedge T)\, g(L)\right]
\end{aligned} \tag{3.58}$$

where we invoked property (vii) of Definition 2.4. Next we let h decrease to zero in (3.58). This yields

$$\begin{aligned}&\mathbb{E}_{\tau,x}\left[f(L)\,g\left(\widetilde{X}(\tau_\infty)\right)\right]=\mathbb{E}_{\tau,x}\left[f(L)\,P(\tau_\infty,\tau_\infty)\,g(L)\right]\\&=\mathbb{E}_{\tau,x}\left[f(L)\,g(L)\right].\end{aligned}\tag{3.59}$$

Since f and g are arbitrary in $C_b^+(E)$, the equality in (3.59) implies that

$$\mathbb{E}_{\tau,x}\left[h\left(L,\widetilde{X}(\tau_\infty)\right)\right]=\mathbb{E}_{\tau,x}\left[h(L,L)\right]\tag{3.60}$$

for all bounded Borel measurable functions $h\in L^\infty(E\times E,\mathcal{E}\otimes\mathcal{E})$. In particular we may take a bounded continuous metric $h(x,y)=d(x,y)$, $(x,y)\in E\times E$. From (3.60) it follows that

$$\mathbb{E}_{\tau,x}\left[d\left(L,\widetilde{X}(\tau_\infty)\right)\right]=\mathbb{E}_{\tau,x}\left[d(L,L)\right]=0,$$

and hence

$$L=\lim_{n\to\infty}\widetilde{X}(\tau_n)=\widetilde{X}(\tau_\infty),\quad \mathbb{P}_{\tau,x}\text{-almost surely.}\tag{3.61}$$

Essentially speaking this proves Theorem 2.9 in case we are dealing with conservative Feller propagators, i.e. Feller propagators with the property that $P(s,t)\,\mathbf{1}=\mathbf{1}$, $0\leqslant s\leqslant t\leqslant T$. In order to be correct the process, or rather the family of probability spaces in (3.28) has to be replaced with (3.52).

This completes the proof of Theorem 2.9 in case the Feller propagator is phrased in terms of probabilities $P(\tau,x;t,E)=1$, $0\leqslant\tau\leqslant t\leqslant T$, $x\in E$. □

The case $P(s,t)\,\mathbf{1}\leqslant\mathbf{1}$ is treated next. It will complete the proof of Theorem 2.9.

Proof. [Continuation of the proof of Theorem 2.9 in case of sub-probabilities.] We have to modify the proof in case a point of absorption is required. Most of the proof for the case that $P(\tau,x;t,E)=1$ can be repeated with the probability transition function $N(\tau,x;t,B)$, $B\in\mathcal{E}^\triangle$. This function was defined in (2.100) of Remark 2.10. However, we need to show that the $E^\triangle$-valued process $\widetilde{X}$ does not enter the absorption state $\triangle$ prior to reentering the state space E. This requires an extra argument. We will use a stopping time argument and Doob's optional sampling time theorem to achieve this: see Proposition 3.2 in which the transition function $N(\tau,x;t,B)$ is also employed.

For further use we will also need a Skorohod space with a point of absorption $\triangle$. The space $\Omega^{\triangle,\prime}$ consists of those $\omega \in \left(E^{\triangle}\right)^{[0,T]}$ whose restrictions to $D\bigcap[0,T]$ have left and right limits in $E^{\triangle}$, and which are such that for some $t = t(\omega) \in [0,T]$ the range $\{\omega(s) : s \in D\bigcap[0,t']\}$ is totally bounded in E for all $t' \in (0,t)$, and such that $\omega(s) = \triangle$ for $s \in D\bigcap[t(\omega),T]$. Again using a metric $d_{\triangle}$ on $E^{\triangle} \times E^{\triangle}$ which renders $E^{\triangle}$ Polish, it can be shown that $\Omega^{\triangle,\prime}$ is a measurable subset of $\Omega = \left(E^{\triangle}\right)^{[0,T]}$. In fact $\Omega^{\triangle,\prime}$ can be written as

$$\begin{aligned}
&\Omega^{\triangle,\prime}\\
&= \bigcup_{r\in D\bigcap[0,T]} \left\{\omega \in \Omega : s \mapsto \omega(s), s \in D\bigcap[0,r] \text{ has left and right limits in } E\right\}\\
&\quad\bigcap \bigcap_{m=1}^{\infty} \bigcup_{\substack{r_1<r_2,\, r_2-r_1<1/m\\ r_1,\, r_2\in D\bigcap[0,T]}} \left(\left\{\omega \in \Omega :\, \omega\left(D\bigcap[0,r_1]\right) \text{ is totally bounded in } E\right\}\right.\\
&\qquad\left.\bigcap \left\{\omega \in \Omega :\, \omega(s) = \triangle \text{ for all } s \in D\bigcap[r_2,T]\right\}\right). \qquad (3.62)
\end{aligned}$$

From (3.62) it follows that $\Omega^{\triangle,\prime}$ is a measurable subset of $\Omega = \left(E^{\triangle}\right)^{[0,T]}$. Again it turns out that $\mathbb{P}_{\tau,x}\left(\Omega^{\triangle,\prime}\right) = 1$. This fact follows from Proposition 3.2 and the fact that for all $t \in D\bigcap[0,T]$

$$\begin{aligned}
&\mathbb{P}_{\tau,x}\Big[\omega \in \Omega :\ s \mapsto \omega(s),\ s \in D\bigcap[0,t]\\
&\qquad\qquad \text{has left and right limits in } E, \text{ and } X(t) \in E\Big]\\
&= \mathbb{P}_{\tau,x}\left[\omega \in \Omega :\ \omega(t) \in E\right]. \qquad (3.63)
\end{aligned}$$

The equality in (3.63) follows in the same way as the corresponding result in case $P(\tau,x;t,B)$, $B \in \mathcal{E}$, but now with $N(\tau,x;t,B)$, $B \in \mathcal{E}^{\triangle}$. Again the construction which led to the process in (3.52) can be performed to get a strong Markov process of the form:

$$\left\{\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}_T^{\tau}, \widetilde{\mathbb{P}}_{\tau,x}\right), \left(\widetilde{X}(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), \left(E^{\triangle}, \mathcal{E}^{\triangle}\right)\right\}, \quad (3.64)$$

where $\widetilde{\Omega}$ is the Skorohod space $D\left([0,T], E^{\triangle}\right)$.

Since for functions $f \in C_b(E)$ we have

$$P(\tau,t)\, f(x) = \int P(\tau,x;t,dy)\, f(y) = \int N(\tau,x;tdy)\, f(y) \qquad (3.65)$$

provided $f(\triangle) = 0$, it follows that the process $\widetilde{X}$ is quasi-left continuous on its life time ζ; see Definition 2.15. For the definition of $N(\tau,x;t,B)$ see

Remark 2.10. In order to be correct the process, or rather the family of probability spaces in (3.28) has to be replaced with (3.64).

The arguments in Proposition 3.2 below then complete the proof of Theorem 2.9 in case the Feller propagator is phrased in terms of sub-probabilities $P(\tau, x; t, E) \leqslant 1$, $0 \leqslant \tau \leqslant t \leqslant T$, $x \in E$. □

In the final part of the proof of Theorem 2.9 we needed the following proposition. The proposition says that an orbit $s \mapsto \left(s, \widetilde{X}(s)\right)$ is contained in a compact subset of $[\tau, t] \times E$ on the event $\left\{\widetilde{X}(t) \in E\right\}$.

Proposition 3.2. *Suppose the transition function $P(\tau, x; t, B)$, which satisfies the equation of Chapman-Kolmogorov, consists of sub-probability Borel measures. Let $N(\tau, x; t, B)$, $B \in \mathcal{E}^{\triangle}$ be the Feller transition function as constructed in Remark 2.10, which now consists of genuine Borel probability measures on the Borel field $\mathcal{E}^{\triangle}$ of $E^{\triangle}$. As in (3.28) construct the corresponding Markov process*

$$\left\{\left(\Omega, \widetilde{\mathcal{F}}_T^{\tau}, \mathbb{P}_{\tau,x}\right), \left(\widetilde{X}(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), \left(E^{\triangle}, \mathcal{E}^{\triangle}\right)\right\}. \tag{3.66}$$

Fix $(\tau, x) \in [0, T] \times E$ and $t \in [\tau, T]$. Then the orbit

$$\left\{\left(s, \widetilde{X}(s)\right) : \tau \leqslant s \leqslant t,\ X(t) \in E\right\}$$

is $\mathbb{P}_{\tau,x}$-almost surely a relatively compact subset of $[\tau, t] \times E$.

Proof. A proof can be based on a stopping time argument and Doob's optional sampling theorem. Let the life time $\zeta : \Omega \to [0, T]$ be defined by

$$\zeta = \begin{cases} \inf\left\{s > 0 : \widetilde{X}(s) = \triangle\right\}, & \text{if } \widetilde{X}(s) = \triangle \text{ for some } s \leqslant T, \\ T & \text{otherwise.} \end{cases}$$

Then ζ is an $\left(\widetilde{\mathcal{F}}_t^{\tau}\right)_{t\in[\tau,t]}$-stopping time and we have:

$$\begin{aligned}
&\mathbb{P}_{\tau,x}\left[\widetilde{X}(t) \in E\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{P}_{\zeta\wedge t, \widetilde{X}(\zeta\wedge t)}\left[\widetilde{X}(t) \in E\right]\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{P}_{\zeta\wedge t, \widetilde{X}(\zeta\wedge t)}\left[\widetilde{X}(t) \in E\right], \zeta \leqslant t\right] \\
&\quad + \mathbb{E}_{\tau,x}\left[\mathbb{P}_{\zeta\wedge t, \widetilde{X}(\zeta\wedge t)}\left[\widetilde{X}(t) \in E\right], \zeta > t\right]
\end{aligned}$$

$$= \mathbb{E}_{\tau,x}\left[\mathbb{P}_{\zeta,\widetilde{X}(\zeta)}\left[\widetilde{X}(t) \in E\right], \zeta \leqslant t\right] + \mathbb{E}_{\tau,x}\left[\mathbb{P}_{t,\widetilde{X}(t)}\left[\widetilde{X}(t) \in E\right], \zeta > t\right]$$
$$= \mathbb{E}_{\tau,x}\left[\mathbb{P}_{\zeta,\triangle}\left[\widetilde{X}(t) \in E\right], \zeta \leqslant t\right] + \mathbb{P}_{\tau,x}\left[\widetilde{X}(t) \in E, \zeta > t\right]$$

(see Remark 2.10)

$$= \mathbb{E}_{\tau,x}\left[N(\zeta, \triangle; t, E), \zeta \leqslant t\right] + \mathbb{P}_{\tau,x}\left[\widetilde{X}(t) \in E, \zeta > t\right]$$
$$= \mathbb{P}_{\tau,x}\left[\widetilde{X}(t) \in E, \zeta > t\right]. \tag{3.67}$$

From (3.67) it follows that on the event $\left\{\widetilde{X}(t) \in E\right\}$ the orbits

$$\left\{\left(s, \widetilde{X}(s)\right) : s \in [\tau, t]\right\}$$

are $\mathbb{P}_{\tau,x}$-almost surely contained in compact subsets of $[\tau, t] \times E$.

This completes the proof of Proposition 3.2. □

In the proof of Proposition 3.4 we need the following result. Notice that in this Proposition 3.3 as well as in Proposition 3.4 the conservative property (3.68) is employed. Proposition 3.2 contains a result which can be used in the non-conservative situation. The possibility of non-conservativeness plays a role in the proof of Theorem 2.12 as well: see the inequalities in (3.122) and (3.123), and their consequences. This proposition could be called a $\mathbb{P}_{\tau,x}$-almost sure $\mathcal{T}_\beta$-equi-continuity result.

Proposition 3.3. *Let (τ, x) be an element in $[0, T] \times E$, and assume*

$$\mathbb{P}_{\tau,x}\left[X(t) \in E\right] = P(\tau, t)\, \mathbf{1}_E(x) = P(\tau, x; t, E) = 1 \tag{3.68}$$

for all $t \in [\tau, T]$. Let $(f_m)_{m\in\mathbb{N}}$ be a sequence in $C_b^+([\tau, T] \times E)$ which decreases pointwise to zero. Denote by D the collection of positive dyadic numbers. Then the following equality holds $\mathbb{P}_{\tau,x}$-almost surely:

$$\inf_{m\in\mathbb{N}} \sup_{t\in D\bigcap[\tau,T]} \sup_{s\in D\bigcap[0,t]} \mathbb{E}_{s,X(s)}\left[f_m(t, X(t))\right] = 0. \tag{3.69}$$

Consequently, the collection of linear functionals $\Lambda_{s,t} : C_b([\tau, T] \times E) \to \mathbb{C}$ defined by $\Lambda_{s,t}(f) = \mathbb{E}_{s,X(s)}\left[f(t, X(t))\right]$, $f \in C_b([\tau, T] \times E)$, $\tau \leqslant s \leqslant t \leqslant T$, s, $t \in D$, is $\mathbb{P}_{\tau,x}$-almost surely equi-continuous for the strict topology $\mathcal{T}_\beta$.

Let (s_n, t_n) be any sequence in $[\tau, T] \times [\tau, T]$ such that $s_n \leqslant t_n$, $n \in \mathbb{N}$. Then the collection

$$\left\{\Lambda_{s,t} : \tau \leqslant s \leqslant t \leqslant T,\ s,\ t \in D \text{ or } (s, t) = (s_n, t_n) \text{ for some } n \in \mathbb{N}\right\}$$

is $\mathbb{P}_{\tau,x}$-almost surely equi-continuous as well.

Proof. Let $(f_m)_{m\in\mathbb{N}} \subset C_b^+([\tau,T]\times E)$ be as in Proposition 3.3. For every $m \in \mathbb{N}$ and $t \in [\tau,T]$ we define the $\mathbb{P}_{\tau,x}$-martingale $s \mapsto M_{t,m}(s)$, $s \in [\tau,T]$, by $M_{t,m}(s) = \mathbb{E}_{s\wedge t, X(s\wedge t)}\left[f_m(t,X(t))\right]$. Then the process

$$s \mapsto \sup_{t\in[\tau,T]} \mathbb{E}_{s\wedge t,X(s\wedge t)}\left[f_m(t,X(t))\right] = \sup_{t\in D\bigcap[\tau,T]} \mathbb{E}_{s\wedge t,X(s\wedge t)}\left[f_m(t,X(t))\right]$$

is a $\mathbb{P}_{\tau,x}$-submartingale. Fix $\eta > 0$. By Doob's submartingale inequality we have

$$\begin{aligned}
&\eta\mathbb{P}_{\tau,x}\left[\sup_{t\in D\bigcap[\tau,T]}\ \sup_{s\in D\bigcap[t,T]} \mathbb{E}_{s,X(s)}\left[f_m(t,X(t))\right] \geqslant \eta\right] \\
&= \eta\mathbb{P}_{\tau,x}\left[\sup_{t\in D\bigcap[\tau,T]}\ \sup_{s\in D\bigcap[\tau,T]} M_{m,t}(s) \geqslant \eta\right] \\
&= \eta\mathbb{P}_{\tau,x}\left[\sup_{s\in D\bigcap[\tau,T]}\ \sup_{t\in D\bigcap[\tau,T]} M_{m,t}(s) \geqslant \eta\right] \\
&\leqslant \mathbb{E}_{\tau,x}\left[\sup_{t\in D\bigcap[\tau,T]} M_{m,t}(T)\right] = \mathbb{E}_{\tau,x}\left[\sup_{t\in D\bigcap[\tau,T]} \mathbb{E}_{t,X(t)}\left[f_m(t,X(t))\right]\right] \\
&= \mathbb{E}_{\tau,x}\left[\sup_{t\in D\bigcap[\tau,T]} f_m(t,X(t))\right]. \qquad (3.70)
\end{aligned}$$

Since the orbit $\{(t,X(t)) : t \in D\bigcap[\tau,T]\}$ is $\mathbb{P}_{\tau,x}$-almost surely contained in a compact subset of E, Dini's lemma implies that

$$\sup_{t\in D\bigcap[\tau,T]} f_m(t,X(t)) \quad \text{decreases to } 0\ \mathbb{P}_{\tau,x}\text{-almost surely,}$$

which implies

$$\lim_{m\to\infty} \mathbb{E}_{\tau,x}\left[\sup_{t\in D\bigcap[\tau,T]} f_m(t,X(t))\right] = 0. \qquad (3.71)$$

A combination of (3.70) and (3.71) yields (3.69). So the first part of Proposition 3.3 has been established.

The second assertion follows from (3.69) together with Theorem 2.3.

The third assertion follows from the fact that for $f \in C_b^+([\tau,T]\times E)$ and $\tau \leqslant s_n \leqslant t_n \leqslant T$ the inequality

$$\mathbb{E}_{s_n,X(s_n)}\left[f(t_n,X(t_n))\right] \leqslant \sup_{t\in D\bigcap[\tau,T]}\ \sup_{s\in D\bigcap[\tau,t]} \mathbb{E}_{s,X(s)}\left[f(t,X(t))\right]$$

holds $\mathbb{P}_{\tau,x}$-almost surely.

This shows Proposition 3.3. □

The next proposition was used in the proof of Theorem 2.9. This proposition contains an interesting continuity result of Feller evolutions.

Proposition 3.4. *Let $(\tau, x) \in [0,T] \times E$, and assume the conservative property (3.68). In addition, let $f \in C_b([0,T] \times E)$ and let $((s_n, t_n))_{n \in \mathbb{N}}$ be sequence in $[\tau, T] \times [\tau, T]$ such that $s_n \leqslant t_n$, $n \in \mathbb{N}$ and such that $\lim_{n\to\infty} (s_n, t_n) = (s, t)$. Then the limit*

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{E}_{s_n, X(s_n)} \left[f\left(t_n, X\left(t_n\right)\right)\right] = \lim_{n\to\infty} \left[P\left(s_n, t_n\right) f\left(t_n, \cdot\right)\right]\left(X\left(s_n\right)\right) \\ &= \left[P\left(s,t\right) f\left(t,\cdot\right)\right]\left(X\left(s\right)\right) = \mathbb{E}_{s,X(s)}\left[f\left(t, X(t)\right)\right] \end{aligned} \tag{3.72}$$

exists $\mathbb{P}_{\tau,x}$-almost surely. In particular if $s_n = t_n$ for all $n \in \mathbb{N}$, then $s = t$ and

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{E}_{t_n, X(t_n)} \left[f\left(t_n, X\left(t_n\right)\right)\right] = \lim_{n\to\infty} f\left(t_n, X\left(t_n\right)\right) \\ &= f\left(t, X(t)\right), \quad \mathbb{P}_{\tau,x}\text{-almost surely.} \end{aligned} \tag{3.73}$$

In addition, by taking $t_n = t$ and letting the sequence $(s_n)_{n\in\mathbb{N}}$ decrease or increase to $s \in [\tau, t]$ it follows that the process $s \mapsto \mathbb{E}_{s,X(s)}\left[f\left(t, X(t)\right)\right]$ is $\mathbb{P}_{\tau,x}$-almost surely a left and right continuous martingale. Moreover, the equalities

$$\mathbb{E}_{\tau,x}\left[f\left(t, X(t)\right) \,\middle|\, \mathcal{F}^{\tau}_{s+}\right] = \mathbb{E}_{s,X(s)}\left[f\left(t, X(t)\right)\right] = \mathbb{E}_{\tau,x}\left[f\left(t, X(t)\right) \,\middle|\, \mathcal{F}^{\tau}_{s}\right] \tag{3.74}$$

hold $\mathbb{P}_{\tau,x}$-almost surely.

The equalities in (3.39) then follow from (3.74) together with the Monotone Class Theorem.

Proof. In the proof of Proposition 3.4 we will employ the properties of the process in (3.7) to its full extent. In addition we will use Proposition 3.3 which implies that continuity properties of the process

$$\begin{aligned} (s,t) \mapsto &\alpha \int_t^\infty e^{-\alpha(\rho - t)} \mathbb{E}_{s,X(s)}\left[f\left(\rho \wedge T, X\left(\rho \wedge T\right)\right)\right] d\rho \\ &= \alpha \int_t^\infty e^{-\alpha(\rho-t)} P\left(s, \rho\wedge T\right) f\left(\rho \wedge T, \cdot\right)\left(X(s)\right) d\rho \\ &= \alpha P\left(s,t\right) R(\alpha) f\left(t,\cdot\right)\left(X(s)\right) \\ &= \int_0^\infty e^{-\rho} \mathbb{E}_{s,X(s)}\left[f\left(\left(t + \frac{\rho}{\alpha}\right) \wedge T, X\left(\left(t+\frac{\rho}{\alpha}\right)\wedge T\right)\right)\right] d\rho, \end{aligned} \tag{3.75}$$

$0 \leqslant s \leqslant t \leqslant T$, $\mathbb{P}_{\tau,x}$-almost surely carry over to the process

$$
\begin{aligned}
(s,t) \mapsto & P(s,t) f(t,\cdot)(X(s)) = \mathbb{E}_{s,X(s)}\left[f(t,X(t))\right] \\
&= \lim_{\alpha\to\infty} \alpha \int_t^\infty e^{-\alpha(\rho-t)} P(s,\rho\wedge T) f(\rho\wedge T,\cdot)(X(s))\, d\rho \\
&= \lim_{\alpha\to\infty} \int_0^\infty e^{-\rho} P\left(s, \left(t+\frac{\rho}{\alpha}\right)\wedge T\right) f\left(\left(t+\frac{\rho}{\alpha}\right)\wedge T,\cdot\right)(X(s))\, d\rho \\
&= \lim_{\alpha\to\infty} \int_0^\infty e^{-\rho} \mathbb{E}_{s,X(s)}\left[f\left(\left(t+\frac{\rho}{\alpha}\right)\wedge T, X\left(\left(t+\frac{\rho}{\alpha}\right)\wedge T\right)\right)\right] d\rho.
\end{aligned} \tag{3.76}
$$

Let $(s_n,t_n)_{n\in\mathbb{N}}$ be a sequence in $[\tau,T]\times[\tau,T]$ for which $s_n \leqslant t_n$. Put

$$\Lambda_{\alpha,s,t} f = \alpha \int_t^\infty e^{-\alpha(\rho-t)} P(s,\rho\wedge T) f(\rho\wedge T,\cdot)(X(s))\, d\rho.$$

The equality in (3.75) in conjunction with Proposition 3.3 shows that the collection of functionals

$$\left\{\Lambda_{\alpha,s,t} : \tau \leqslant s \leqslant t \leqslant T,\ s,\, t\in D \text{ or } (s,t) = (s_n,t_n) \text{ for some } n\in\mathbb{N}, \alpha \geqslant 1\right\}$$

is $\mathbb{P}_{\tau,x}$-almost surely $\mathcal{T}_\beta$-equi-continuous. Therefore the family of its limits $\Lambda_{t,s} = \lim_{\alpha\to\infty} \Lambda_{\alpha,s,t}$ inherits the continuity properties from the family

$$\left\{\Lambda_{\alpha,s,t} : \tau \leqslant s \leqslant t \leqslant T,\ s,\, t\in D \text{ or } (s,t) = (s_n,t_n) \text{ for some } n\in\mathbb{N}\right\}$$

where $\alpha \in (0,\infty)$ is fixed.

We still have to prove that

$$\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f(t_n, X(t_n))\right] = \mathbb{E}_{s,X(s)}\left[f(t,X(t))\right] \tag{3.77}$$

$\mathbb{P}_{\tau,x}$-almost surely, whenever $f \in C_b([\tau,t]\times E)$ and the sequence $(s_n,t_n)_{n\in\mathbb{N}}$ in $[\tau,T]\times[\tau,T]$ is such that $\lim_{n\to\infty}(s_n,t_n) = (s,t)$ and $s_n \leqslant t_n$ for all $n\in\mathbb{N}$. In view of the first equality in (3.22) and the previous arguments it suffices to prove this equality for processes of the form

$$(s,t) \mapsto \alpha \int_t^\infty e^{-\alpha(\rho-t)} \mathbb{E}_{s,X(s)}\left[f(\rho\wedge T, X(\rho\wedge T))\right] d\rho$$

instead of

$$(s,t) \mapsto \mathbb{E}_{s,X(s)}\left[f(t,X(t))\right].$$

It is easy to see that this convergence reduces to treating the case where, for $\rho \in (\tau,T]$ fixed and for $s_n \to s$, $s\in[\tau,\rho]$,

$$\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] = \mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho))\right]. \tag{3.78}$$

Here we will distinguish two cases: s_n increases to s and s_n decreases to s. In both cases we will prove the equality in (3.78). In case of an increasing the result follows more or less directly from the martingale martingale property and from the left continuity on the diagonal. In case of a decreasing sequence we employ the fact that a subspace of the form $\{P(\rho,u)\,g : u \in (\rho,T],\, g \in C_b(E)\}$ is $\mathcal{T}_\beta$-dense in $C_b(E)$. First we consider the situation where s_n increases to $s \in [\tau,\rho]$. Then we have

$$\begin{aligned}
\mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] &= \mathbb{E}_{\tau,x}\left[f(\rho,X(\rho)) \mid \mathcal{F}^\tau_{s_n}\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{\tau,x}\left[f(\rho,X(\rho)) \mid \mathcal{F}^\tau_{s}\right] \mid \mathcal{F}^\tau_{s_n}\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho))\right] \mid \mathcal{F}^\tau_{s_n}\right] \\
&= \mathbb{E}_{s_n,X(s_n)}\left[\mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho))\right]\right] \\
&= \left(P(s_n,s)\,\mathbb{E}_{s,\cdot}\left[f(\rho,X(\rho))\right]\right)(X(s_n)). \qquad (3.79)
\end{aligned}$$

In (3.79) we let $n \to \infty$ and use the left continuity of the propagator (see property (v) in Definition 2.4) to conclude

$$\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] = \mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho))\right]. \qquad (3.80)$$

The equality in (3.80) shows the $\mathbb{P}_{\tau,x}$-almost sure left continuity of the process $s \mapsto \mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho))\right]$ on the interval $[\tau,\rho]$. Next assume that the sequence $(s_n)_{n\in\mathbb{N}}$ decreases to $s \in [\tau,\rho]$. Then we get $\mathbb{P}_{\tau,x}$-almost surely

$$\mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho))\right] = P(s,\rho)\,f(\rho,\cdot)(X(\rho))$$

(employ (vi) of Definition 2.4)

$$\begin{aligned}
&= \lim_{n\to\infty} P(s_n,\rho)\,f(\rho,\cdot)(X(s_n)) = \lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] \\
&= \mathbb{E}_{s,X(s)}\left[\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] \mid \mathcal{F}^s_{s+}\right] \\
&= \lim_{n\to\infty} \mathbb{E}_{s,X(s)}\left[\mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] \mid \mathcal{F}^s_{s+}\right] \\
&= \lim_{n\to\infty} \mathbb{E}_{s,X(s)}\left[\mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho)) \mid \mathcal{F}^s_{s_n}\right] \mid \mathcal{F}^s_{s+}\right]
\end{aligned}$$

(tower property of conditional expectation)

$$\begin{aligned}
&= \mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho)) \mid \mathcal{F}^s_{s+}\right] \\
&= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s,X(s)}\left[f(\rho,X(\rho)) \mid \mathcal{F}^s_{s+}\right] \mid \mathcal{F}^\tau_{s+}\right] \\
&= \mathbb{E}_{\tau,x}\left[\lim_{n\to\infty} \mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] \mid \mathcal{F}^\tau_{s+}\right] \\
&= \lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s_n,X(s_n)}\left[f(\rho,X(\rho))\right] \mid \mathcal{F}^\tau_{s+}\right]
\end{aligned}$$

(Markov property)

$$= \lim_{n\to\infty} \mathbb{E}_{\tau,x}\left[\mathbb{E}_{\tau,x}\left[f(\rho, X(\rho)) \mid \mathcal{F}^{\tau}_{s_n}\right] \mid \mathcal{F}^{\tau}_{s+}\right]$$

(tower property of conditional expectation)

$$= \mathbb{E}_{\tau,x}\left[f(\rho, X(\rho)) \mid \mathcal{F}^{\tau}_{s+}\right]. \tag{3.81}$$

The equality in (3.81) is the same as the first equality in (3.74). The second equality is a consequence of the Markov property with respect to the filtration $(\mathcal{F}^{\tau}_t)_{t\in[\tau,T]}$.

This completes the proof of Proposition 3.3. □

3.1.2 *Proof of Theorem 2.10*

Here we have to prove that Markov processes with certain continuity properties give rise to Feller evolutions.

Proof. [Proof of Theorem 2.10.] Let the operators $P(\tau,t)$, $\tau \leqslant t$, be as in (2.95). We have to prove that this collection is a Feller evolution. The properties (i), (iii) and (iv) of Definition 2.4 are obvious. The propagator property (ii) is a consequence of the Markov property of the process in (2.94). To be precise, let $f \in C_b(E)$ and $0 \leqslant \tau < s < t \leqslant T$. Then we have:

$$\begin{aligned} P(\tau,s)\,P(s,t)\,f(x) &= \mathbb{E}_{s,x}\left[P(s,t)\,f(X(s))\right] = \mathbb{E}_{\tau,x}\left[\mathbb{E}_{s,X(s)}\left[f(X(t))\right]\right] \\ &= \mathbb{E}_{\tau,x}\left[\mathbb{E}_{\tau,x}\left[f(X(t)) \mid \mathcal{F}^{\tau}_{s}\right]\right] \\ &= \mathbb{E}_{\tau,x}\left[f(X(t))\right] = P(\tau,t)\,f(x). \end{aligned} \tag{3.82}$$

Let f be any function in $C_b(E)$. The continuity of the function $(\tau,t,x) \mapsto P(\tau,t)\,f(x)$, $0 \leqslant \tau \leqslant t \leqslant T$, $x \in E$, implies the properties (v) through (vii) of Definition 2.4. Let $f \in C_b([0,T]\times E)$. In addition we have to prove that the function $(\tau,t,x) \mapsto P(\tau,t)\,f(t,\cdot)(x)$ is continuous. The proof of this fact requires the following steps:

(1) The Feller evolution $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ is $\mathcal{T}_\beta$-equi-continuous.
(2) Define the operators $R(\alpha) : C_b([0,T]\times E) \to C_b([0,T]\times E)$, $\alpha > 0$, as in (4.6) in Chapter 4;

$$\begin{aligned} R(\alpha)f(t,x) &= \int_t^\infty e^{-\alpha(\rho-t)} P(t,\rho\wedge T)\,f(\rho\wedge T,\cdot)(x)d\rho \\ &= \int_0^\infty e^{-\alpha\rho} S(\rho)f(t,x)\,d\rho,\ \ f \in C_b([0,T]\times E), \end{aligned} \tag{3.83}$$

where, by definition,

$$S(\rho)f(t,x) = P(t,(\rho+t)\wedge T)f((\rho+t)\wedge T,\cdot)(x),\ \ f\in C_b([0,T]\times E). \tag{3.84}$$

Since the family of operators $\{S(\rho):\ \rho\geqslant 0\}$ has the semigroup property, i.e. $S(\rho_1)\,S(\rho_2) = S(\rho_1+\rho_2)$, ρ_1, $\rho_2\geqslant 0$, the family $\{R(\alpha):\ \alpha>0\}$ has the resolvent property: see (3.85) below. Moreover, the functions $(\tau,t,x)\mapsto P(\tau,t)\left[R(\alpha)f(t,\cdot)\right](x)$, $0\leqslant\tau\leqslant t\leqslant T$, $x\in E$, $\alpha>0$, are continuous for all $f\in C_b([0,T]\times E)$.

(3) The family $\{R(\alpha):\ \alpha>0\}$ is a resolvent family, and hence the range of $R(\alpha)$ does not depend on $\alpha>0$. The $\mathcal{T}_\beta$-closure of its range coincides with $C_b([0,T]\times E)$.

From (3), (1) and (2) it then follows that functions of the form $P(\tau,t)\,f(t,\cdot)(x)$, $0\leqslant\tau\leqslant t\leqslant T$, $f\in C_b([0,T]\times E)$, are continuous. So we have to prove (1) through (3).

Let $(\psi_m)_{m\in\mathbb{N}}$ be a sequence of functions in $C_+(E)$ which decreases pointwise to zero. Since, by assumption, the functions $(\tau,t,x)\mapsto P(\tau,t)\,\psi_m(x)$, $m\in\mathbb{N}$, are continuous, the sequence $P(\tau,t)\,\psi_m(x)$ decreases uniformly on compact subsets to 0. By Theorem 2.7 it follows that the Feller evolution $\{P(\tau,t):\ 0\leqslant\tau\leqslant t\leqslant T\}$ is $\mathcal{T}_\beta$-equi-continuous. This proves (1).

Let $f\in C_b([0,T]\times E)$, and fix $\alpha>0$. Then the function $P(\tau,t)\,R(\alpha)f$ can be written in the form

$$P(\tau,t)\left[R(\alpha)f(t,\cdot)\right](x) = \int_t^\infty e^{-\alpha(\rho-t)}P(\tau,\rho\wedge T)\,f(\rho\wedge T,\cdot)(x)d\rho,$$

which by inspection is continuous, because for fixed $\rho\in[0,T]$ the function $(\tau,x)\mapsto P(\tau,\rho)\,f(\rho,\cdot)(x)$ is continuous. This proves Assertion (2).

The family $\{R(\alpha):\ \alpha>0\}$ is a resolvent family, i.e. it satisfies:

$$R(\beta) = R(\alpha) + (\alpha-\beta)R(\alpha)R(\beta),\quad \alpha,\ \beta>0. \tag{3.85}$$

Consequently, the range $R(\alpha)C_b([0,T]\times E)$ does not depend on $\alpha>0$. Next fix $f\in C_b([0,T]\times E)$. Then $\lim_{\alpha\to\infty}\alpha R(\alpha)f(t,x) = f(t,x)$ for all $(t,x)\in[0,T]\times E$. By dominated convergence it also follows that

$$\begin{aligned}
&\lim_{\alpha\to\infty}\int_{[0,T]\times E}\alpha R(\alpha)f(t,x)\,d\mu(t,x)\\
&= \lim_{\alpha\to\infty}\int_{[0,T]\times E}\int_0^\infty e^{-\rho}P\left(t,\left(t+\frac{\rho}{\alpha}\right)\wedge T\right)f\left(\left(t+\frac{\rho}{\alpha}\right)\wedge T,\cdot\right)(x)d\rho\,d\mu(t,x)\\
&= \int_{[0,T]\times E} f(t,x)\,d\mu(t,x),
\end{aligned} \tag{3.86}$$

where μ is a complex Borel measure on $[0,T] \times E$. From (3.86) and Corollary 2.1 we see that the space $R(\alpha)C_b([0,T] \times E)$ is $\mathcal{T}_\beta$-weakly dense in $C_b([0,T] \times E)$. It follows that it is $\mathcal{T}_\beta$-dense. Let K be a compact subset of E. Since the Feller evolution is $\mathcal{T}_\beta$-equi-continuous there exists a bounded function $u \in H^+([0,T] \times E)$ such that

$$\sup_{(t,x)\in[0,T]\times K} |P(\tau,t) f(x)| \leqslant \|uf\|_\infty, \quad f \in C_b(E). \tag{3.87}$$

Fix $\varepsilon > 0$. For $\alpha_0 > 0$ and $f \in C_b([0,T] \times E)$ fixed, there exists a function $g \in C_b([0,T] \times E)$ such that

$$\sup_{s\in[0,T]} \sup_{y\in E} |u(s,y)(f(s,y) - \alpha_0 R(\alpha_0) g(s,y))| \leqslant \varepsilon. \tag{3.88}$$

From (3.87) and (3.88) we infer:

$$\begin{aligned} &\sup_{0\leqslant\tau\leqslant t\leqslant T} \sup_{x\in K} |P(\tau,t)[f(t,\cdot) - \alpha_0 R(\alpha_0) g(t,\cdot)](x)| \\ &\leqslant \sup_{0\leqslant s\leqslant T} \sup_{y\in E} |u(y)(f(s,y) - \alpha_0 R(\alpha_0)(s,y))| \leqslant \varepsilon. \end{aligned} \tag{3.89}$$

As a consequence of (3.89) the function $(\tau,t,x) \mapsto P(\tau,t) f(t,\cdot)(x)$ inherits its continuity properties from functions of the form

$$(\tau,t,x) \mapsto P(\tau,t) R(\alpha_0) f(t,\cdot)(x), \quad 0 \leqslant \tau \leqslant t \leqslant T, \ x \in E.$$

Since the latter functions are continuous, the same is true for the function $P(\tau,t) f(t,\cdot)(x)$.

This concludes the proof of Theorem 2.10. □

As a corollary we mention the following: its proof follows from the arguments leading to the observation that for all $f \in C_b([0,T] \times E)$ the function $(\tau,t,x) \mapsto P(\tau,t) f(t,\cdot)(x)$ is continuous. It will be used in the proof of Theorem 4.3 in Chapter 4.

Corollary 3.1. *Let the family $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ be a Feller evolution in $C_b(E)$. Extend these operators to the space $C_b([0,T] \times E)$ by the formula $\widetilde{P}(\tau,t) f(\tau,x) = P(\tau,t) f(t,\cdot)(x)$, $f \in C_b([0,T] \times E)$. Then the family $\left\{\widetilde{P}(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\right\}$ is again $\mathcal{T}_\beta$-equi-continuous. In addition define the $\mathcal{T}_\beta$-continuous semigroup $\{S(t) : t \geqslant 0\}$ on $C_b([0,T] \times E)$ by*

$$S(t)f(\tau,x) = P(\tau,(\tau+t)\wedge T) f((\tau+t)\wedge T,\cdot)(x), \quad f \in C_b([0,T] \times E). \tag{3.90}$$

Then the semigroup $\{S(t) : t \geqslant 0\}$ is $\mathcal{T}_\beta$-equi-continuous.

In the sequel we will not use the notation $\widetilde{P}(\tau, t)$ for the extended Feller evolution very much: we will simply ignore the difference between $\widetilde{P}(\tau, t)$ and $P(\tau, t)$. For more details on the semigroup defined in (3.90) see (4.5) below.

Proof. [Proof of Corollary 3.1.] Let $f \in C_b([0,T] \times E)$. From the proof of Theorem 2.10 (see the very end) we infer that the function $(\tau, t, x) \mapsto \widetilde{P}(\tau, t) f(\tau, x)$ is continuous. Let $(\psi_m)_{m\in\mathbb{N}}$ be a sequence of functions in $C_b([0,T] \times E)$ which decreases pointwise to 0. Let $u \in H^+([0,T] \times [0, T \times E)$. Then the functions $\widetilde{P}(\tau, t)(\psi_m f)(x)$ also decrease uniformly to 0. From Corollary 2.3 it follows that the family $\left\{\widetilde{P}(\tau, t) : \tau \leqslant t \leqslant T\right\}$ is $\mathcal{T}_\beta$-equi-continuous. From the representation (3.90) of the semigroup $\{S(t) : t \geqslant 0\}$, it is also clear that this semigroup is $\mathcal{T}_\beta$-equi-continuous. This completes the proof of Corollary 3.1. □

3.1.3 *Proof of Theorem 2.11*

In this part and in Theorem 2.12 we will see the intimate relationship which exists between solutions to the martingale problem and the corresponding (strong) Markov processes.

Proof. [Proof of Theorem 2.11.] In the proof of Theorem 2.11 we will use the fact that an operator L generates a Feller evolution if and only if it generates the corresponding Markov process: see Proposition 4.1 below. So we may assume that the corresponding Markov process is that of Theorem 2.9: see (2.90). Among other things this means that it is right continuous, and has left limits in E on its life time. In addition, it is quasi-left continuous on its life time: see Definition 2.15. Let $f \in C_b([0,T] \times E)$ belong to the domain of $D_1 + L$. We will show that the process in (2.96) is a $\mathbb{P}_{\tau,x}$-martingale. Therefore, fix $s \in [\tau, t]$, and put

$$M_{\tau,f}(s) = f(s, X(s)) - f(\tau, X(\tau)) - \int_\tau^s \left(\frac{\partial}{\partial\rho} + L(\rho)\right) f(\rho, \cdot)(X(\rho))\, d\rho.$$

Then by the Markov property we have

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[M_{\tau,f}(t) \,\middle|\, \mathcal{F}_s^\tau\right] - M_{\tau,f}(s) \\
&= \mathbb{E}_{\tau,x}\left[M_{s,f}(t) \,\middle|\, \mathcal{F}_s^\tau\right] = \mathbb{E}_{s,X(s)}\left[M_{s,f}(t)\right] \\
&= \mathbb{E}_{s,X(s)}\left[f(t, X(t))\right] - \mathbb{E}_{s,X(s)}\left[f(s, X(s))\right] \\
&\qquad - \int_s^t \mathbb{E}_{s,X(s)}\left[\left(\frac{\partial}{\partial\rho} + L(\rho)\right) f(\rho, \cdot)(X(\rho))\right] d\rho
\end{aligned}$$

(the operator L generates the involved Markov process)

$$
\begin{aligned}
&= \mathbb{E}_{s,X(s)}\left[f\left(t,X(t)\right)\right] - \mathbb{E}_{s,X(s)}\left[f\left(s,X(s)\right)\right] - \int_s^t \frac{d\mathbb{E}_{s,X(s)}\left[f\left(\rho,X(\rho)\right)\right]}{d\rho}\,d\rho \\
&= \mathbb{E}_{s,X(s)}\left[f\left(t,X(t)\right)\right] - \mathbb{E}_{s,X(s)}\left[f\left(s,X(s)\right)\right] - \mathbb{E}_{s,X(s)}\left[f\left(\rho,X(\rho)\right)\right]\Big|_{\rho=s}^{\rho=t} \\
&= 0. \qquad (3.91)
\end{aligned}
$$

The equality in (3.91) proves the first part of Theorem 2.10. Proposition 3.5 below proves more than what is claimed in Theorem 2.10. Therefore the proof of Theorem 2.10 is completed by Proposition 3.5. □

Proposition 3.5. *Let the Markov family of probability spaces be as in Theorem 2.9, formula (2.90). Let $\vee_t$, $\wedge_t$, $\vartheta_t : \Omega \to \Omega$, $t \in [0,T]$, be time transformations with the following respective defining properties: $X(s) \circ \vee_t = X(s \vee t)$, $X(s) \circ \wedge_t = X(s \wedge t)$, and $X(s) \circ \vartheta_t = X((s+t) \wedge T)$, for all s, $t \in [0,T]$. Let the σ-fields $\mathcal{F}_{t_2}^{t_1}$, $0 \leqslant t_1 \leqslant t_2 \leqslant T$, be defined by $\mathcal{F}_{t_2}^{t_1} = \sigma\left(X(s) : t_1 \leqslant s \leqslant t_2\right)$. Fix $t \in [0,T]$. Then the mapping $\vee_t$ is $\mathcal{F}_{t_2 \vee t}^{t_1 \vee t}$-$\mathcal{F}_{t_2}^{t_1}$-measurable, the mapping $\wedge_t$ is $\mathcal{F}_{t_2 \wedge t}^{t_1 \wedge t}$-$\mathcal{F}_{t_2}^{t_1}$-measurable, and ϑ_t is $\mathcal{F}_{(t_2+t)\wedge T t}^{(t_1+t)\wedge T}$-$\mathcal{F}_{t_2}^{t_1}$-measurable.*

Fix $\tau \in [0,T]$, and $\tau \leqslant t_1 \leqslant t_2 \leqslant T$. Let μ be a Borel probability measure on $\mathcal{E}$, and define the probability measure $\mathbb{P}_{\tau,\mu}$ on $\mathcal{F}_T^\tau$ by the formula $\mathbb{P}_{\tau,\mu}(A) = \int_E \mathbb{P}_{\tau,x}(A)d\mu(x)$, $A \in \mathcal{F}_T^\tau$. Let $\overline{\left(\mathcal{F}_{t_2}^{t_1}\right)}^{\tau,\mu}$ be the $\mathbb{P}_{\tau,\mu}$-completion of the σ-field $\mathcal{F}_{t_2}^{t_1}$. Then ($\mathbb{P}_{\tau,\mu}$-a.s. means $\mathbb{P}_{\tau,\mu}$-almost surely)

$$
\overline{\left(\mathcal{F}_{t_2}^{t_1}\right)}^{\tau,\mu} = \left\{A \in \overline{\left(\mathcal{F}_T^\tau\right)}^{\tau,\mu} : \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2} = \mathbf{1}_A,\ \mathbb{P}_{\tau,\mu}\text{-a.s.}\right\}, \qquad (3.92)
$$

and

$$
\overline{\left(\mathcal{F}_{t_2+}^{t_1}\right)}^{\tau,\mu} = \bigcap_{\varepsilon \in (0,T-t_2]} \left\{A \in \overline{\left(\mathcal{F}_T^\tau\right)}^{\tau,\mu} : \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2+\varepsilon} = \mathbf{1}_A,\ \mathbb{P}_{\tau,\mu}\text{-a.s.}\right\}. \qquad (3.93)
$$

In addition the following equalities are $\mathbb{P}_{\tau,\mu}$-almost surely valid for all bounded random variables F which are $\overline{\left(\mathcal{F}_T^\tau\right)}^{\tau,\mu}$-measurable:

$$
\mathbb{E}_{\tau,\mu}\left[F \mid \mathcal{F}_{t+}^\tau\right] = \mathbb{E}_{\tau,\mu}\left[F \mid \mathcal{F}_t^\tau\right], \quad \text{and} \qquad (3.94)
$$

$$
\mathbb{E}_{\tau,\mu}\left[F \mid \overline{\left(\mathcal{F}_{t+}^\tau\right)}^{\tau,\mu}\right] = \mathbb{E}_{\tau,\mu}\left[F \mid \mathcal{F}_{t+}^\tau\right]. \qquad (3.95)
$$

If the variable F is $\overline{\left(\mathcal{F}_{t+}^\tau\right)}^{\tau,\mu}$-measurable, then the equalities

$$
F = \mathbb{E}_{\tau,\mu}\left[F \mid \overline{\left(\mathcal{F}_{t+}^\tau\right)}^{\tau,\mu}\right] = \mathbb{E}_{\tau,\mu}\left[F \mid \mathcal{F}_{t+}^\tau\right] = \mathbb{E}_{\tau,\mu}\left[F \mid \mathcal{F}_t^\tau\right] \qquad (3.96)
$$

hold $\mathbb{P}_{\tau,\mu}$-almost surely. If the random variable F is $\overline{\left(\mathcal{F}_T^t\right)}^{\tau,\mu}$-measurable and bounded, then $\mathbb{P}_{\tau,\mu}$-almost surely

$$\mathbb{E}_{\tau,\mu}\left[F \mid \overline{\left(\mathcal{F}_{t+}^{\tau}\right)}^{\tau,\mu}\right] = \mathbb{E}_{t,X(t)}\left[F\right]. \tag{3.97}$$

Finally, if F is $\overline{\left(\mathcal{F}_{t+}^t\right)}^{\tau,\mu}$-measurable, then

$$F = \mathbb{E}_{t,X(t)}\left[F\right], \quad \mathbb{P}_{\tau,\mu}\text{-almost surely.} \tag{3.98}$$

In particular such variables are $\mathbb{P}_{\tau,x}$-almost surely functions of the space-time variable $(t, X(t))$.

Proof. Let F be a bounded $\mathcal{F}_{s_2}^{s_1}$-measurable variable. The measurability properties of the time operator $\vee_t$ follow from the fact that $F \circ \vee_t$ is $\mathcal{F}_{s_2 \vee t}^{s_1 \vee t}$-measurable. Similar statements hold for the operators $\wedge_t$ and ϑ_t.

The equality

$$\mathcal{F}_{t_2}^{t_1} = \left\{A \in \mathcal{F}_T^{\tau} : \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2} = \mathbf{1}_A, \ \mathbb{P}_{\tau,\mu}\text{-a.s.}\right\} \tag{3.99}$$

is clear, and so the left-hand side is included in the right-hand side of (3.92). This can be seen as follows. Let $A \in \overline{\mathcal{F}_{t_2}^{t_1}}^{\tau,\mu}$. Then there exist subsets A_1 and $A_2 \in \mathcal{F}_{t_2}^{t_1}$ such that $A_1 \subset A \subset A_2$ and $\mathbb{P}_{\tau,\mu}\left[A_2 \backslash A_1\right] = 0$. Then we have

$$\begin{aligned} &\mathbf{1}_{A_1} - \mathbf{1}_{A_2} = \mathbf{1}_{A_1} - \mathbf{1}_{A_2} \circ \vee_{t_1} \circ \wedge_{t_2} \\ &\quad \leqslant \mathbf{1}_A - \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2} \leqslant \mathbf{1}_{A_2} - \mathbf{1}_{A_1} \circ \vee_{t_1} \circ \wedge_{t_2} = \mathbf{1}_{A_2} - \mathbf{1}_{A_1}. \end{aligned} \tag{3.100}$$

From (3.100) we see that $\mathbf{1}_A - \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2}$, $\mathbb{P}_{\tau,x}$-almost surely, and hence the left-hand side of (3.92) is included in the right-hand side. Since by the same argument the σ-field $\left\{A \in \overline{\left(\mathcal{F}_T^{\tau}\right)}^{\tau,\mu} : \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2} = \mathbf{1}_A, \ \mathbb{P}_{\tau,\mu}\text{-a.s.}\right\}$ is $\mathbb{P}_{\tau,\mu}$-complete and since

$$\left\{A \in \mathcal{F}_T^{\tau} : \mathbf{1}_A \circ \vee_{t_1} \circ \wedge_{t_2} = \mathbf{1}_A, \ \mathbb{P}_{\tau,\mu}\text{-a.s.}\right\} \subset \mathcal{F}_{t_2}^{t_1}, \tag{3.101}$$

we also obtain that the right-hand side of (3.92) is contained in the left-hand side. The equality in (3.93) is an immediate consequence of (3.92), and the definition of $\mathcal{F}_{t_2+}^{t_1}$.

By the Monotone Class Theorem and an approximation argument the proof of (3.94) can be reduced to the case where $F = \prod_{j=1}^n f_j\left(X\left(t_j\right)\right)$ with $\tau \leqslant t_1 < \cdots < t_k \leqslant t < t_{k+1} < \cdots < t_n \leqslant T$, and $f_j \in C_b(E)$, $1 \leqslant j \leqslant n$. Then by properties of conditional expectation and the Markov property with respect to the filtration $\left(\mathcal{F}_t^{\tau}\right)_{t \in [\tau,T]}$ we have

$$\mathbb{E}_{\tau,\mu}\left[F \mid \mathcal{F}_{t+}^{\tau}\right] = \mathbb{E}_{\tau,\mu}\left[\prod_{j=1}^n f_j\left(X\left(t_j\right)\right) \mid \mathcal{F}_{t+}^{\tau}\right]$$

$$= \prod_{j=1}^{k} f_j\left(X\left(t_j\right)\right) \mathbb{E}_{\tau,\mu}\left[\mathbb{E}_{\tau,\mu}\left[\prod_{j=k+1}^{n} f_j\left(X\left(t_j\right)\right) \,\middle|\, \mathcal{F}_{t_{k+1}}^{\tau}\right] \,\middle|\, \mathcal{F}_{t+}^{\tau}\right]$$

(Markov property)

$$= \prod_{j=1}^{k} f_j\left(X\left(t_j\right)\right) \mathbb{E}_{\tau,\mu}\left[g\left(X\left(t_{k+1}\right)\right) \,\middle|\, \mathcal{F}_{t+}^{\tau}\right], \tag{3.102}$$

where $g(y) = f_{k+1}(y)\mathbb{E}_{t_{k+1},y}\left[\prod_{j=k+1}^{n} f_j\left(X\left(t_j\right)\right)\right]$. Again we may suppose that the function g belongs to $C_b(E)$. Then we get, for $t < s < t_{k+1}$,

$$\mathbb{E}_{\tau,\mu}\left[g\left(X\left(t_{k+1}\right)\right) \,\middle|\, \mathcal{F}_{t+}^{\tau}\right] = \mathbb{E}_{\tau,\mu}\left[\mathbb{E}_{\tau,\mu}\left[g\left(X\left(t_{k+1}\right)\right) \,\middle|\, \mathcal{F}_{s}^{\tau}\right] \,\middle|\, \mathcal{F}_{t+}^{\tau}\right]$$

(Markov property)

$$\begin{aligned}
&= \mathbb{E}_{\tau,\mu}\left[\mathbb{E}_{s,X(s)}\left[g\left(X\left(t_{k+1}\right)\right)\right] \,\middle|\, \mathcal{F}_{t+}^{\tau}\right] \\
&= \lim_{s\downarrow t} \mathbb{E}_{\tau,\mu}\left[\mathbb{E}_{s,X(s)}\left[g\left(X\left(t_{k+1}\right)\right)\right] \,\middle|\, \mathcal{F}_{t+}^{\tau}\right] \\
&= \mathbb{E}_{\tau,\mu}\left[\mathbb{E}_{t,X(t)}\left[g\left(X\left(t_{k+1}\right)\right)\right] \,\middle|\, \mathcal{F}_{t+}^{\tau}\right] \\
&= \mathbb{E}_{t,X(t)}\left[g\left(X\left(t_{k+1}\right)\right)\right]
\end{aligned}$$

(again Markov property)

$$= \mathbb{E}_{\tau,\mu}\left[g\left(X\left(t_{k+1}\right)\right) \,\middle|\, \mathcal{F}_{t}^{\tau}\right]. \tag{3.103}$$

Inserting the result of (3.103) into (3.102) and reverting the arguments which led to (3.102) with $\mathcal{F}_t^{\tau}$ instead of $\mathcal{F}_{t+}^{\tau}$ shows the equality in (3.94) for $F = \prod_{j=1}^{n} f_j\left(X\left(t_j\right)\right)$ where the functions f_j, $1 \leqslant j \leqslant n$, belong to $C_b(E)$. As mentioned earlier this suffices to obtain (3.94) for all bounded random variables F which are $\overline{\left(\mathcal{F}_T^{\tau}\right)}^{\tau,\mu}$-measurable. Here we use the fact that for any σ-field $\mathcal{F} \subset \overline{\left(\mathcal{F}_T^{\tau}\right)}^{\tau,\mu}$, and any bounded $\overline{\left(\mathcal{F}_T^{\tau}\right)}^{\tau,\mu}$-measurable random variable F an equality of the form $F = \mathbb{E}_{\tau,\mu}\left[F \,\middle|\, \mathcal{F}\right]$ holds $\mathbb{P}_{\tau,\mu}$-almost surely. This argument also shows that the equality in (3.95) is a consequence of (3.94). The equalities in (3.96) follow from the definition of conditional expectation and the equalities (3.94) and (3.95). The equality in (3.97) also follows from (3.94) and (3.95) together with the Markov property. Finally, the equality in (3.103) is a consequence of (3.102) and the definition of conditional expectation.

Altogether this proves Proposition 3.5. □

3.1.4 *Proof of Theorem 2.12*

In this subsection we will establish the fact that unique solutions to the martingale problem yield strong Markov processes.

Proof. [Proof of Theorem 2.12.] The proof of this result is quite technical. The first part follows from a well-known theorem of Kolmogorov on projective systems of measures: see Theorem 3.1. In the second part we must show that the indicated path space has full measure, so that no information is lost. The techniques used are reminiscent the material found in for example [Blumenthal and Getoor (1968)], Theorem 9.4. p. 46. The result in Theorem 2.12 is a consequence of the propositions 3.6, 3.7, and 3.8 below. □

In Theorem 2.12 as anywhere else in the book $L = \{L(s) : 0 \leqslant s \leqslant T\}$ is considered as a linear operator with domain $D(L)$ and range $R(L)$ in the space $C_b([0,T] \times E)$. Suppose that the domain $D(L)$ of L is $\mathcal{T}_\beta$-dense in $C_b([0,T] \times E)$. The problem we want to address is the following. Give necessary and sufficient conditions on the operator L in order that for every $(\tau, x) \in [0,T] \times E$ there exists a unique probability measure $\mathbb{P}_{\tau,x}$ on $\mathcal{F}_T^\tau$ with the following properties:

(i) For every $f \in D(L)$, which is $C^{(1)}$-differentiable in the time variable the process

$$f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t (D_1 f + Lf)(s, X(s))\, ds, \quad t \in [\tau, T],$$

is a $\mathbb{P}_{\tau,x}$-martingale;

(ii) $\mathbb{P}_{\tau,x}[X(\tau) = x] = 1$.

Here we suppose $\Omega = D([0,\infty], E^\triangle)$ is the Skohorod space associated with $E^\triangle$, as described in Definition 2.9, and $\mathcal{F}_T^\tau$ is the σ-field generated by the state variables $X(t)$, $t \in [\tau, T]$. The probability measures $\mathbb{P}_{\tau,x}$ are defined on the σ-field $\mathcal{F}_T^\tau$. The following procedure extends them to $\mathcal{F}_T^0$. If the event A belongs to $\mathcal{F}_T^0$, then we put $\mathbb{P}_{\tau,x}[A] = \mathbb{E}_{\tau,x}[\mathbf{1}_A \circ \vee_\tau]$. The composition $\mathbf{1}_A \circ \vee_\tau$ is defined in (2.73). With this convention in mind the equality in (ii) may be replaced by

(ii)′ $\mathbb{P}_{\tau,x}[X(s) = x] = 1$ for all $s \in [0, \tau]$.

Let $P(\Omega)$ be the set of all probability measures on $\mathcal{F}_T^0$ and define the subset $P_0'(\Omega)$ of $P(\Omega)$ by

$$P_0'(\Omega) = \bigcup_{(\tau,x)\in[0,T]\times E^\triangle} \Big\{ \mathbb{P} \in P(\Omega) : \mathbb{P}[X(\tau) = x] = 1$$

and for every $f \in D(L) \bigcap D(D_1)$ the process

$$f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t (D_1 + L) f(s, X(s))\, ds, \quad t \in [\tau, T],$$

$$\text{is a } \mathbb{P}\text{-martingale} \Big\}. \tag{3.104}$$

Instead of $D(L) \bigcap D(D_1)$ we often write $D^{(1)}(L)$: see the comments following Definition 2.7. Let $(v_j : j \in \mathbb{N})$ be a sequence of continuous functions defined on $[0, T] \times E$ with the following properties:

(i) $v_0 = \mathbf{1}_E$, $v_1 = \mathbf{1}_{\{\triangle\}}$;
(ii) $\|v_j\|_\infty \leqslant 1$, v_j belongs to $D^{(1)}(L) = D(L) \bigcap D(D_1)$, and $v_j(s, \triangle) = 0$ for $j \geqslant 2$;
(iii) The linear span of v_j, $j \geqslant 0$, is dense in $C_b\left([0, T] \times E^\triangle\right)$ for the strict topology $\mathcal{T}_\beta$.

In addition let $(f_k : k \in \mathbb{N})$ be a sequence in $D^{(1)}(L)$ such that the linear span of $\{(f_k, (D_1 + L) f_k) : k \in \mathbb{N}\}$ is $\mathcal{T}_\beta$ dense in the graph $G(D_1 + L) := \{(f, (D_1 + L) f) : f \in D(L)\}$ of the operator $D_1 + L$. Moreover, let $(s_j : j \in \mathbb{N})$ be an enumeration of the set $\mathbb{Q} \bigcap [0, T]$. A subset $P'(\Omega)$, which is closely related to P_0', may be described as follows (see (3.54) as well):

$$\begin{aligned} &P'(\Omega) \\ &= \bigcap_{n=1}^\infty \bigcap_{k=1}^\infty \bigcap_{m=0}^\infty \bigcap_{(j_1, \dots, j_{m+1}) \in \mathbb{N}^{m+1}} \bigcap_{0 \leqslant s_{j_1} < \dots < s_{j_{m+1}} \leqslant T} \{\mathbb{P} \in P(\Omega) : \\ &\quad \mathbb{P}\left[X(s_{j_k}) \in E,\ 1 \leqslant k \leqslant m+1\right] = \mathbb{P}\left[X\left(s_{j_{m+1}}\right) \in E\right], \quad \text{and} \\ &\int \left(f_k\left(s_{j_{m+1}}, X\left(s_{j_{m+1}}\right)\right) - f_k\left(s_{j_m}, X\left(s_{j_m}\right)\right)\right) \prod_{k=1}^m v_{j_k}\left(s_{j_k}, X\left(s_{j_k}\right)\right) d\mathbb{P} \\ &= \int \left(\int_{s_{j_m}}^{s_{j_{m+1}}} (D_1 + L) f_k(s, X(s))\, ds\right) \prod_{k=1}^m v_{j_k}\left(s_{j_k}, X\left(s_{j_k}\right)\right) d\mathbb{P} \Big\}. \end{aligned} \tag{3.105}$$

Let $P(\Omega)$ be the collection of probability measures on $\mathcal{F}_T^0$. For a concise formulation of the relevant distance between probability measures in $P(\Omega)$ we introduce kind of Lévy numbers. Let $\mathbb{P}_1$ and $\mathbb{P}_2 \in P(\Omega)$. Then we write, for $\Lambda \subset [0, T]$, Λ finite or countable,

$$L_\Lambda(\mathbb{P}_2, \mathbb{P}_1) = \lim_{\ell \to \infty} \inf \left\{ \eta > 0 : \mathbb{P}_2 \left[(X(s))_{s \in \Lambda} \subset \bigcup_{j=1}^{\ell} B\left(x_j, 2^{-m}\right) \right] \right.$$

$$\geqslant \left(1-\eta 2^{-m}\right) \mathbb{P}_1 \left[(X(s))_{s\in\Lambda} \subset E \right], \text{ for all } m \in \mathbb{N} \Big\} \quad (3.106)$$

where $B(x,\varepsilon)$ is a ball in E centered at x and with radius $\varepsilon > 0$. Perhaps a more appropriate name for a Lévy number would be a "tightness number". Notice that in (3.106) $\lim_{\ell\to\infty}$ may be replaced with $\inf_{\ell\in\mathbb{N}}$. In fact we shall prove that, if for the operator L the martingale problem is solvable, that then the set $P'(\Omega)$ is complete metrizable and separable for the metric $d(\mathbb{P}_1, \mathbb{P}_2)$ given by

$$\begin{aligned} & d_L(\mathbb{P}_1, \mathbb{P}_2) \\ & = \sum_{\Lambda\subset\mathbb{N}, |\Lambda|<\infty} 2^{-|\Lambda|} \sum_{(\ell_j)_{j\in\Lambda}} \left| \int \prod_{j\in\Lambda} 2^{-j-\ell_j} v_j\left(s_{\ell_j}, X\left(s_{\ell_j}\right)\right) d\left(\mathbb{P}_2 - \mathbb{P}_1\right) \right| \\ & \quad + \sum_{k=1}^{\infty} 2^{-k} \left(L_{\mathbb{Q}\bigcap[0,s_k]}\left(\mathbb{P}_2, \mathbb{P}_1\right) + L_{\mathbb{Q}\bigcap[0,s_k]}\left(\mathbb{P}_1, \mathbb{P}_2\right)\right). \end{aligned} \quad (3.107)$$

If a sequence of probability measures $(\mathbb{P}_n)_{n\in\mathbb{N}}$ converges to $\mathbb{P}$ with respect to the metric in (3.107), then the first term on the right-hand side says that the finite-dimensional distributions of $\mathbb{P}_n$ converge to the finite-dimensional distributions of $\mathbb{P}$. The second term says, that the limit $\mathbb{P}$ is a measure indeed, and that the paths of the process are $\mathbb{P}$-almost surely totally bounded. The following result should be compared to the comments in 6.7.4. of [Stroock and Varadhan (1979)], pp. 167–168. It is noticed that in Proposition 3.6 the uniqueness of the martingale problem is used to prove the separability.

Proposition 3.6. *The set $P'(\Omega)$ supplied with the metric d_L defined in (3.107) is a separable complete metrizable Hausdorff space.*

Proof. Let $(\mathbb{P}_n : n \in \mathbb{N})$ be a Cauchy sequence in $(P'(\Omega), d)$. Then for every $m \in \mathbb{N}$, for every m-tuple $(j_1, \dots, j_m)$ in $\mathbb{N}^m$ and for every m-tuple $(s_{j_1}, \dots, s_{j_m}) \in \mathbb{Q}^m \bigcap [0,T]$ the limit $\lim_{\ell\to\infty} \int \prod_{k=1}^m v_{j_k}\left(s_{j_k}, X\left(s_{j_k}\right)\right) d\mathbb{P}_{n_\ell}$ exists. We shall prove that for every every $m \in \mathbb{N}$, for every m-tuple $(j_1, \dots, j_m)$ in $\mathbb{N}^m$ and for every m-tuple $(t_{j_1}, \dots, t_{j_m}) \in [0,T]^m$ the limit

$$\lim_{n\to\infty} \int \prod_{k=1}^m u_{j_k}\left(t_{j_k}, X\left(t_{j_k}\right)\right) d\mathbb{P}_n \quad (3.108)$$

exists for all sequences $(u_j)_{j\in\mathbb{N}}$ in $C_b([0,T] \times E)$. Since, in addition,

$$\lim_{n\to\infty} \lim_{m\to\infty} L_{\mathbb{Q}\bigcap[0,s_k]}\left(\mathbb{P}_n, \mathbb{P}_m\right) = \lim_{m\to\infty} \lim_{n\to\infty} L_{\mathbb{Q}\bigcap[0,s_k]}\left(\mathbb{P}_n, \mathbb{P}_m\right) = 0, \quad (3.109)$$

for all $k \in \mathbb{N}$, it follows that the sequence $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is tight in the sense that the paths $\{X(s) : s \in \mathbb{Q} \bigcap [0, s_k]\}$ are $\mathbb{P}_n$-almost surely totally bounded uniformly in $\mathbb{P}_n$ for all n simultaneously. The latter means that for every $\varepsilon > 0$ there exists $n(\varepsilon) \in \mathbb{N}$ and integers $(\ell_m(\varepsilon))_{m\in\mathbb{N}}$ such that

$$\begin{aligned} &\mathbb{P}_{n_2}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset \bigcup_{j=1}^{\ell_m(\varepsilon)} B\left(x_j, 2^{-m}\right)\right] \\ &\quad\geqslant \left(1-\varepsilon 2^{-m}\right)\mathbb{P}_{n_1}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset E\right] \end{aligned} \tag{3.110}$$

for all n_2, $n_1 \geqslant n(\varepsilon)$, and for all $m \in \mathbb{N}$. By enlarging $\ell_m(\varepsilon)$ we may and do assume that

$$\begin{aligned} &\mathbb{P}_{n}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset \bigcup_{j=1}^{\ell_m(\varepsilon)} B\left(x_j, 2^{-m}\right)\right] \\ &\quad\geqslant \left(1-\varepsilon 2^{-m}\right)\mathbb{P}_{n(\varepsilon)}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset E\right], \end{aligned} \tag{3.111}$$

and

$$\begin{aligned} &\mathbb{P}_{n}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset \bigcup_{j=1}^{\ell_m(\varepsilon)} B\left(x_j, 2^{-m}\right)\right] \\ &\quad\geqslant \left(1-\varepsilon 2^{-m}\right)\mathbb{P}_{n}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset E\right] \end{aligned} \tag{3.112}$$

for all $n \in \mathbb{N}$. It follows that

$$\begin{aligned} &\mathbb{P}_{n}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset \bigcap_{m=1}^{\infty}\bigcup_{j=1}^{\ell_m(\varepsilon)} B\left(x_j, 2^{-m}\right)\right] \\ &\quad\geqslant \left(1-\varepsilon\right)\mathbb{P}_{n}\left[(X(s))_{s\in\mathbb{Q}\bigcap[0,s_k]} \subset E\right], \end{aligned} \tag{3.113}$$

for all $n \in \mathbb{N}$. But then there exists, by Kolmogorov's extension theorem, a probability measure $\mathbb{P}$ such that

$$\lim_{n\to\infty}\int\prod\nolimits_{k=1}^{m} u_{j_k}\left(t_{j_k}, X\left(t_{j_k}\right)\right) d\mathbb{P}_n = \int\prod\nolimits_{k=1}^{m} u_{j_k}\left(t_{j_k}, X\left(t_{j_k}\right)\right) d\mathbb{P}, \tag{3.114}$$

for all $m \in \mathbb{N}$, for all $(j_1, \dots, j_m) \in \mathbb{N}^m$ and for all $(t_{j_1}, \dots, t_{j_m}) \in [0,T]^m$. From the description (3.104) of $P'(\Omega)$ it then readily follows that $\mathbb{P}$ is a member of $P'(\Omega)$. So the existence of the limit in (3.108) remains to be verified, together with the following facts: the limit $\mathbb{P}$ is a martingale solution, and $D([0,\infty], E^{\triangle})$ has full $\mathbb{P}$-measure. Let t be in $\mathbb{Q}\bigcap[0,T]$. Since, for every $j \in \mathbb{N}$, the process

$$v_j(s, X(s)) - v_j(0, X(0)) - \int_0^s (D_1 + L)\, v_j\left(\sigma, X(\sigma)\right) d\sigma, \quad s \in [0,T],$$

is a martingale for the measure $\mathbb{P}_{n_\ell}$, we infer

$$\begin{aligned}&\int \int_0^t \left(D_1+L\right) v_j\left(s, X(s)\right) ds d\mathbb{P}_{n_\ell} \\ &= \int v_j\left(t, X(t)\right) d\mathbb{P}_{n_\ell} - \int v_j\left(0, X(0)\right) d\mathbb{P}_{n_\ell},\end{aligned}$$

and hence the limit $\lim\limits_{\ell\to\infty} \int \int_0^t \left(D_1+L\right) v_j\left(s, X(s)\right) ds d\mathbb{P}_{n_\ell}$ exists. Next let t_0 be in $[0,T]$. Again using the martingale property we see

$$\begin{aligned}&\int v_j\left(t_0, X\left(t_0\right)\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right) \\ &= \int \left(\int_0^t \left(D_1+L\right) v_j\left(s, X(s)\right) ds\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right) \\ &\quad + \int v_j\left(0, X(0)\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right) \\ &\quad - \int \left(\int_{t_0}^t \left(D_1+L\right) v_j\left(s, X(s)\right) ds\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right), \qquad (3.115)\end{aligned}$$

where t is any number in $\mathbb{Q}\bigcap[0,T]$. From (3.115) we infer

$$\begin{aligned}&\left|\int v_j\left(t_0, X\left(t_0\right)\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right)\right| \\ &\leqslant \left|\int \left(\int_0^t \left(D_1+L\right) v_j\left(s, X(s)\right) ds\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right)\right| \\ &\quad + \left|\int v_j\left(0, X(0)\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right)\right| + 2\left|t-t_0\right| \left\|\left(D_1+L\right) v_j\right\|_\infty . \qquad (3.116)\end{aligned}$$

If we let ℓ and k tend to infinity, we obtain

$$\limsup_{\ell,k\to\infty} \left|\int v_j\left(t_0, X\left(t_0\right)\right) d\left(\mathbb{P}_{n_\ell} - \mathbb{P}_{n_k}\right)\right| \leqslant 2\left|t-t_0\right| \left\|\left(D_1+L\right) v_j\right\|_\infty . \qquad (3.117)$$

Consequently for every $s \in [0,T]$ the limit $\lim_{\ell\to\infty} \int v_j\left(s, X(s)\right) d\mathbb{P}_{n_\ell}$ exists. The inequality

$$\begin{aligned}&\left|\int v_j\left(t, X(t)\right) d\mathbb{P}_{n_\ell} - \int v_j\left(t_0, X\left(t_0\right)\right) d\mathbb{P}_{n_\ell}\right| \\ &= \left|\int \int_{t_0}^t \left(D_1+L\right) v_j\left(s, X(s)\right) ds\, d\mathbb{P}_{n_\ell}\right| \\ &\leqslant \left|t-t_0\right| \left\|\left(D_1+L\right) v_j\right\|_\infty\end{aligned}$$

shows that the functions $t \mapsto \lim_{\ell\to\infty} \int v_j(t, X(t))\, d\mathbb{P}_{n_\ell}$, $j \in \mathbb{N}$, are continuous. Since the linear span of $(v_j : j \geqslant 2)$ is dense in $C_b([0,T] \times E)$ for the strict topology, it follows that for every $v \in C_b([0,T] \times E)$ and for every $t \in [0,T]$ the limit

$$t \mapsto \lim_{\ell\to\infty} \int v(t, X(t))\, d\mathbb{P}_{n_\ell}, \quad t \in [0,T], \tag{3.118}$$

exists and that this limit, as a function of t, is continuous. The following step consists in proving that for every $t_0 \in [0,\infty)$ the equality

$$\lim_{t\to t_0} \limsup_{\ell\to\infty} \int |v_j(t, X(t)) - v_j(t_0, X(t_0))|\, d\mathbb{P}_{n_\ell} = 0 \tag{3.119}$$

holds. For $t > s$ the following (in-)equalities are valid:

$$\begin{aligned}
&\left(\int |v_j(t, X(t)) - v_j(s, X(s))|\, d\mathbb{P}_{n_\ell}\right)^2 \leqslant \int |v_j(t, X(t)) - v_j(s, X(s))|^2\, d\mathbb{P}_{n_\ell} \\
&= \int |v_j(t, X(t))|^2\, d\mathbb{P}_{n_\ell} - \int |v_j(s, X(s))|^2\, d\mathbb{P}_{n_\ell} \\
&\quad - 2\Re \int (v_j(t, X(t)) - v_j(s, X(s)))\, \overline{v}_j(s, X(s)) d\mathbb{P}_{n_\ell} \\
&= \int |v_j(t, X(t))|^2\, d\mathbb{P}_{n_\ell} - \int |v_j(s, X(s))|^2\, d\mathbb{P}_{n_\ell} \\
&\quad - 2\Re \int \left(\int_s^t (D_1 + L)\, v_j(\sigma, X(\sigma))\, d\sigma\right) \overline{v}_j(s, X(s)) d\mathbb{P}_{n_\ell} \\
&\leqslant \int |v_j(t, X(t))|^2\, d\mathbb{P}_{n_\ell} - \int |v_j(s, X(s))|^2\, d\mathbb{P}_{n_\ell} \\
&\quad + 2(t-s) \left\| (D_1 + L)\, v_j \right\|_\infty .
\end{aligned} \tag{3.120}$$

Hence (3.118) together with (3.120) implies (3.119). By (3.119), we may apply Kolmogorov's extension theorem to prove that there exists a probability measure $\mathbb{P}$ on $\Omega' := \left(E^{\triangle}\right)^{[0,T]}$ with the property that

$$\int \prod_{k=1}^m v_{j_k}(s_{j_k}, X(s_{j_k}))\, d\mathbb{P} = \lim_{n\to\infty} \int \prod_{k=1}^m v_{j_k}(s_{j_k}, X(s_{j_k}))\, d\mathbb{P}_n \tag{3.121}$$

holds for all $m \in \mathbb{N}$ and for all $(s_{j_1}, \dots, s_{j_m}) \in [0,T]^m$. It then follows that the equality in (3.121) is also valid for all m-tuples $f_1, \dots, f_m$ in $C_b\left([0,T] \times E^{\triangle}\right)$ instead of for $v_{j_1}, \dots, v_{j_m}$. This is true because the linear span of the sequence $(v_j)_{j\in\mathbb{N}}$ is $\mathcal{T}_\beta$-dense in $C_b\left([0,T] \times E^{\triangle}\right)$. In addition we conclude that the processes

$$f(t, X(t)) - f(0, X(0)) - \int_0^t (D_1 + L)\, f(s, X(s))\, ds,$$

$t \in [0,T]$, $f \in D^{(1)}(L)$, are $\mathbb{P}$-martingales. We still have to show that $D([0,T], E^{\triangle})$ has $\mathbb{P}$-measure 1. From (3.119) it essentially follows that set of $\omega \in \left(E^{\triangle}\right)^{[0,T]}$ for which the left and right hand limits exist in $E^{\triangle}$ has "full" $\mathbb{P}$-measure. First let $f \geqslant 0$ be in $C_b([0,T] \times E)$. Then the process

$$[G_\lambda f](t) := \mathbb{E}\left[\int_t^\infty e^{-\lambda\sigma} f(\sigma \wedge T, X(\sigma \wedge T))\, d\sigma \mid \mathcal{F}_t^0\right]$$

is a $\mathbb{P}$-supermartingale with respect to the filtration $\left(\mathcal{F}_t^0\right)_{t\in[0,T]}$. It follows that the limits $\lim_{t\uparrow t_0} [G_\lambda f](t)$ and $\lim_{t\downarrow t_0} [G_\lambda f](t)$ both exist $\mathbb{P}$-almost surely for all $t_0 \geqslant 0$ and for all $f \in C_b([0,T] \times E)$. In particular these limits exist $\mathbb{P}$-almost surely for all $f \in D^{(1)}(L)$. By the martingale property it follows that, for $f \in D^{(1)}(L)$,

$$\begin{aligned}
&\left|f(t, X(t)) - \lambda e^{\lambda t}[G_\lambda f](t)\right| \\
&= \left|\lambda e^{\lambda t}\mathbb{E}\left[\int_t^\infty e^{-\lambda\sigma}\left(f(\sigma \wedge T, X(\sigma \wedge T)) - f(t, X(t))\right) d\sigma \mid \mathcal{F}_t^0\right]\right| \\
&= \left|\lambda e^{\lambda t}\mathbb{E}\left[\int_t^\infty e^{-\lambda\sigma}\left(\int_t^\sigma (D_1 + L) f(s, X(s))\, ds\right) d\sigma \mid \mathcal{F}_t^0\right]\right| \\
&\leqslant \lambda e^{\lambda t}\int_t^\infty e^{-\lambda\sigma}(\sigma - t)\left\|(D_1 + L) f\right\|_\infty d\sigma = \lambda^{-1}\left\|(D_1 + L) f\right\|_\infty. \quad (3.122)
\end{aligned}$$

Consequently, we may conclude that, for all s, $t \geqslant 0$,

$$\begin{aligned}
&|f(t, X(t)) - f(s, X(s))| \\
&\leqslant 2\lambda^{-1}\left\|(D_1 + L) f\right\|_\infty + \left|\lambda e^{\lambda t}[G_\lambda f](t) - \lambda e^{\lambda s}[G_\lambda f](s)\right|, \qquad (3.123)
\end{aligned}$$

Again using (3.111), (3.112) and (3.113) it follows that the path

$$\left\{X(s) : s \in \mathbb{Q} \bigcap [0,t],\, X(t) \in E\right\}$$

is $\mathbb{P}$-almost surely totally bounded. By separability and $\mathcal{T}_\beta^{\mathcal{T}}$-density of $D^{(1)}(L)$ it follows that the limits $\lim_{t\downarrow s} X(t)$ and $\lim_{s\uparrow t} X(s)$ exist in E $\mathbb{P}$-almost surely for all s respectively $t \in [0,T]$, for which $X(s)$ respectively $X(t)$ belongs to E. See the arguments which led to (3.14) and (3.15) in the proof of Theorem 2.9. Put $Z(s)(\omega) = \lim_{t\downarrow s, t\in\mathbb{Q}\bigcap[0,T]} X(t)(\omega)$. Then, for $\mathbb{P}$-almost all ω the mapping $s \mapsto Z(s)(\omega)$ is well-defined, possesses left limits in $t \in [0,T]$ for those paths $\omega \in \Omega$ for which $\omega(t) \in E$ and is right continuous. In addition we have

$$\begin{aligned}
\mathbb{E}\left[f(s, Z(s))g(s)\right] &= \mathbb{E}\left[f(s, X(s+))g(s, X(s))\right] \\
&= \lim_{t\downarrow s}\mathbb{E}\left[f(t, X(t))g(s, X(s))\right] = \mathbb{E}\left[f(s, X(s))g(s, X(s))\right],
\end{aligned}$$

for all f, $g \in C_b([0,T] \times E)$ and for all $s \in [0,T]$: see (3.119). But then we may conclude that $X(s) = Z(s)$ $\mathbb{P}$-almost surely for all $s \in [0,T]$. Hence we may replace X with Z and consequently (see the arguments in the proof of Theorem 2.9, and see Theorem 9.4 in [Blumenthal and Getoor (1968)], p. 49)

$$\mathbb{P}[\Omega] = 1, \text{ and so } P \in P'(\Omega) = P_0'(\Omega) \tag{3.124}$$

where $\Omega = D\left([0,T] \times E^{\triangle}\right)$. For the definition of $D\left([0,T] \times E^{\triangle}\right)$ see Definition 2.9, and for the definition of $P'(\Omega)$, and $P_0'(\Omega)$ the reader is referred to (3.105) and (3.104).

We also have to prove the separability. Denote by Convex the collection of all mappings

$$\alpha : \mathcal{P}_f(\mathbb{N}) \times \mathcal{P}_f\left(\mathbb{Q}\bigcap[0,T]\right) \to \mathbb{Q}\bigcap[0,1],$$

which take only finitely many non-zero values, such that

$$\sum_{\Lambda' \in \mathcal{P}_f(\mathbb{N})} \alpha\left(\Lambda', \Lambda\right) = 1, \quad \Lambda \in \mathcal{P}_f\left(\mathbb{Q}\bigcap[0,T]\right),$$

and let $\{w_{\Lambda'} : \Lambda' \in \mathcal{P}_f(\mathbb{N})\}$ be a countable family of functions from $\mathbb{Q}\bigcap[0,T]$ to $E^{\triangle}$ such that for every finite subset $\Lambda = \{s_{j_1}, \dots, s_{j_n}\} \in \mathcal{P}_f(\mathbb{Q}\bigcap[0,T])$ the collection

$$\left\{\left(w_{\Lambda'}(s_{j_1}), \dots, w_{\Lambda'}(s_{j_n})\right) : \Lambda' \in \mathcal{P}_f(\mathbb{N})\right\}$$

is dense in $\left(E^{\triangle}\right)^{(s_{j_1},\dots,s_{j_n})} = \left(E^{\triangle}\right)^{\Lambda}$. For example the value of $w_{\Lambda'}(s_{j_\ell})$ could be x_{k_ℓ}, $1 \leqslant \ell \leqslant n$, where $\Lambda' = (k_1, \dots, k_n)$. Here $(x_k)_{k\in\mathbb{N}}$ is a dense sequence in $\mathbb{E}^{\triangle}$. The countable collection of probability measures

$$\{\mathbb{P}_{\alpha,w,\Lambda} : \alpha \in \text{Convex}, \ \Lambda \in \mathcal{P}_f(\mathbb{N})\}$$

determined by

$$\mathbb{E}_{\alpha,w,\Lambda}\left[F\left((s, X(s))_{s\in\Lambda}\right)\right] = \sum_{\Lambda' \in \mathcal{P}_f(\mathbb{N})} \alpha\left(\Lambda', \Lambda\right) F\left((s, w_{\Lambda'}(s))_{s\in\Lambda}\right)$$

is dense in $P(\Omega)$ endowed with the metric d_L. Since $P'(\Omega)$ is a closed subspace of $P(\Omega)$, it is separable as well.

Finally we observe that $X(t) \in E$, $\tau < s < t$, implies $X(s) \in E$. This follows from the assumption that the Skorohod space $D\left([0,T], E^{\triangle}\right)$ is the sample space on which we consider the martingale problem: see Definition 2.9. In particular it is assumed that $X(s) = \triangle$, $\tau < s \leqslant t$, implies $X(t) = \triangle$, and $L(\rho)f(\rho,\cdot)(X(\rho)) = 0$ for $s < \rho < t$. Consequently, once we have $X(s) = \triangle$, and $t \in (s,T]$, then $X(t) = \triangle$, and by transposition $X(t) \in E$, $s \in [\tau, t)$ implies $X(s) \in E$.

This completes the proof of Proposition 3.6. □

In the following proposition we see that under the condition of λ-dominance the function $(\tau, s, x) \mapsto \mathbb{E}_{\tau,x}\left[u\left(s, X(s)\right)\right]$ is continuous whenever the function $(s, x) \mapsto u(s, x)$ belongs to $C_b\left([0, T] \times E\right)$, and the martingale problem is well-posed.

Proposition 3.7. *Suppose that for every $(\tau, x) \in [0, T] \times E$ the martingale problem is uniquely solvable. In addition, suppose that there exists $\lambda > 0$ such that the operator $D_1 + L$ is sequentially λ-dominant: see Definition 4.3. Define the map $F : P''(\Omega) \to [0, T] \times E$ by $F(\mathbb{P}) = (\tau, x)$, where $\mathbb{P} \in P''(\Omega)$ is such that $\mathbb{P}(X(s) = x) = 1$, for $s \in [0, \tau]$. Then F is a homeomorphism from the Polish space $P''(\Omega)$ onto $[0, T] \times E$. In fact it follows that for every $u \in C_b\left([0, T] \times E\right)$ and for every $s \in [\tau, T]$, the function $(\tau, s, x) \mapsto \mathbb{E}_{\tau,x}\left[u\left(s, X(s)\right)\right]$, $0 \leqslant \tau \leqslant s \leqslant T$, $x \in E$, is continuous.*

Here $P''(\Omega) := \{\mathbb{P}_{\tau,x} : (\tau, x) \in [0, T] \times E\}$.

Proof. Since the martingale problem is uniquely solvable for every $(\tau, x) \in [0, T] \times E$ the map F is a one-to-one map from the Polish space $\left(P''(\Omega), d_L\right)$ onto $[0, T] \times E$ (see Proposition 3.6 and (3.107)). Let for $(\tau, x) \in [0, T] \times E$ the probability $\mathbb{P}_{\tau,x}$ be the unique solution to the martingale problem:

(i) For every $f \in D^{(1)}(L)$ the process

$$f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t (D_1 + L) f(s, X(s)) ds, \quad t \in [\tau, T],$$

is a $\mathbb{P}_{\tau,\mu}$-martingale;

(ii) The $\mathbb{P}_{\tau,\mu}$-distribution of $X(\tau)$ is the measure μ. If $\mu = \delta_x$, then we write $\mathbb{P}_{\tau,\delta_x} = \mathbb{P}_{\tau,x}$, and $\mathbb{P}_{\tau,x}\left[X(\tau) = x\right] = 1$.

Then, by definition $F\left(\mathbb{P}_{\tau,x}\right) = (\tau, x)$, $(\tau, x) \in [0, T] \times E$. Moreover, since for every $(\tau, x) \in [0, T] \times E$ the martingale problem is uniquely solvable we see $P'(\Omega) = \{\mathbb{P}_{\tau,\mu} : (\tau, \mu) \in [0, T] \times P(E)\}$. Here $P(E)$ is the collection of Borel probability measures on E. This equality of probability spaces can be seen as follows. If the measure $\mathbb{P}_{\tau,\mu}$ is a solution to the martingale problem, then it is automatically a member of $P'(\Omega)$. If $\mathbb{P}$ is a member of $P'(\Omega)$ which starts at time τ, then by uniqueness of solutions we have:

$$\mathbb{P}\left[A \mid \sigma\left(X(\tau)\right)\right] \big|_{X(\tau)=x} = \mathbb{P}_{\tau,x}\left[A\right], \quad A \in \mathcal{F}_T^\tau. \tag{3.125}$$

In addition, $\mathbb{P} = \mathbb{P}_{\tau,\mu}$, where $\mu(B) = \mathbb{P}\left[X(\tau) \in B\right]$, $B \in \mathcal{E}$. Let $\left((t_\ell, x_\ell)\right)_{\ell\in\mathbb{N}}$ be a sequence in $[0, T] \times E$ with the property that $\lim_{\ell\to\infty} d_L\left(\mathbb{P}_{t_\ell,x_\ell}, \mathbb{P}_{\tau,x}\right) = 0$ for some $(\tau, x) \in [0, T] \times E$. Then for some random variable ε

the orbit $\{(s,X(s)) : s \in (\tau-\varepsilon,\tau+\varepsilon)\}$ is totally bounded $\mathbb{P}_{t_\ell,x_\ell}$-almost surely for all t_ℓ and τ simultaneously. It follows that the sequence $\{x_\ell = X(t_\ell) : \ell \in \mathbb{N}\}\bigcup\{x\}$ is contained in a compact subset of E. Then $\lim_{\ell\to\infty} |v_j(t_\ell,x_\ell) - v_j(\tau,x)| = 0$, for all $j \in \mathbb{N}$, where, as above, the span of the sequence $(v_j)_{j\geqslant 2}$ is $\mathcal{T}_\beta$-dense in $C([0,T]\times E)$. It follows that $\lim_{\ell\to\infty}(t_\ell,x_\ell) = (t,x)$ in $[0,T]\times E$. Consequently the mapping F is continuous. Since F is a continuous bijective map from one Polish space

$$P''(\Omega) := \{\mathbb{P}_{\tau,x} : (\tau,x) \in [0,T]\times E\} \tag{3.126}$$

onto another such space $[0,T]\times E$, its inverse is continuous as well. Among other things this implies that, for every $s \in \mathbb{Q}\bigcap[0,\infty)$ and for every $j \geqslant 2$, the function $(\tau,x) \mapsto \int v_j(s,X(s))\,d\mathbb{P}_{\tau,x}$ belongs to $C_b([0,T]\times E)$. Since the linear span of the sequence $(v_j : j \geqslant 2)$ is $\mathcal{T}_\beta$-dense in $C_b([0,T]\times E)$ it also follows that for every $v \in C_b([0,T]\times E)$, the function $(\tau,x) \mapsto \int v(s,X(s))\,d\mathbb{P}_{\tau,x}$ belongs to $C_b([0,T]\times E)$. Next let $s_0 \in [0,T]$ be arbitrary. For every $j \geqslant 2$ and every $s \in \mathbb{Q}\bigcap[0,T]$, $s > s_0$, we have by the martingale property:

$$\begin{aligned}
&\sup_{(\tau,x)\in[0,s_0]\times E} |\mathbb{E}_{\tau,x}(v_j(s,X(s))) - \mathbb{E}_{\tau,x}(v_j(s_0,X(s_0)))| \\
&= \sup_{(\tau,x)\in[0,s_0]\times E} \left|\int_{s_0}^{s} \mathbb{E}_{\tau,x}(Lv_j(\sigma,X(\sigma)))\,d\sigma\right| \\
&\leqslant (s-s_0)\,\|(D_1+L)\,v_j\|_\infty .
\end{aligned} \tag{3.127}$$

Consequently, for every $s \in [0,T]$, the function $(\tau,x) \mapsto \mathbb{E}_{\tau,x}[v_j(s,X(s))]$, $j \geqslant 1$, belongs to $C_b([0,T]\times E)$. It follows that, for every $v \in C_b([0,T]\times E)$ and every $s \in [0,T]$, the function $(\tau,x) \mapsto \mathbb{E}_{\tau,x}[v(s,X(s))]$ belongs to $C_b([0,T]\times E)$. These arguments also show that the function $(\tau,s,x) \mapsto \mathbb{E}_{\tau,x}[v(s,X(s))]$, $0 \leqslant \tau \leqslant s \leqslant T$, $x \in E$, is continuous for every $v \in C_b([0,T]\times E)$. The continuity in the three variables (τ,s,x) requires the sequential λ-dominance of the operator D_1+L for some $\lambda > 0$. The arguments run as follows. Using the Markov process

$$\{(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x}),(X(t),\tau\leqslant t\leqslant T),(\vee_t:\tau\leqslant t\leqslant T),(E,\mathcal{E})\} \tag{3.128}$$

we define the semigroup $\{S(\rho) : \rho \geqslant 0\}$ as follows

$$\begin{aligned}
S(\rho)f(\tau,x) &= P(\tau,(\rho+s)\wedge T)\,f((\rho+s)\wedge T,\cdot)(x) \\
&= \mathbb{E}_{\tau,x}[f((\rho+s)\wedge T, X((\rho+s)\wedge T))].
\end{aligned} \tag{3.129}$$

Here $(\tau,x) \in [0,T]\times E$, $\rho \geqslant 0$, and $f \in C_b([0,T]\times E)$. Let $\lambda > 0$ and $f \in C_b([0,T]\times E)$. We want to establish a relationship between the semigroup $\{S(\rho) : \rho \geqslant 0\}$ and the operator $D_1 + L$. Therefore we first prove that the process

$$\begin{aligned} t \mapsto & e^{-\lambda t} f\left(t \wedge T, X\left(t \wedge T\right)\right) - e^{-\lambda \tau} f\left(\tau, X(\tau)\right) \\ & + \int_\tau^t e^{-\lambda\rho} \left(\lambda I - D_1 - L\right) f\left(\rho \wedge T, X\left(\rho \wedge T\right)\right) d\rho, \quad t \geqslant \tau, \end{aligned} \tag{3.130}$$

is a $\mathbb{P}_{\tau,x}$-martingale with respect to the filtration $(\mathcal{F}_t^\tau)_{t\in[\tau,T]}$. Let $\tau \leqslant s < t \leqslant T$, and $y \in E$. Then integration by parts shows:

$$\begin{aligned} & e^{-\lambda t} f\left(t, X(t)\right) - e^{-\lambda s} f\left(s, X(s)\right) + \int_s^t e^{-\lambda\rho} \left(\lambda I - D_1 - L\right) f\left(\rho, X(\rho)\right) d\rho \\ & = e^{-\lambda t} f\left(t, X(t)\right) - e^{-\lambda s} f\left(s, X(s)\right) + \lambda \int_s^t e^{-\lambda\rho} f\left(\rho, X(\rho)\right) d\rho \\ & - e^{-\lambda t} \int_s^t \left(D_1 + L\right) f\left(\rho, X(\rho)\right) d\rho - \lambda \int_s^t e^{-\lambda\rho} \left(f\left(\rho, X(\rho)\right) - f\left(s, X(s)\right)\right) d\rho. \end{aligned} \tag{3.131}$$

Then by the martingale property the $\mathbb{P}_{s,y}$-expectation of the expression in (3.131) is zero. By employing the Markov property we obtain

$$\begin{aligned} & \mathbb{E}_{\tau,x} \left[e^{-\lambda t} f\left(t, X(t)\right) - e^{-\lambda \tau} f\left(\tau, X(\tau)\right) \right. \\ & \qquad \left. + \int_\tau^t e^{-\lambda\rho} \left(\lambda I - D_1 - L\right) f\left(\rho, X(\rho)\right) d\rho \;\Big|\; \mathcal{F}_s^\tau \right] \\ & - \left(e^{-\lambda s} f\left(s, X(s)\right) - e^{-\lambda \tau} f\left(\tau, X(\tau)\right) \right. \\ & \qquad \left. + \int_\tau^s e^{-\lambda\rho} \left(\lambda I - D_1 - L\right) f\left(\rho, X(\rho)\right) d\rho \right) \\ & = \mathbb{E}_{\tau,x} \left[e^{-\lambda t} f\left(t, X(t)\right) - e^{-\lambda s} f\left(s, X(s)\right) \right. \\ & \qquad \left. + \int_s^t e^{-\lambda\rho} \left(\lambda I - D_1 - L\right) f\left(\rho, X(\rho)\right) d\rho \;\Big|\; \mathcal{F}_s^\tau \right] \end{aligned}$$

(Markov property)

$$\begin{aligned} & = \mathbb{E}_{s,X(s)} \left[e^{-\lambda t} f\left(t, X(t)\right) - e^{-\lambda s} f\left(s, X(s)\right) \right. \\ & \qquad \left. + \int_s^t e^{-\lambda\rho} \left(\lambda I - D_1 - L\right) f\left(\rho, X(\rho)\right) d\rho \right] = 0 \end{aligned} \tag{3.132}$$

where in the final step in (3.132) we used the fact that the $\mathbb{P}_{s,y}$-expectation, $y \in E$, of the expression in (3.131) vanishes. Consequently, the process in

(3.130) is a $\mathbb{P}_{\tau,x}$-martingale. From the fact that the process in (3.130) is a $\mathbb{P}_{\tau,x}$-martingale we infer by taking expectations that for $t \geqslant 0$

$$\begin{aligned} &e^{-\lambda(t+\tau)}\mathbb{E}_{\tau,x}\left[f\left((t+\tau)\wedge T, X\left((t+\tau)\wedge T\right)\right)\right] - e^{-\lambda\tau}\mathbb{E}_{\tau,x}\left[f\left(\tau, X(\tau)\right)\right] \\ &\quad + \int_\tau^{t+\tau} e^{-\lambda\rho}\mathbb{E}_{\tau,x}\left[\left(\lambda I - D_1 - L\right) f\left(\rho\wedge T, X\left(\rho\wedge T\right)\right)\right] d\rho = 0. \end{aligned} \tag{3.133}$$

The equality in (3.133) is equivalent to

$$\begin{aligned} &\mathbb{E}_{\tau,x}\left[f\left(\tau, X(\tau)\right)\right] - e^{-\lambda t}\mathbb{E}_{\tau,x}\left[f\left((t+\tau)\wedge T, X\left((t+\tau)\wedge T\right)\right)\right] \\ &= \int_\tau^{t+\tau} e^{-\lambda(\rho-\tau)}\mathbb{E}_{\tau,x}\left[\left(\lambda I - D_1 - L\right) f\left(\rho\wedge T, X\left(\rho\wedge T\right)\right)\right] d\rho = 0. \end{aligned} \tag{3.134}$$

In terms of the semigroup $\{S(\rho) : \rho \geqslant 0\}$ the equality in (3.134) can be rewritten as follows:

$$f(\tau,x) - e^{-\lambda t}S(t)f\left(\tau,x\right) = \int_0^t e^{-\lambda\rho}S(\rho)\left(\lambda I - D_1 - L\right) f\left(\tau,x\right)\, d\rho. \tag{3.135}$$

By letting $t \to \infty$ in (3.135) we see

$$\begin{aligned} f\left(\tau,x\right) &= \int_0^\infty e^{-\lambda\rho}S(\rho)\left(\lambda I - D_1 - L\right) f\left(\tau,x\right)\, d\rho \\ &= R(\lambda)\left(\lambda I - D_1 - L\right) f\left(\tau,x\right)\, d\rho \end{aligned} \tag{3.136}$$

where the definition of $R(\lambda)$, $\lambda > 0$, is self-explanatory. Define the operator $L^{(1)} : D\left(L^{(1)}\right) = R(\lambda)C_b\left([0,T]\times E\right) \to C_b\left([0,T]\times E\right)$ by $L^{(1)}R(\lambda)f = \lambda R(\lambda)f - f$, $f \in C_b\left([0,T]\times E\right)$. Then by definition we see $\left(\lambda I - L^{(1)}\right)R(\lambda)f = f$, and thus $R(\lambda)\left(\lambda I - L^{(1)}\right)R(\lambda)f = R(\lambda)f$, $f \in C_b\left([0,T]\times E\right)$. Put $g = \left(\lambda I - L^{(1)}\right)R(\lambda)f - f$. Then by the resolvent identity we see that $R(\alpha)g = 0$ for all $\alpha > 0$, and hence $S(\rho)g\left(\tau,x\right) = \mathbb{E}_{\tau,x}\left[g\left((\rho+\tau)\wedge T, X\left((\rho+\tau)\wedge T\right)\right)\right] = 0$ for all $\rho > 0$. By the right-continuity of the process $\rho \mapsto X(\rho)$, we see that $g = 0$. Consequently, $\left(\lambda I - L^{(1)}\right)R(\lambda)f - f = 0$, $f \in C_b\left([0,T]\times E\right)$. If $f \in D^{(1)}(L)$, then (3.136) reads $f = R(\lambda)\left(\lambda I - D_1 - L\right)f$, and hence $f \in D\left(L^{(1)}\right)$, and $\left(\lambda I - L^{(1)}\right)f = \left(\lambda I - D_1 - L\right)f$, or what amounts to the same $f \in D\left(L^{(1)}\right)$, and $L^{(1)}f = D_1 f + Lf$. In other words the operator $L^{(1)}$ extends $D_1 + L$. As in (2.42) define the sub-additive mapping $U_\lambda^1 : C_b\left([0,T]\times E, \mathbb{R}\right) \to L^\infty\left([0,T]\times E, \mathbb{R}\right)$ by

$$U_\lambda^1 f = \sup_{K\in\mathcal{K}(E)} \inf_{g\in D^{(1)}(L)} \left\{g \geqslant f\mathbf{1}_K : \lambda g - D_1 g - Lg \geqslant 0\right\}. \tag{3.137}$$

Since $L^{(1)}$ extends $D_1 + L$, from (3.137) we get

$$U_\lambda^1 f \geqslant \sup_{K\in\mathcal{K}(E)} \inf_{g\in D\left(L^{(1)}\right)} \left\{ g \geqslant f\mathbf{1}_K : \lambda g - L^{(1)} g \geqslant 0 \right\}. \tag{3.138}$$

Then, as explained in Proposition 2.4, formula (2.49), we have

$$\sup\left\{ \left(\mu R(\lambda+\mu)\right)^k f;\, \mu > 0,\, k \in \mathbb{N} \right\} \leqslant U_\lambda^1(f), \quad f \in C_b\left([0,T]\times E,\mathbb{R}\right). \tag{3.139}$$

As is indicated in the proof of (iii) $\Longrightarrow$ (i) of Theorem 4.3 the following equality also holds:

$$\sup\left\{ \left(\mu R(\lambda+\mu)\right)^k f;\, \mu > 0,\, k \in \mathbb{N} \right\} = \sup\left\{ e^{-\lambda\rho} S(\rho) f : \rho \geqslant 0 \right\}, \tag{3.140}$$

where $f \in C_b\left([0,T]\times E,\mathbb{R}\right)$. For this observation the reader is referred to the formulas (4.20), (4.21) and (4.22). Next let $(f_n)_{n\in\mathbb{N}} \subset C_b\left([0,T]\times E\right)$ be a sequence which decreases pointwise to zero. Using the sequential λ-dominance of the operator $D_1 + L$ and using the equality in (3.139) and the inequality in (3.140) we see that $\sup_{\rho\geqslant 0} e^{-\lambda\rho} S(\rho) f_n(\tau,x)$ decreases to zero uniformly on compact subsets of $[0,T]\times E$: see Definition 4.3. From Proposition 2.3 it follows that the semigroup $\left\{e^{-\lambda\rho}S(\rho) : \rho \geqslant 0\right\}$ is $\mathcal{T}_\beta$-equi-continuous. In addition, by the arguments above, every operator $S(\rho)$, $\rho \geqslant 0$, assigns to a function $f \in D^{(1)}(L) = D(D_1) \bigcap D(L)$ a function $S(\rho) f \in C_b\left([0,T]\times E\right)$. By the $\mathcal{T}_\beta$-continuity of $S(\rho)$, and by the fact that $D^{(1)}(L)$ is $\mathcal{T}_\beta$-dense in $C_b\left([0,T]\times E\right)$, the mapping $S(\rho)$ extends to a $\mathcal{T}_\beta$-continuous linear continuous operator from $C_b\left([0,T]\times E\right)$ to itself. This extension is again denoted by $S(\rho)$. In addition, for $v \in D^{(1)}(L)$, the function $(\tau,\rho,x) \mapsto S(\rho)v(\tau,x)$ is continuous on $[0,T]\times[0,\infty)\times E$; see (3.127). Fix $f \in C_b\left([0,T]\times E\right)$. Using the sequential λ-dominance and its consequence of $\mathcal{T}_\beta$-equi-continuity of the semigroup $\left\{e^{-\lambda\rho}S(\rho) : \rho \geqslant 0\right\}$ we see that the function $(\tau,s,x) \mapsto S(\rho)f(\tau,x)$ is continuous on $[0,T]\times[0,\infty)\times E$, and hence the same is true for the function $(\tau,s,x) \mapsto \mathbb{E}_{\tau,x}\left[f\left(s,X(s)\right)\right]$. Here we again used the $\mathcal{T}_\beta$-density of $D^{(1)}(L)$ in $C_b\left([0,T]\times E\right)$.

This completes the proof of Proposition 3.7. □

Notice that in the proof of the implication (iii) $\Longrightarrow$ (i) of Theorem 4.3 arguments very similar to the ones in the final part of the proof of Proposition 3.7 will be employed.

The following corollary establishes an important relation between unique solutions to the martingale problem and Feller semigroups.

Corollary 3.2. *Suppose that the martingale problem is well posed for the operator $D_1 + L$, and that the operator $D_1 + L$ is sequentially λ-dominant* *for some* $\lambda > 0$. *Let* $\{(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}) : (\tau, x) \in [0,T] \times E\}$ *be the solutions to the martingale problem starting at* x *at time* τ. *Let the process in (3.128) be the corresponding Markov process, and let the semigroup* $\{S(\rho) : \rho \geqslant 0\}$, *as defined in (3.129), be the corresponding Feller semigroup. Then this semigroup is* $\mathcal{T}_\beta$*-equi-continuous, and its generator extends* $D_1 + L$.

Proof. From Proposition 2.3 it follows that for some $\lambda > 0$ the semigroup $\left\{e^{-\lambda\rho} S(\rho) : \rho \geqslant 0\right\}$ is $\mathcal{T}_\beta$-equi-continuous: see the proof of Proposition 3.7. Since $S(\rho) = S(T)$ for $\rho \geqslant T$, we see that the semigroup $\{S(\rho) : \rho \geqslant 0\}$ itself is $\mathcal{T}_\beta$-equi-continuous. Moreover, it is a Feller semigroup in the sense that it consists of $\mathcal{T}_\beta$-continuous linear operators, and $\mathcal{T}_\beta\text{-}\lim_{t\to s} S(t)f = S(s)f$, $f \in C_b([0,T] \times E)$. From the proof of Proposition 3.7 it follows that the generator of the semigroup $\{S(\rho) : \rho \geqslant 0\}$ extends $D_1 + L$.

This proves Corollary 3.2. □

The proof of the following proposition may be copied from [Ikeda and Watanabe (1998)], Theorem 5.1. p. 205. For completeness we insert a proof as well.

Proposition 3.8. *Suppose that for every* $(\tau, x) \in [0,T] \times E$ *the martingale problem, posed on the Skorohod space* $D\left([0,T], E^\triangle\right)$ *as follows,*

(i) *For every* $f \in D^{(1)}(L)$ *the process*

$$f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t (D_1 + L) f(s, X(s)) ds, \quad t \in [\tau, T],$$

is a $\mathbb{P}$*-martingale;*

(ii) $\mathbb{P}(X(\tau) = x) = 1$,

has a unique solution $\mathbb{P} = \mathbb{P}_{\tau,x}$.

Then the process

$$\{(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}), (X(t), \tau \leqslant t \leqslant T), (\vee_t : \tau \leqslant t \leqslant T), (E, \mathcal{E})\}, \tag{3.141}$$

is a strong Markov process with respect to the right-continuous filtration $\left(\mathcal{F}_{t+}^\tau\right)_{t\in[\tau,T]}$.

For the definition of $\mathcal{F}_{S+}^\tau$ the reader is referred to (2.97) in Remark 2.8; also see (2.85) in Definition 2.14.

Proof. Fix $(\tau, x) \in [0,T] \times E$ and let S be a stopping time and choose a realization $A \mapsto \mathbb{E}_{\tau,x}\left[\mathbf{1}_A \circ \vee_S \,\middle|\, \mathcal{F}_{S+}^\tau\right]$, $A \in \mathcal{F}_T^\tau$. Fix any $\omega \in \Omega$ for which

$$A \mapsto Q_{s,y}[A] := \mathbb{E}_{\tau,x}\left[\mathbf{1}_A \circ \vee_S \,\middle|\, \mathcal{F}_{S+}^\tau\right](\omega),$$

is defined for all $A \in \mathcal{F}_T^\tau$. Here, by definition, $(s,y) = (S(\omega), \omega(S(\omega)))$. Notice that this construction can be performed for $\mathbb{P}_{\tau,x}$-almost all ω. Let f be in $D^{(1)}(L) = D(D_1) \bigcap D(L)$ and fix $T \geqslant t_2 > t_1 \geqslant 0$. Moreover, fix $C \in \mathcal{F}_{t_1}^\tau$. Then $\vee_S^{-1}(C)$ is a member of $\mathcal{F}_{t_1 \vee S+}^{\tau \vee S}$. Put

$$M_f(t) = f(t, X(t)) - f(X(\tau)) - \int_\tau^t (D_1 + L) f(s, X(s)) ds, \quad t \in [\tau, T].$$

We have

$$\mathbb{E}_{s,y}\left[M_f(t_2)\mathbf{1}_C\right] = \mathbb{E}_{s,y}\left[M_f(t_1)\mathbf{1}_C\right]. \tag{3.142}$$

We also have

$$\int \left(f\left(t_2, X(t_2)\right) - f(\tau, X(\tau)) - \int_\tau^{t_2} Lf(X(s))ds\right) \mathbf{1}_C dQ_{s,y} \tag{3.143}$$

$$\begin{aligned}
&= \mathbb{E}_{\tau,x}\Bigg[\Bigg(f\left(t_2 \vee S, X\left(t_2 \vee S\right)\right) - f(S, X(S)) \\
&\qquad - \int_\tau^{t_2} (D_1 + L) f\left(s \vee S, X(s \vee S)\right) ds\Bigg) \left(\mathbf{1}_C \circ \vee_S\right) \Bigm| \mathcal{F}_{S+}^\tau\Bigg](\omega) \\
&= \mathbb{E}_{\tau,x}\Bigg[\Bigg(f\left(t_2 \vee S, X(t_2 \vee S)\right) - f(S, X(S)) \\
&\qquad - \int_S^{t_2 \vee S} (D_1 + L) f\left(X(s)\right) ds\Bigg) \left(\mathbf{1}_C \circ \vee_S\right) \Bigm| \mathcal{F}_{S+}^\tau\Bigg](\omega) \\
&= \mathbb{E}_{\tau,x}\Bigg[\mathbb{E}_{\tau,x}\Bigg[\Bigg(f\left(t_2 \vee S, X(t_2 \vee S)\right) - f(S, X(S)) \\
&\quad - \int_S^{t_2 \vee S} (D_1 + L) f\left(s, X(s)\right) ds\Bigg) \Bigm| \mathcal{F}_{t_1 \vee S+}^\tau\Bigg] \mathbf{1}_C \circ \vee_S \Bigm| \mathcal{F}_S^\tau\Bigg](\omega).
\end{aligned} \tag{3.144}$$

By Doob's optional sampling theorem, and right-continuity of paths, the process

$$f\left(t \vee S, X(t \vee S)\right) - f(S, X(S)) - \int_S^{t \vee S} (D_1 + L) f\left(s, X(s)\right) ds$$

is a $\mathbb{P}_{\tau,x}$-martingale with respect to the filtration consisting of the σ-fields $\mathcal{F}_{t \vee S+}^\tau$, $t \in [\tau, T]$. So from (3.143) we obtain:

$$\int \left(f(t_2, X(t_2)) - f(\tau, X(\tau)) - \int_\tau^{t_2} Lf(X(s))ds\right) \mathbf{1}_C dQ_{s,y}$$

$$= \mathbb{E}_{\tau,x}\left[\left(f\left(t_1 \vee S, X(t_1 \vee S)\right) - f(S, X(S))\right.\right. \tag{3.145}$$
$$\left.\left. - \int_S^{t_1 \vee S} \left(D_1 + L\right) f\left(s, X(s)\right) ds\right) \left(\mathbf{1}_C \circ \vee_S\right) \,\Big|\, \mathcal{F}_{S+}^{\tau}\right](\omega)$$
$$= \int \left(f\left(t_1, X(t_1)\right) - f\left(\tau, X(\tau)\right) - \int_\tau^{t_1} \left(D_1 + L\right) f(s, X(s)) ds\right) \mathbf{1}_C dQ_{s,y}.$$

It follows that, for $f \in D(L)$, the process $M_f(t)$ is a $\mathbb{P}_{s,y}$- as well as a $Q_{s,y}$-martingale. Since $\mathbb{P}_{s,y}[X(s) = y] = 1$ and since

$$\begin{aligned} Q_{s,y}[X(s) = y] &= \mathbb{E}_{\tau,x}\left[\mathbf{1}_{\{X(S)=y\}} \circ \vee_S \,\Big|\, \mathcal{F}_{S+}^{\tau}\right](\omega) \\ &= \mathbb{E}_{\tau,x}\left[\mathbf{1}_{\{X(S)=y\}} \,\Big|\, \mathcal{F}_{S+}^{\tau}\right](\omega) = \mathbf{1}_{\{X(S)=y\}}(\omega) = 1, \end{aligned} \tag{3.146}$$

we conclude that the probabilities $\mathbb{P}_{s,y}$ and $Q_{s,y}$ are the same. Equality (3.146) follows, because, by definition, $y = X(S)(\omega) = \omega(S(\omega))$. Since $\mathbb{P}_{s,y} = Q_{s,y}$, it then follows that

$$\mathbb{P}_{S(\omega),X(S)(\omega)}[A] = \mathbb{E}_{\tau,x}\left[\mathbf{1}_A \circ \vee_S \,\Big|\, \mathcal{F}_{S+}^{\tau}\right](\omega), \quad A \in \mathcal{F}_T^{\tau}.$$

Or putting it differently:

$$\mathbb{P}_{S,X(S)}\left[\mathbf{1}_A \circ \vee_S\right] = \mathbb{E}_{\tau,x}\left[\mathbf{1}_A \circ \vee_S \,\Big|\, \mathcal{F}_{S+}^{\tau}\right], \quad A \in \mathcal{F}_T^{\tau}. \tag{3.147}$$

However this is exactly the strong Markov property.

This concludes the proof of Proposition 3.8. □

The following proposition can be proved in the same manner as Theorem 5.1 Corollary in [Ikeda and Watanabe (1998)], p. 206.

Proposition 3.9. *If an operator family $L = \{L(s) : 0 \leqslant s \leqslant T\}$ generates a Feller evolution $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$, then the martingale problem is uniquely solvable for L.*

Proof. Let $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ be a Feller evolution generated by L and let

$$\left\{\left(\Omega, \mathcal{F}_T^{\tau}, \mathbb{P}_{\tau,x}\right), \left(X(t), \tau \leqslant t \leqslant T\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), (E, \mathcal{E})\right\}, \tag{3.148}$$

be the associated strong Markov process (see Theorem 2.9) If f belongs to $D^{(1)}(L)$, then the process

$$M_f(t) := f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t \left(D_1 + L\right) f(s, X(s)) ds, \quad t \in [\tau, T],$$

is a $\mathbb{P}_{\tau,x}$-martingale for all $(\tau, x) \in [0, T] \times E$. This can be seen as follows. Fix $T \geqslant t_2 > t_1 \geqslant 0$. Then

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[M_f(t_2) \mid \mathcal{F}_{t_1}^{\tau}\right] - M_f(t_1) \\
&= \mathbb{E}_{\tau,x}\left[f\left(t_2, X(t_2)\right) - \int_{t_1}^{t_2} \left(D_1 + L\right) f(X(s)) ds \mid \mathcal{F}_{t_1}^{\tau}\right] - f(t_1, X(t_1))
\end{aligned}$$

(Markov property)

$$\begin{aligned}
&= \mathbb{E}_{t_1,X(t_1)}\left[f\left(t_2, X(t_2)\right) - \int_{t_1}^{t_2} \left(D_1 + L\right) f(s, X(s)) ds\right] - f(t_1, X(t_1)) \\
&= \mathbb{E}_{t_1,X(t_1)}\left[f\left(t_2, X(t_2)\right)\right] - \int_{t_1}^{t_2} \mathbb{E}_{t_1,X(t_1)}\left[\left(D_1 + L\right) f(s, X(s))\right] ds \\
&\qquad - f(t_1, X(t_1))
\end{aligned}$$

(see Proposition 4.1 in Chapter 4)

$$\begin{aligned}
&= \mathbb{E}_{t_1,X(t_1)}\left[f\left(t_2, X(t_2)\right)\right] - \int_{t_1}^{t_2} \frac{d}{ds}\mathbb{E}_{t_1,X(t_1)}\left[f(s, X(s))\right] ds - f(t_1, X(t_1)) \\
&= 0. \qquad (3.149)
\end{aligned}$$

Hence from (3.149) it follows that the process $M_f(t)$, $t \geqslant 0$, is a $\mathbb{P}_{\tau,x}$-martingale. Next we shall prove the uniqueness of the solutions of the martingale problem associated to the operator L. Let $\mathbb{P}_{\tau,x}^1$ and $\mathbb{P}_{\tau,x}^2$ be solutions "starting" in $x \in E$ at time τ. We have to show that these probabilities coincide. Let f belong to $D^{(1)}(L)$ and let $S : \Omega \to [\tau, T]$ be an $\left(\mathcal{F}_{t+}^{\tau}\right)_{t\in[\tau,T]}$-stopping time. Then, via partial integration, we infer

$$\begin{aligned}
&\lambda \int_0^\infty e^{-\lambda t} \left\{ f\left((t+S) \wedge T, X\left((t+S) \wedge T\right)\right) \right. \\
&\qquad \left. - \int_S^{t+S} \left(D_1 + L\right) f\left(\rho \wedge T, X\left(\rho \wedge T\right)\right) d\rho - f\left(S, X(S)\right) \right\} dt + f\left(S, X(S)\right) \\
&= \lambda \int_0^\infty e^{-\lambda t} \left\{ f\left((t+S) \wedge T, X\left((t+S) \wedge T\right)\right) \right. \\
&\qquad \left. - \int_S^{t+S} \left(D_1 + L\right) f\left(\rho \wedge T, X\left(\rho \wedge T\right)\right) d\rho \right\} dt \\
&= \lambda \int_0^\infty e^{-\lambda t} f\left((t+S) \wedge S, X\left((t+S) \wedge T\right)\right) dt \\
&\qquad - \lambda \int_0^\infty e^{-\lambda t} \int_0^t \left(D_1 + L\right) f\left((t+S) \wedge T, X\left((\rho+S) \wedge T\right)\right) d\rho\, dt
\end{aligned}$$

$$
\begin{aligned}
&= \lambda \int_0^\infty e^{-\lambda t} f\left((t+S)\wedge T, X\left((t+S)\wedge T\right)\right)\, dt \\
&\quad - \lambda \int_0^\infty \left(\int_\rho^\infty e^{-\lambda t} dt\right) \left(D_1+L\right) f\left((\rho+S)\wedge T, X\left((\rho+S)\wedge T\right)\right)\, d\rho \\
&= \int_0^\infty e^{-\lambda t} \left[(\lambda I - D_1 - L) f\right]\left((t+S)\wedge T, X\left((t+S)\wedge T\right)\right)\, dt.
\end{aligned} \tag{3.150}
$$

From Doob's optional sampling theorem together with (3.150) we obtain:

$$
\begin{aligned}
&\int_0^\infty e^{-\lambda t}\mathbb{E}^1_{\tau,x}\left[(\lambda I - D_1 - L) f\left((t+S)\wedge T, X\left((t+T)\wedge T\right)\right) \,\Big|\, \mathcal{F}^\tau_{S+}\right] dt \\
&= \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}^1_{\tau,x}\Bigg[\Bigg\{ f\left((t+S)\wedge T, X\left((t+S)\wedge T\right)\right) \\
&\quad - \int_S^{t+S} \left(D_1+L\right) f\left(\rho\wedge T, X\left(\rho\wedge T\right)\right) d\rho - f\left(S, X(S)\right)\Bigg\} \,\Bigg|\, \mathcal{F}^\tau_{S+}\Bigg] dt \\
&\quad + f\left(S, X(S)\right) \\
&= f\left(S, X(S)\right), \quad \mathbb{P}^1_{\tau,x}\text{-almost surely.}
\end{aligned} \tag{3.151}
$$

By the same token we also have $\mathbb{P}^2_{\tau,x}$-almost surely

$$
\begin{aligned}
&\lambda \int_0^\infty e^{-\lambda t}\mathbb{E}^2_{\tau,x}\Bigg[\Bigg\{ f\left((t+S)\wedge T, X\left((t+S)\wedge T\right)\right) \\
&\quad - \int_S^{t+S} \left(D_1+L\right) f\left(\rho\wedge T, X\left(\rho\wedge T\right)\right) d\rho - f\left(S, X(S)\right)\Bigg\} \,\Bigg|\, \mathcal{F}^\tau_{S+}\Bigg] dt \\
&\quad + f\left(S, X(S)\right) \\
&= \int_0^\infty e^{-\lambda t}\mathbb{E}^2_{\tau,x}\left[(\lambda I - D_1 - L) f\left((t+S)\wedge T, X\left((t+S)\wedge T\right)\right) \,\Big|\, \mathcal{F}^\tau_{S+}\right] dt.
\end{aligned} \tag{3.152}
$$

As in (3.84), (3.129) and (4.5) in Chapter 4 we write:

$S(\rho)f(t,x) = P\left(t, (\rho+t)\wedge T\right) f\left((\rho+t)\wedge T, \cdot\right)(x), \quad f \in C_b\left([0,T]\times E\right),$
$\rho \geqslant 0$, $(t,x) \in [0,T]\times E$. Then the family $\{S(\rho) : \rho \geqslant 0\}$ is a $\mathcal{T}_\beta$-continuous semigroup. Its resolvent is given by

$$
\begin{aligned}
\left[R(\lambda)f\right](\tau,x) &= \int_0^\infty e^{-\lambda t}\left[P\left(\tau, (\tau+t)\wedge T\right) f\left((\tau+t)\wedge T, \cdot\right)\right](x) dt \\
&= \int_0^\infty e^{-\lambda t} S(t) f(\tau,x) dt,
\end{aligned} \tag{3.153}
$$

for $x \in E$, $\lambda > 0$, and $f \in C_b([0,T] \times E)$. Let $L^{(1)}$ be its generator. Then, as will be shown in Theorem 4.1 below, $L^{(1)}$ is the $\mathcal{T}_\beta$-closure of $D_1 + L$, and

$$\begin{aligned} \left(\lambda I - L^{(1)}\right) R(\lambda) f &= f, \quad f \in C_b([0,T] \times E), \\ R(\lambda)\left(\lambda I - L^{(1)}\right) f &= f, \quad f \in D\left(L^{(1)}\right). \end{aligned} \tag{3.154}$$

Since $L^{(1)}$ is the $\mathcal{T}_\beta$-closure of $D_1 + L$, the equalities in (3.151) and (3.152) also hold for $L^{(1)}$ instead of $D_1 + L$. Among other things we see that

$$R\left(\lambda I - L^{(1)}\right) = C_b([0,T] \times E), \quad \lambda > 0.$$

From (3.151) and (3.152), with $L^{(1)}$ instead of $D_1 + L$, (3.153), and (3.154) it then follows that for $g \in C_b([0,T] \times E)$ we have

$$\begin{aligned} &\int_0^\infty e^{-\lambda t} \mathbb{E}^1_{\tau,x}\left[g((t+S) \wedge T, X((t+T) \wedge T)) \,\middle|\, \mathcal{F}^\tau_{S+}\right] dt \\ &= \int_0^\infty e^{-\lambda t} [S(t)g](S, X(S))\, dt \\ &= \int_0^\infty e^{-\lambda t} \mathbb{E}^2_{\tau,x}\left[g((t+S) \wedge T, X((t+T) \wedge T)) \,\middle|\, \mathcal{F}^\tau_{S+}\right] dt. \end{aligned} \tag{3.155}$$

Here the first equality in (3.155) holds $\mathbb{P}^1_{\tau,x}$-almost surely, and the second one holds $\mathbb{P}^2_{\tau,x}$-almost surely. Since Laplace transforms are unique, g belongs to $C_b([0,T] \times E)$, and paths are right continuous, we conclude

$$\begin{aligned} &\mathbb{E}^1_{\tau,x}\left[g((t+S) \wedge T, X((t+S) \wedge T)) \,\middle|\, \mathcal{F}^\tau_{S+}\right] \\ &= [S(t)g](S, X(S)) \\ &= \mathbb{E}^2_{\tau,x}\left[g((t+S) \wedge T, X((t+S) \wedge T)) \,\middle|\, \mathcal{F}^\tau_{S+}\right], \end{aligned} \tag{3.156}$$

whenever g belongs to $C_b([0,T] \times E)$, $t \in [0,\infty]$ and S is an $\left(\mathcal{F}^\tau_{t+}\right)_{t \in [\tau,T]}$-stopping time. The first equality in (3.156) holds $\mathbb{P}^1_{\tau,x}$-almost surely and the second $\mathbb{P}^2_{\tau,x}$-almost surely. In (3.156) we take for S a fixed time $s \in [\tau, T-t]$ and we substitute $\rho = t + s$. Then we get

$$\mathbb{E}^1_{\tau,x}\left[g((\rho, X(\rho))) \,\middle|\, \mathcal{F}^\tau_{s+}\right] = [S(\rho - s)g](s, X(s)) = \mathbb{E}^2_{\tau,x}\left[g(\rho, X(\rho)) \,\middle|\, \mathcal{F}^\tau_{s+}\right]. \tag{3.157}$$

For $s = \tau$ the equalities in (3.157) imply

$$\mathbb{E}^1_{\tau,x}\left[g((\rho, X(\rho))) \,\middle|\, \mathcal{F}^\tau_{\tau+}\right] = [S(\rho - \tau)g](\tau, X(\tau)) = \mathbb{E}^2_{\tau,x}\left[g(\rho, X(\rho)) \,\middle|\, \mathcal{F}^\tau_{\tau+}\right], \tag{3.158}$$

and by taking expectations in (3.158) we get

$$\mathbb{E}^1_{\tau,x}\left[g((\rho, X(\rho)))\right] = [S(\rho - \tau)g](\tau, x) = \mathbb{E}^2_{\tau,x}\left[g(\rho, X(\rho))\right] \tag{3.159}$$

where we used the fact that $X(\tau) = x$ $\mathbb{P}^1_{\tau,x}$- and $\mathbb{P}^2_{\tau,x}$-almost surely. It follows that the one-dimensional distributions of $\mathbb{P}^1_{\tau,x}$ and $\mathbb{P}^2_{\tau,x}$ coincide. By induction with respect to n and using (3.157) several times we obtain:

$$\mathbb{E}^1_{\tau,x}\left[\prod_{j=1}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right] = \mathbb{E}^2_{\tau,x}\left[\prod_{j=1}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right] \tag{3.160}$$

for $n = 1, 2, \ldots$ and for $f_1, \ldots, f_n$ in $C_b\left([0,T] \times E\right)$. But then the probabilities $\mathbb{P}^1_{\tau,x}$ and $\mathbb{P}^2_{\tau,x}$ are the same.

This proves Proposition 3.9. □

The following proposition establishes a close link between unique solutions to the martingale problem and generators of strong Markov processes.

Proposition 3.10. *Let L be a densely defined operator for which the martingale problem is uniquely solvable. Then there exists a unique closed linear extension L_0 of L, which is the generator of a Feller semigroup.*

Proof. *Existence.* Let $\{\mathbb{P}_{\tau,x} : (\tau, x) \in [0,T] \times E\}$ be the solution for L. Put

$$\begin{aligned} \left[S(t)f\right](\tau,x) &= \mathbb{E}_{\tau,x}\left[f\left((\tau+t) \wedge T, X\left((\tau+t) \wedge T\right)\right)\right], \\ \left[R(\lambda)f\right](\tau,x) &= \int_0^\infty e^{-\lambda s}\left[S(s)f\right](\tau,x)ds, \\ L_0\left(R(\lambda)f\right) &:= \lambda R(\lambda)f - f, \quad f \in C_b\left([0,T] \times E\right). \end{aligned}$$

Here $t \in [0,T]$ and $\lambda > 0$ are fixed. Then, as follows from the proof of Theorem 4.1 the operator L_0 extends $D_1 + L$ and generates a $\mathcal{T}_\beta$-continuous Feller semigroup.

Uniqueness. Let L_1 and L_2 be closed linear extensions of L, which both generate Feller evolutions. Let

$$\left\{\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}^1_{\tau,x}\right), \left(X(t) : t \in [0,T]\right), \left(\vee_t : t \in [0,T]\right), (E, \mathcal{E})\right\}$$

respectively

$$\left\{\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}^2_{\tau,x}\right), \left(X(t) : t \in [0,T]\right), \left(\vee_t : t \in [0,T]\right), (E, \mathcal{E})\right\}$$

be the corresponding Markov processes. For every $f \in D(L)$, the process

$$f(t, X(t)) - f(\tau, X(\tau)) - \int_\tau^t \left(D_1 + L\right) f(s, X(s))ds, \quad t \geqslant 0,$$

is a martingale with respect to $\mathbb{P}^1_{\tau,x}$ as well as with respect to $\mathbb{P}^2_{\tau,x}$. Uniqueness implies $\mathbb{P}^1_{\tau,x} = \mathbb{P}^2_{\tau,x}$ and hence $L_1 = L_2$.

So the proof of Proposition 3.10 is now complete. □

Proof. [Proof of Theorem 2.12: conclusion.] In this final part of the proof we mainly collect the results, which we proved in Theorem 2.9, and Propositions 3.6, 3.7, 3.2, 3.8, 3.9, and 3.10. The main work we have to do is to organize these matters into a proof of Theorem 2.12. More details follow. As in (3.126) let $P''(\Omega) = \{\mathbb{P}_{\tau,x} : (\tau, x) \in [0, T] \times E\}$, be the collection of unique solutions to the martingale problem. Then the process

$$\left\{(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})_{(\tau,x)\in[0,T]\times E}, (X(t), t \in [0,T]), (\vee_t : t \in [0,T]), (E, \mathcal{E})\right\}$$

is strong Markov process, and the function $P(\tau, x; t, B)$ defined by

$$P(\tau, x; t, B) = \mathbb{P}_{\tau,x}[X(t) \in B], \quad 0 \leqslant \tau \leqslant t \leqslant T, \ x \in E, \ B \in \mathcal{E},$$

is a Feller evolution. Here the state variables $X(t) : \Omega \to E^\triangle$ are defined by $X(t) = \omega(t)$, $\omega \in \Omega = D\left([0, T], E^\triangle\right)$. The sample path space is supplied with the standard filtration $(\mathcal{F}_t^\tau)_{\tau\leqslant t\leqslant T}$. The strong Markov property follows from Proposition 3.8. The Feller property is a consequence of Proposition 3.7 (which in turn is based on Proposition 3.6 where completeness and separability of the space $P''(\Omega)$ is heavily used). Its $\mathcal{T}_\beta$-continuity and $\mathcal{T}_\beta$-equi-continuity is explained in Corollary 3.2 to Proposition 3.8. Define the Feller semigroup $\{S(\rho) : \rho \geqslant 0\}$ on $C_b([0, T] \times E)$ as in (3.129), and let $L^{(1)}$ be its generator. From Corollary 3.2 we see that $L^{(1)}$ extends the operator $D_1 + L$. Since the martingale problem is uniquely solvable for the operator L, it follows that the martingale problem is uniquely solvable for the operator $L^{(1)}$ (but now as a time-homogeneous martingale problem). Therefore, Proposition 3.9 implies that the operator $L^{(1)}$ is the unique extension of $D_1 + L$ which generates a Feller semigroup. It follows that $L^{(1)} - D_1$ is the unique $\mathcal{T}_\beta$-extension of L which generates a Feller evolution. This Feller evolution is given by the original solution to the martingale problem: this claim follows from Theorem 2.9.

Finally, this completes the proof of Theorem 2.12. □

3.1.5 *Proof of Theorem 2.13*

In this subsection we will show that under certain conditions, like possessing the Korovkin property, satisfying the maximum principle, and $\mathcal{T}_\beta$-equi-continuity a $\mathcal{T}_\beta$-densely defined operator in $C_b(E)$ has a unique extension which generates a (strong) Markov process.

Proof. [Proof of Theorem 2.13.] Let E_0 be a subset of $[0, T] \times E$ which is Polish for the relative topology. First suppose that the operator $D_1 + L$ possesses the Korovkin property on E_0. Also suppose that it satisfies the maximum principle on E_0. By Proposition 4.4 and its Corollary 4.2 there exists a family of linear operators $\{R(\lambda) : \lambda > 0\}$ such that for all $(\tau_0, x_0) \in E_0$ and $g \in C_b(E_0)$ the following equalities hold:

$$\begin{aligned}
&\lambda R(\lambda) g(\tau_0, x_0) \\
&= \inf_{h \in D^{(1)}(L)} \max_{(\tau,x)\in E_0} \left\{ h(\tau_0, x_0) + \left[g - \left(I - \frac{1}{\lambda}(D_1 + L) \right) h \right](\tau, x) \right\} \\
&= \inf_{h \in D^{(1)}(L)} \left\{ h(\tau_0, x_0) : \left(I - \frac{1}{\lambda}(D_1 + L) \right) h \geqslant g \quad \text{on } E_0 \right\} \\
&= \sup_{h \in D^{(1)}(L)} \left\{ h(\tau_0, x_0) : \left(I - \frac{1}{\lambda}(D_1 + L) \right) h \leqslant g \quad \text{on } E_0 \right\} \\
&= \sup_{h \in D^{(1)}(L)} \min_{(\tau,x)\in E_0} \left\{ h(\tau_0, x_0) + \left[g - \left(I - \frac{1}{\lambda}(D_1 + L) \right) h \right](\tau, x) \right\}.
\end{aligned} \tag{3.161}$$

As will be shown in Proposition 4.4 the family $\{R(\lambda) : \lambda > 0\}$ has the resolvent property: $R(\lambda) - R(\mu) = (\lambda - \mu) R(\mu) R(\lambda)$, $\lambda > 0$, $\mu > 0$. It also follows that $R(\lambda)(\lambda I - D_1 - L) f = f$ on E_0 for $f \in D^{(1)}(L)$. This equality is an easy consequence of the inequalities in (3.161): see Corollary 4.2. Fix $\lambda > 0$ and $f \in C_b([0,T] \times E)$. We will prove that $f = \mathcal{T}_\beta\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha) f$. If f is of the form $f = R(\lambda) g$, $g \in C_b(E_0)$, then by the resolvent property we have

$$\alpha R(\alpha) f - f = \alpha R(\alpha) R(\lambda) g - R(\lambda) g = \frac{\alpha}{\alpha - \lambda} R(\lambda) g - R(\lambda) g - \frac{\alpha R(\alpha) g}{\alpha - \lambda}. \tag{3.162}$$

Since $\|\alpha R(\alpha) g\|_\infty \leqslant \|g\|_\infty$, the equality in (3.162) yields

$$\|\cdot\|_\infty\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha) f - f = 0 \quad \text{for } f \text{ of the form } f = R(\lambda) g,\ g \in C_b(E_0).$$

Since $g = R(\lambda)(\lambda I - D_1 - L) g$ on E_0, $g \in D^{(1)}(L)$, it follows that

$$\lim_{\alpha\to\infty} \|\alpha R(\alpha) g - g\|_\infty = 0 \quad \text{for } g \in D^{(1)}(L) = D(D_1) \bigcap D(L). \tag{3.163}$$

As will be proved in Corollary 4.3 there exists $\lambda_0 > 0$ such that the family $\{\lambda R(\lambda) : \lambda \geqslant \lambda_0\}$ is $\mathcal{T}_\beta$-equi-continuous. Hence for $u \in H^+(E_0)$ there exists $v \in H^+(E_0)$ that for $\alpha \geqslant \lambda_0$ we have

$$\|u \alpha R(\alpha) g\|_\infty \leqslant \|v g\|_\infty, \quad g \in C_b(E_0). \tag{3.164}$$

Fix $\varepsilon > 0$, and choose for $f \in C_b(E_0)$ and $u \in H^+(E_0)$ given $g \in D^{(1)}(L)$ in such a way that

$$\|u(f - g)\|_\infty + \|v(f - g)\|_\infty \leqslant \frac{2}{3}\varepsilon. \tag{3.165}$$

Since $D(L)$ is $\mathcal{T}_\beta$-dense in $C_b([0,T]\times E)$ such a choice of g is possible by. The inequality (3.165) and the identity

$$\begin{aligned}
&\alpha R(\alpha)f - f = \alpha R(\alpha)(f-g) - (f-g) + \alpha R(\alpha)g - g \quad \text{yield} \\
&\|u(\alpha R(\alpha)f - f)\|_\infty \\
&\leqslant \|u(\alpha R(\alpha)(f-g))\|_\infty + \|u(f-g)\|_\infty + \|u\alpha R(\alpha)g - g\|_\infty \\
&\leqslant \|v(f-g)\|_\infty + \|u(f-g)\|_\infty + \|u\alpha R(\alpha)g - g\|_\infty \\
&\leqslant \frac{2}{3}\varepsilon + \|u(\alpha R(\alpha)g - g)\|_\infty .
\end{aligned} \tag{3.166}$$

From (3.163) and (3.166) we infer $\mathcal{T}_\beta$-$\lim_{\alpha\to\infty}\alpha R(\alpha)f = f$, $f \in C_b(E_0)$. Of course the same arguments apply if $E_0 = [0,T]\times E$. The detailed arguments which prove the fact that the operator $D_1 + L$, confined to E_0, extends to the unique generator of a Feller semigroup are found in the proof of Theorem 4.4.

Let $E_0 = [0,T] \times E_0'$ where E_0' is a Polish subspace of E. Let $\mathcal{E}_0$ and $\mathcal{E}_0'$ be the Borel field of E_0 respectively E_0'. We still have to show that the martingale problem for the operator L restricted to E_0 is well posed. Saying that the martingale is well posed for $L\restriction_{E_0}$ is the same as saying that the martingale problem is well posed for the operator $(D_1 + L)\restriction_{E_0}$. More precisely, if

$$\left\{(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})_{(\tau,x)\in E_0}, (X(t), t \in [0,T]), (E_0', \mathcal{E}_0')\right\} \tag{3.167}$$

is a solution to the martingale problem associated to $L\restriction_{E_0}$, then the time-homogeneous family

$$\left\{\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbb{P}^{(0)}_{\tau,x}\right)_{(\tau,x)\in E_0}, (Y(t), t \geqslant 0), (E_0, \mathcal{E}_0)\right\} \tag{3.168}$$

is a solution to the martingale problem associated with $(D_1 + L)\restriction_{E_0}$. Here $\widetilde{\Omega} = [0,T]\times\Omega$, $Y(t)(\tau,\omega) = ((\tau+t)\wedge T, X((\tau+t)\wedge T))$, $(\tau,\omega) \in [0,T]\times\Omega = \widetilde{\Omega}$, and the measure $\mathbb{P}^{(0)}_{\tau,x}$ is determined by the equality

$$\mathbb{E}^{(0)}_{\tau,x}\left[\prod_{j=1}^n f_j(Y(t_j))\right] = \mathbb{E}_{\tau,x}\left[\prod_{j=1}^n f_j((\tau+t_j)\wedge T, X((\tau+t_j)\wedge T))\right] \tag{3.169}$$

where the functions f_j, $1 \leqslant j \leqslant n$, are bounded Borel measurable functions on E_0, and where $0 \leqslant t_1 < \dots < t_n$. Conversely, if the measures $\mathbb{P}^{(0)}_{\tau,x}$ in (3.168) are known, then those in (3.167) are also determined by (3.169):

$$\mathbb{E}_{\tau,x}\left[\prod_{j=1}^n f_j(t_j, X(t_j))\right] = \mathbb{E}^{(0)}_{\tau,x}\left[\prod_{j=1}^n f_j(Y(t_j-\tau))\right] \tag{3.170}$$

where the functions f_j, $1 \leqslant j \leqslant n$, are again bounded Borel measurable functions on E_0, and where $\tau \leqslant t_1 < \cdots < t_n \leqslant T$. In fact in (3.170) the functions f_j, $1 \leqslant j \leqslant n$, only need to be defined on E_0. It follows that instead of considering the time-inhomogeneous martingale problem associated with $L\restriction_{E_0}$ we may consider the time-homogeneous martingale problem associated with $(D_1 + L)\restriction_{E_0}$. However, the martingale problem for the time-homogeneous case is taken care of in the final part of Theorem 4.4.

So combining the above observations with Theorem 4.4 completes the proof of Theorem 2.13. □

Remark 3.1. It is left as an exercise for the reader to prove that the process in (3.5) is a supermartingale indeed.

Remark 3.2. Let $(\psi_m)_{m\in\mathbb{N}}$ be a sequence in $C_b^+([\tau,T]\times E)$ which decreases pointwise to the zero function. Since the orbit

$$\left\{\left(t,\widetilde{X}(t)\right) : t \in [\tau,T]\right\}$$

is $\mathbb{P}_{\tau,x}$-almost surely compact, or, equivalently, totally bounded, we know that

$$\inf_{m\in\mathbb{N}} \sup_{t\in[\tau,T]} \psi_m\left(t,\widetilde{X}(t)\right) = 0, \quad \mathbb{P}_{\tau,x}\text{-almost surely.}$$

3.1.6 *Some historical remarks*

The Lévy numbers in (3.106) are closely related to the Lévy metric, which in turn is related to approach structures. The definition of Lévy metric and Lévy-Prohorov metric can be found in Encyclopaedia of Mathematics, edited by Hazewinkel [Hazewinkel (2001)]. Lévy numbers could also have called tightness numbers. In the area of convergence of measures the Encyclopaedia contains contributions by V. M. Zolotarev. In fact special sections are devoted to the Lévy metric, the Lévy-Prokhorov metric, and related topics like convergence of probability measures on complete metrizable spaces. The Lévy metric goes back to Lévy: see [Lévy (1937)]. The Lévy-Prohorov metric generalizes the Lévy metric, and has its origin in Prohorov [Prohorov (1956)]. Whereas in [van Casteren (1992)] we used only the first term in the distance d_L of formula (3.107) this is not adequate in the non-compact case. The reason for this is that the second term in the right-hand side of the definition of the metric $d_L(\mathbb{P}_2,\mathbb{P}_1)$ in (3.107) ensures us that the limiting "functionals" are probability measures indeed.

Here we use a concept which, for distribution functions, is due to Lévy. For probability measures on a metric space the corresponding metric originates from Prohorov. This metric is often called the Lévy-Prohorov metric. For completeness we insert the definition of the latter metric.

Definition 3.2. Let (E, d) be a metric space with its Borel sigma field $\mathcal{E}$. Let $\mathcal{P}(E)$ denote the collection of all probability measures on the measurable space $(E, \mathcal{E})$. For a subset $A \subseteq E$, define the ε-neighborhood of A by

$$A^{\varepsilon} := \{x \in E : \text{there exists } y \in A \text{ such that } d(x,y) < \varepsilon\} = \bigcup_{y \in A} B(y,\varepsilon)$$

where $B(y,\varepsilon)$ is the open ball of radius ε centered at y. The Lévy-Prohorov metric $d_{\mathrm{LP}} : \mathcal{P}(E)^2 \to [0,+\infty)$ is defined by setting the distance between two probability measures μ and ν as

$$\begin{aligned} &d_{\mathrm{LP}}(\mu,\nu) && (3.171)\\ &= \inf\left\{\varepsilon > 0 : \mu(A) \leqslant \nu\left(A^{\varepsilon}\right) + \varepsilon \,\text{and}\, \nu(A) \leqslant \mu\left(A^{\varepsilon}\right) + \varepsilon \,\text{for all}\, A \in \mathcal{E}\right\}. \end{aligned}$$

For probability measures μ and ν we clearly have

$$d_{\mathrm{LP}}(\mu,\nu) = \inf\left\{\varepsilon > 0 : \mu(A) \leqslant \nu\left(A^{\varepsilon}\right) + \varepsilon, \text{ for all } A \in \mathcal{E}\right\} \leqslant 1. \quad (3.172)$$

Some authors omit one of the two inequalities or choose only open or closed subsets A; either inequality implies the other, but restricting to open or closed sets changes the metric as defined in (3.171). The Lévy-Prohorov metric is also called the Prohorov metric. The interested reader should compare the definition of Lévy-Prohorov metric with that of approach structure as exhibited in e.g. [Lowen (1997)]. When discussing convergence of measures and constructing appropriate metrics the reader is also referred to [Billingsley (1999)], [Parthasarathy (2005)], [Rachev (1991)], [Zolotarev (1983)], and others like Bickel, Klaassen, Ritov and Wellner in [Bickel *et al.* (1993)], appendices A6–A9. A book which uses the notion of Korovkin set to a great extent is [Altomare and Campiti (1994)]. For applications of Korovkin sets to ergodic theory see e.g. [Marsden and Riemenschneider (1974)], [Nishishiraho (1998)], [Labsker (1982)], Chapter 7 and 8 in [Donner (1982)], and [Krengel (1985)]. Another book of interest is [Bergelson *et al.* (1996)] edited by Bergelson, March and Rosenblatt. For the convergence results we also refer to the original book by Korovkin [Korovkin (1960)]. The reader also might want to consult (the references in) Bukhalov [Bukhvalov (1988)]. In the terminology of test sets, or Korovkin sets, our space $D^{(1)} = D\left(D_1\right) \bigcap D\left(L\right)$ in $C_b\left([0,T] \times E\right)$ is a Korovkin set for the resolvent family $(\lambda I - D_1 - L)^{-1}$, $\lambda > 0$. From the proof of Theorem 2.13 it follows that we only need the Korovkin property for some fixed $\lambda_0 > 0$: see the definitions 4.4 and 2.12. In the finite-dimensional setting these Korovkin sets may be relatively small: see e.g. Özarslan and Duman [Özarslan and Duman (2007)]. Section 5.2 in the recent book on functional analysis by Dzung Minh Ha [Ha (2007)] carries the title "Korovkin's theorem and the Weierstrass approximation theorem".

3.1.7 *Kolmogorov extension theorem*

In this subsection we present the Kolmogorov extension (or existence) theorem for Polish spaces. It reads as follows. Let $\{E_t : t \in T\}$ be a family of Polish spaces equipped with their Borel σ-field $\mathcal{E}_t$. We identify each $\mathcal{E}_F$ with the collection $\widehat{\mathcal{E}}_F$ of F-cylinder sets in $E_T = \prod_{t\in T} E_t$. That is, $\widetilde{\mathcal{E}}_F$ consists of all sets of the form $A \times \prod_{t\in T\setminus F} E_t$, where A belongs to $\mathcal{E}_F$. By definition the product σ-field $\otimes_{t\in T}\mathcal{E}_t$ is the σ-field generated by $\left\{\widehat{\mathcal{E}}_F : F \text{ is a finite subset of } T\right\}$. Define $\widehat{P}_F$ on $\widehat{\mathcal{E}}_F$ by

$$\widehat{P}_F\left(A \times \prod_{t\in T\setminus F} E_t\right) = P_F(A), \quad A \in \mathcal{E}_F.$$

Regard the family of finite subsets of T as a net directed upward by inclusion. The family $\left\{\left(E^F, \mathcal{E}_F, P_F\right) : F \subset T \text{ finite}\right\}$ is called (Kolmogorov) consistent if for every $t_0 \in T$ and for finite every finite subset $F \subset T$ the equality $P_{F\bigcup\{t_0\}}\left[A \times E_{t_0}\right] = P_F\left[A\right]$ for all Borel subsets $A \in \mathcal{E}_F$. Moreover, it is implicitly assumed that $P_{t_{\sigma(1)},\dots,t_{\sigma(n)}}\left[A_{\sigma(1)},\dots,A_{\sigma(n)}\right] = P_{t_1,\dots,t_n}\left[A_1,\dots,A_n\right]$, $A_j \in \mathcal{E}_{t_j}$, $1 \leqslant j \leqslant n$, whenever $\{t_1,\dots,t_n\}$ is a subset of n elements of T, and whenever σ is a permutation on n elements. The consistency property is equivalent to saying that the probability measures $\widehat{P}_F$ and $\widehat{P}_{F'}$ coincide on $\widehat{\mathcal{E}}_F$ whenever F is a subset of the finite subset F' of T. Consistent families of probability spaces are also called projective systems of probability measures or cylindrical measures.

Theorem 3.1. *Let $\{E_t : t \in T\}$ be a family of Polish spaces equipped with their Borel σ-field $\mathcal{E}_t$. For each finite subset F of T let P_F be a Borel probability measure on $E_F = \prod_{t\in F} E_t$ with its product (Borel) σ-field $\mathcal{E}_F$. Assume the distributions $\left\{\widehat{P}_F : F \subset T,\ F \text{ finite}\right\}$ are Kolmogorov consistent. Then there is a unique probability measure on the infinite product σ-field $\mathcal{E}_T = \otimes_{t\in T}\mathcal{E}_t$ that extends each measure $\widehat{P}_F$.*

A proof of this result can be found in a course text by Kim C. Border [Border (1998)]. An important tool which is a basic ingredient of the proof is the inner regularity of the measures $P_t = P_{\{t\}}$, $t \in T$. Since for every $t \in T$ the space E_t is Polish the measure P_t is inner regular in the sense that $P_t[A] = \sup_{K \subset A,\, K \text{ compact}} P_t[K]$ for all $A \in \mathcal{E}_t$. For this result the reader may consult [Aliprantis and Border (1994)], Theorem 11.20. In fact Theorem 3.1 is also true if the spaces E_t, $t \in T$, are merely topological Hausdorff spaces and the measures P_t are inner regular. In addition, the text by Border contains an example of a situation where Kolmogorov's extension theorem does not hold. The example is due to Andersen and Jessen [Andersen and Jessen (1948)]; for related topics see Dudley [Dudley (2002)].

This page is intentionally left blank

Chapter 4

Space-time operators and miscellaneous topics

In this chapter we discuss a number of issues related to time dependent Markov processes. Topics include space-time operators, dissipative operators, continuity of sample paths, measurability properties of hitting times. Another feature of the present chapter is the fact that to a Feller propagator on $C_b(E)$ we can associate a Feller semigroup in the space $C_b([0,T] \times E)$: see formula (4.5).

4.1 Space-time operators

In this section we will discuss in more detail the generators of the time-space Markov process (see (2.76):

$$\left\{ \left(\Omega, \mathcal{F}_T^{\tau}, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant \tau\right), \left(\vee_t : \tau \leqslant t \leqslant T\right), \left(E, \mathcal{E}\right) \right\}. \tag{4.1}$$

In Definition 2.7 we have introduced the family of generators of the corresponding Feller evolution $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ given by $P(\tau,t) f(x) = \mathbb{E}_{\tau,x}\left[f(X(t))\right]$, $f \in C_b(E)$. In fact for any fixed $t \in [0,T]$ we will consider the Feller evolution as an operator from $C_b([0,T] \times E)$ to $C_b([0,t] \times E)$. This is done in the following manner. To a function $f \in C_b([0,T] \times E)$ our Feller evolution assigns the function $(\tau, x) \mapsto P(\tau,t) f(t,\cdot)(x)$. We will also consider the family of operators $L := \{L(t) : t \in [0,T)\}$ as defined in Definition 2.7, and which is considered as a linear operator which acts on a subspace of $C_b([0,T] \times E)$. It is called the (infinitesimal) *generator* of the Feller evolution $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$, if $L(s) f = \mathcal{T}_{\beta}\text{-}\lim_{t \downarrow s} \dfrac{P(s,t)f - f}{t-s}$, $0 \leqslant s \leqslant T$. This means that a function f belongs to $D(L(s))$ whenever $L(s) f := \lim_{t \downarrow s} \dfrac{P(s,t)f - f}{t-s}$ exists in $C_b(E)$, equipped with the strict topology. As explained earlier, such a family of operators is considered as an operator L with domain in the space $C_b([0,T]\times E)$. A function $f\in C_b([0,T]\times E)$ is said to belong to $D(L)$ if for every $s\in[0,T]$ the function $x\mapsto f(s,x)$ is a member of $D(L(s))$ and if the function $(s,x)\mapsto L(s)f(s,\cdot)(x)$ belongs to $C_b(E)$. Instead of $L(s)f(s,\cdot)(x)$ we often write $L(s)f(s,x)$. If a function $f\in D(L)$ is such that the function $s\mapsto f(s,x)$ is differentiable, then we say that f belongs to $D^{(1)}(L)$. We will show that such a generator also generates the corresponding Markov process in the sense of Definition 2.8. For convenience of the reader we repeat here the defining property. A family of operators $L:=\{L(s):\,0\leqslant s\leqslant T\}$, is said to generate a time-inhomogeneous Markov process, as described in (2.76), if for all functions $u\in D(L)$, for all $x\in E$, and for all pairs (τ,s) with $0\leqslant\tau\leqslant s\leqslant T$ the following equality holds:

$$\frac{d}{ds}\mathbb{E}_{\tau,x}\left[u\left(s,X(s)\right)\right]=\mathbb{E}_{\tau,x}\left[\frac{\partial u}{\partial s}(s,x)+L(s)u\left(s,\cdot\right)\left(X(s)\right)\right]. \tag{4.2}$$

Our first result says that generators of Markov processes and the corresponding Feller evolutions coincide.

Proposition 4.1. *Let the Markov process in (4.1) and the Feller evolution $\{P(\tau,t):0\leqslant\tau\leqslant t\leqslant T\}$ be related by $P(\tau,t)f(x)=\mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right]$, $f\in C_b(E)$. Let $L=\{L(s):0\leqslant s\leqslant T\}$ be a family of linear operators with domain and range in $C_b(E)$. If L is a generator of the Feller evolution, then it also generates the corresponding Markov process. Conversely, if L generates a Markov process, then it also generates the corresponding Feller evolution.*

Proof. First suppose that the Feller evolution $\{P(\tau,t):0\leqslant\tau\leqslant t\leqslant T\}$ is generated by the family L. Let the function f belong to the domain of L and suppose that D_1f is continuous on $[0,T]\times E$. Then we have

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\frac{\partial f}{\partial s}\left(s,X(s)\right)+L(s)f\left(s,\cdot\right)\left(X(s)\right)\right]\\
&=P(\tau,s)\frac{\partial f}{\partial s}(s,\cdot)(x)+P(\tau,s)L(s)f(s,\cdot)(x)\\
&=P(\tau,s)\frac{\partial f}{\partial s}(s,\cdot)(x)+P(\tau,s)\left[\lim_{h\downarrow 0}\frac{P(s,s+h)f(s,\cdot)-f(s,\cdot)}{h}\right](x)\\
&=P(\tau,s)\frac{\partial f}{\partial s}(s,\cdot)(x)+\lim_{h\downarrow 0}P(\tau,s)\left[\frac{P(s,s+h)f(s,\cdot)-f(s,\cdot)}{h}\right](x)\\
&=P(\tau,s)\frac{\partial f}{\partial s}(s,\cdot)(x)+\lim_{h\downarrow 0}\left[\frac{P(\tau,s+h)f(s,\cdot)-P(\tau,s)f(s,\cdot)}{h}\right](x)
\end{aligned}$$

$$= P(\tau, s) \frac{\partial f}{\partial s}(s, \cdot)(x) - \lim_{h \downarrow 0} P(\tau, s+h) \left[\frac{f(s+h, \cdot) - f(s, \cdot)}{h} \right](x)$$
$$+ \lim_{h \downarrow 0} \left[\frac{P(\tau, s+h) f(s+h, \cdot) - P(\tau, s) f(s, \cdot)}{h} \right](x).$$
$$= P(\tau, s) \frac{\partial f}{\partial s}(s, \cdot)(x) - P(\tau, s) \frac{\partial f}{\partial s}(s, \cdot)(x)$$
$$+ \lim_{h \downarrow 0} \frac{\mathbb{E}_{\tau,x}[f(s+h, X(s+h))] - \mathbb{E}_{\tau,x}[f(s, X(s))]}{h}$$
$$= \frac{d}{ds} \mathbb{E}_{s,X(s)}[f(s, X(s))]. \tag{4.3}$$

In (4.3) we used the fact that the function $D_1 f$ is continuous and its consequence that $\lim_{h \downarrow 0} \frac{f(s+h, y) - f(s, y)}{h}$ converges uniformly for y in compact subsets of E. We also used the fact that the family of operators $\{P(\tau, t) : t \in [\tau, T]\}$ is equi-continuous for the strict topology.

In the second part we have to show that a generator L of a Feller process (4.1) also generates the corresponding Feller evolution. Therefore we fix $s \in [0, T]$ and take $f \in D(L(s)) \subset C_b(E)$. Using the fact that L generates the Markov process in (4.1) we infer for $h \in (0, T-s)$:

$$\lim_{h \downarrow 0} \frac{P(s, s+h) f(x) - f(x)}{h} = \frac{d}{dh} P(s, s+h) \big|_{h=0}$$
$$= \frac{d}{dh} \mathbb{E}_{s,x}[f(X(s+h))] \big|_{h=0} = \mathbb{E}_{s,x}[L(s) f(X(s))] = L(s) f(x). \tag{4.4}$$

This completes the proof of Proposition 4.1 □

To such a Feller evolution $\{P(\tau, t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ we may also associate a semigroup of operators $S(\rho)$ acting on the space $C_b([0, T] \times E)$ and the corresponding resolvent family $\{R(\alpha) : \Re\alpha > 0\}$. The semigroup $\{S(\rho) : \rho \geqslant 0\}$ is defined by the formula:

$$S(\rho) f(t, x) = P(t, (\rho + t) \wedge T) f((\rho + t) \wedge T, \cdot)(x)$$
$$= \mathbb{E}_{t,x}[f((\rho + t) \wedge T, X((\rho + t) \wedge T))], \tag{4.5}$$

$f \in C_b([0, T] \times E)$, $(t, x) \in [0, T] \times E$. Notice that the operator $S(\rho)$ does not leave the space $C_b(E)$ invariant: i.e. a function of the form $(s, y) \mapsto f(y)$, $f \in C_b(E)$, will be mapped to function $S(\rho) f \in C_b([0, T] \times E)$ which really depends on the time variable. Then the resolvent operator $R(\alpha)$ which also acts as an operator on the space of bounded continuous functions on space-time space $C_b([0, T] \times E)$ is given by

$$R(\alpha) f(t, x) = \int_t^\infty e^{-\alpha(\rho - t)} P(t, \rho \wedge T) f(\rho \wedge T, \cdot)(x) d\rho$$

$$
\begin{aligned}
&= \int_0^\infty e^{-\alpha\rho} P\left(t, (\rho+t)\wedge T\right) f\left((\rho+t)\wedge T, \cdot\right)(x) d\rho \\
&= \int_0^\infty e^{-\alpha\rho} S(\rho) f(t,x)\, d\rho \\
&= \mathbb{E}_{t,x}\left[\int_0^\infty e^{-\alpha\rho} f\left((\rho+t)\wedge T, X\left((\rho+t)\wedge T\right)\right) d\rho\right], \qquad (4.6)
\end{aligned}
$$

$f \in C_b([0,T]\times E)$, $(t,x) \in [0,T]\times E$. In order to prove that the family $\{R(\alpha) : \Re\alpha > 0\}$ is a resolvent family indeed it suffices to establish that the family $\{S(\rho) : \rho \geqslant 0\}$ is a semigroup. Let $f \in C_b([0,T]\times E)$ and fix $0 \leqslant \rho_1$, $\rho_2 < \infty$. Then this fact is a consequence of the following identities:

$$
\begin{aligned}
&S(\rho_1) S(\rho_2) f(t,x) \\
&= P\left(t, (\rho_1+t)\wedge T\right)\left[y \mapsto S(\rho_2) f\left((\rho_1+t)\wedge T, y\right)\right](x) \\
&= P\left(t, (\rho_1+t)\wedge T\right) \\
&\qquad \left[y \mapsto P\left((\rho_1+t)\wedge T, (\rho_2+\rho_1+t)\wedge T\right) f\left((\rho_2+\rho_1+t)\wedge T, y\right)\right](x)
\end{aligned}
$$

(use evolution property)

$$
\begin{aligned}
&= P\left(t, (\rho_2+\rho_1+t)\wedge T\right) f\left((\rho_2+\rho_1+t)\wedge T, \cdot\right)(x) \\
&= S(\rho_2+\rho_1) f(t,x). \qquad (4.7)
\end{aligned}
$$

Let $D_1 : C_b^{(1)}[0,T] \to C_b([0,T])$ be the time derivative operator. Then the space-time operator $D_1 + L$ defined by

$$
(D_1 + L) f(t,x) = D_1 f(t,x) + L(t) f(t,\cdot)(x), \quad f \in D(D_1 + L),
$$

turns out to be the generator of the semigroup $\{S(\rho) : \rho \geqslant 0\}$. We also observe that once the semigroup $\{S(\rho) : \rho \geqslant 0\}$ is known, the Feller evolution $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ can be recovered by the formula:

$$
P(\tau,t) f(x) = S(t-\tau) f(\tau,x), \quad f \in C_b(E), \qquad (4.8)
$$

where at the right-hand side of (4.8) the function f is considered as the function in $C_b([0,T]\times E)$ given by $(s,y) \mapsto f(y)$. The following theorem elaborates on these concepts.

Theorem 4.1. *Let $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ be a Feller propagator. Define the corresponding $\mathcal{T}_\beta$-continuous semigroup $\{S(\rho) : \rho \geqslant 0\}$ as in (4.5). Define the resolvent family $\{R(\alpha) : \alpha > 0\}$ as in (4.6). Let $L^{(1)}$ be its generator. Then $\left(\alpha I - L^{(1)}\right) R(\alpha) f = f$, $f \in C_b([0,T]\times E)$, $R(\alpha)\left(\alpha I - L^{(1)}\right) f = f$, $f \in D\left(L^{(1)}\right)$, and $L^{(1)}$ extends $D_1 + L$. Conversely, if the operator $L^{(1)}$ is defined by $L^{(1)} R(\alpha) f = \alpha R(\alpha) f - f$, $f \in C_b([0,T]\times E)$, then $L^{(1)}$ generates the semigroup $\{S(\rho) : \rho \geqslant 0\}$, and $L^{(1)}$ extends the operator $D_1 + L$.*

Proof. By definition we know that

$$L^{(1)}f = \mathcal{T}_\beta\text{-}\lim_{t\downarrow 0} \frac{1}{t}\left(S(t) - S(0)\right) f, \quad f \in D\left(L^{(1)}\right). \tag{4.9}$$

Here $D\left(L^{(1)}\right)$ is the subspace of those $f \in C_b\left([0,T]\times E\right)$ for which the limit in (4.9) exists. Fix $f \in C_b\left([0,T]\times E\right)$, and $\alpha > 0$. Then

$$\begin{aligned}
&\left(I - e^{-\alpha t}S(t)\right)\int_0^\infty e^{-\alpha\rho}S(\rho) f d\rho \\
&= \int_0^\infty e^{-\alpha\rho}S(\rho)f\, d\rho - \int_0^\infty e^{-\alpha\rho}e^{-\alpha(t+\rho)}S\left(t\right)S\left(\rho\right)f\, d\rho \\
&= \int_0^\infty e^{-\alpha\rho}S(\rho)f\, d\rho - \int_t^\infty e^{-\alpha\rho}S(\rho)f\, d\rho = \int_0^t e^{-\alpha\rho}S(\rho)f\, d\rho.
\end{aligned} \tag{4.10}$$

From (4.10) it follows that $R(\alpha)f \in D\left(L^{(1)}\right)$, and that $\left(\alpha I - L^{(1)}\right)R(\alpha)f = f$. Conversely, let $f \in D\left(L^{(1)}\right)$. Then we have

$$\begin{aligned}
&R(\alpha)\left(\alpha f - L^{(1)}f\right) = R(\alpha)\mathcal{T}_\beta\text{-}\lim_{t\downarrow 0}\frac{1}{t}\left(f - e^{-\alpha t}S(t)f\right)d\rho \\
&= \mathcal{T}_\beta\text{-}\lim_{t\downarrow 0}\frac{1}{t}\left(R(\alpha)f - R(\alpha)e^{-\alpha t}S(t)f\right)d\rho = \mathcal{T}_\beta\text{-}\lim_{t\downarrow 0}\frac{1}{t}\int_0^t e^{-\alpha\rho}S(\rho)fd\rho = f.
\end{aligned} \tag{4.11}$$

The first part of Theorem 4.1 follows from (4.10) and (4.11). In order to show that $L^{(1)}$ extends $D_1 + L$ we recall the definition of generator of a Feller evolution as given in Definition 2.7: $L(s)f = \mathcal{T}_\beta\text{-}\lim_{t\downarrow s}\dfrac{P(s,t)f - f}{t-s}$. So that if $f \in D^{(1)}\left(L\right)$, then $f \in D\left(L^{(1)}\right)$, and $L^{(1)}f = D_1 f + Lf$. Recall that $Lf(s,x) = L(s)f\left(s,\cdot\right)(x)$. Next, if the operator L_0 is defined by $L_0R(\alpha)f = \alpha R(\alpha)f - f$, $f \in C_b\left([0,T]\times E\right)$. Then necessarily we have $L_0 = L^{(1)}$, and hence L_0 generates the semigroup $\{S(\rho) : \rho \geqslant 0\}$. Altogether this proves Theorem 4.1. □

In the next theorem we establish a version of the Lumer-Phillips theorem: see [Lumer and Phillips (1961)], and Theorem 11.22 in [Renardy and Rogers (2004)].

Theorem 4.2. *Let L be a linear operator with domain $D(L)$ and range $R(L)$ in $C_b(E)$. The following assertions are equivalent:*

(i) The operator L is $\mathcal{T}_\beta$-closable and its $\mathcal{T}_\beta$-closure generates a Feller semigroup.

(ii) The operator L verifies the maximum principle, its domain $D(L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$, it is $\mathcal{T}_\beta$-dissipative and sequentially λ-dominant for some $\lambda > 0$, and there exists $\lambda_0 > 0$ such that the range $R(\lambda_0 I - L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$.

In the definitions 4.1 – 4.3 the notions of maximum principle, dissipativity, and sequential λ-dominance are explained. In the proof we will employ the results of Proposition 2.4.

Definition 4.1. An operator L with domain and range in $C_b(E)$ is said to satisfy the maximum principle, if for every $f \in D(L)$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ with the following properties:

$$\lim_{n\to\infty} \Re f(x_n) = \sup_{x\in E} \Re f(x), \quad \text{and} \quad \lim_{n\to\infty} \Re Lf(x_n) \leqslant 0. \tag{4.12}$$

In assertion (b) of Proposition 4.3 it will be shown that (4.12) is equivalent to the inequality in (4.46).

Definition 4.2. An operator L with domain and range in $C_b(E)$ is called dissipative if

$$\|\lambda f - Lf\|_\infty \geqslant \lambda \|f\|_\infty, \quad \text{for all } \lambda > 0, \text{ and for all } f \in D(L). \tag{4.13}$$

An operator L with domain and range in $C_b(E)$ is called $\mathcal{T}_\beta$-dissipative if there exists $\lambda_0 \geqslant 0$ such that for every function $u \in H^+(E)$ there exists a function $v \in H^+(E)$ such that

$$\|v(\lambda f - Lf)\|_\infty \geqslant \lambda \|uf\|_\infty, \quad \text{for all } \lambda \geqslant \lambda_0, \text{ and all } f \in D(L). \tag{4.14}$$

An operator L with domain and range in $C_b(E)$ is called positive $\mathcal{T}_\beta$-dissipative if there exists $\lambda_0 > 0$ such that for every function $u \in H^+(E)$ there exists a function $v \in H^+(E)$ for which

$$\sup_{x\in E} v(x)\Re(\lambda f(x) - Lf(x)) \geqslant \lambda \sup_{x\in E} u(x)\Re f(x), \tag{4.15}$$

for all $\lambda \geqslant \lambda_0$, and for all $f \in D(L)$.

The definition which follows is crucial in proving that an operator L (or its $\mathcal{T}_\beta$-closure) generates a $\mathcal{T}_\beta$-continuous Feller semigroup. The symbol $\mathcal{K}(E)$ stands for the collection of compact subsets of E. The mapping $f \mapsto U_\lambda^1(f)$, $f \in C_b(E,\mathbb{R})$, was introduced in (2.42).

Definition 4.3. Let L be an operator with domain and range in $C_b(E)$ and fix $\lambda > 0$. Let $f \in C_b(E,\mathbb{R})$, $\lambda > 0$, and put

$$U_\lambda^1(f) = \sup_{K\in\mathcal{K}(E)} \inf_{g\in D(L)} \{g \geqslant f\mathbf{1}_K : (\lambda I - L)g \geqslant 0\}. \tag{4.16}$$

The operator L is called sequentially λ-dominant if for every sequence $(f_n)_{n\in\mathbb{N}}$, which decreases pointwise to zero, the sequence $\left(f_n^\lambda = U_\lambda^1\left(f_n\right)\right)_{n\in\mathbb{N}}$ defined as in (4.16) possesses the following properties:

(1) The function f_n^λ dominates f_n: $f_n \leqslant f_n^\lambda$, and
(2) The sequence $\left(f_n^\lambda\right)_{n\in\mathbb{N}}$ converges to zero uniformly on compact subsets of E: $\lim_{n\to\infty} \sup_{x\in K} f_n^\lambda(x) = 0$ for all $K \in \mathcal{K}(E)$.

The functions f_n^λ automatically have the first property, provided that the constant functions belong to $D(L)$ and that $L\mathbf{1} = 0$. The real condition is given by the second property. Some properties of the mapping U_λ^1 : $C_b\left(E,\mathbb{R}\right) \to L^\infty\left(E,\mathcal{E},\mathbb{R}\right)$ were explained in Proposition 2.4.

If in Definition 4.3 U_λ^1 is a mapping from $C_b\left(E,\mathbb{R}\right)$ to itself, then Dini's lemma implies that in (2) uniform convergence on compact subsets of E may be replaced by pointwise convergence on E.

Remark 4.1. Suppose that the operator L in Definition 4.3 satisfies the maximum principle and that $\left(\mu I - L\right) D(L) = C_b(E)$, $\mu > 0$. Then the inverses $R(\mu) = \left(\mu I - L\right)^{-1}$, $\mu > 0$, exist and represent positivity preserving operators. If a function $g \in D(L)$ is such that $\left(\lambda I - L\right) g \geqslant 0$, then $g \geqslant 0$ and $\left(\left(\lambda + \mu\right) I - L\right) g \geqslant \mu g$, $\mu \geqslant 0$. It follows that $g \geqslant \mu R\left(\lambda + \mu\right) g$, $\mu \geqslant 0$. In the literature functions $g \in C_b(E)$ with the latter property are called λ-super-median. For more details see e.g. [Sharpe (1988)]. If the operator L generates a Feller semigroup $\{S(t) : t \geqslant 0\}$, then a function $g \in C_b(E)$ is called λ-super-mean valued if for every $t \geqslant 0$ the inequality $e^{-\lambda t}S(t)g \leqslant g$ holds pointwise. In Lemma (9.12) in [Sharpe (1988)] it is shown that, essentially speaking, these notions are equivalent. In fact the proof is not very difficult. It uses the Hausdorff-Bernstein-Widder theorem about the representation by Laplace transforms of positive Borel measures on $[0,\infty)$ of completely positive functions. It is also implicitly proved in the proof of Theorem 4.3 implication (iii) $\Longrightarrow$ (i): see (in-)equalities (4.131), (4.132), (4.133), (4.134), and (4.140).

Proof. [Proof of Theorem 4.2.] (i) $\Longrightarrow$ (ii). Let $\overline{L}$ be the $\mathcal{T}_\beta$-closure of L, which is the $\mathcal{T}_\beta$-generator of the semigroup $\{S(t) : t \geqslant 0\}$. Then $R\left(\lambda I - \overline{L}\right) = C_b(E)$, and the inverses of $\lambda I - \overline{L}$ which we denote by $R(\lambda)$ exist and satisfy: $R(\lambda)f(x) = \int_0^\infty e^{-\lambda t}S(t)f(x)$. It follows that

$$\lambda\Re\left(R(\lambda)f(x)\right) = \lambda\int_0^\infty e^{-\lambda t}\left(S(t)\Re f\right)(x)dt$$

$$\leqslant \lambda \int_0^\infty e^{-\lambda t} dt \sup_{y \in E} \Re f(y) = \sup_{y \in E} \Re f(y), \tag{4.17}$$

and hence $\sup_{x\in E} \lambda \Re \left(R(\lambda) f(x)\right) \leqslant \sup_{y\in E} \Re f(y)$. The substitution $f = \lambda g - \overline{L} g$ yields:

$$\lambda \sup_{x\in E} \Re g(x) \leqslant \sup_{y\in E} \Re \left(\lambda g(y) - \overline{L} g(y)\right), \quad g \in D\left(\overline{L}\right). \tag{4.18}$$

In other words, the operator $\overline{L}$ satisfies the maximum principle, and so does the operator L: see Proposition 4.3 assertion (b) below. Since the operator L is $\mathcal{T}_\beta$-dissipative, the resolvent families $\{R(\lambda) : \lambda \geqslant \lambda_0\}$, $\lambda_0 > 0$, are $\mathcal{T}_\beta$-equi-continuous. Hence every operator $R(\lambda)$ can be written as an integral: $R(\lambda) f(x) = \int f(y) r(\lambda, x, dy)$, $f \in C_b(E)$. For this the reader may consider the arguments in (the proof of) Proposition 2.4. Moreover, we have that for every $\lambda_0 > 0$, the family $\left\{e^{-\lambda_0 t} S(t) : t \geqslant 0\right\}$ is $\mathcal{T}_\beta$-equi-continuous, and in addition, $\lim_{t\downarrow 0} S(t) f(x) = f(x)$, $f \in C_b(E)$. It then follows that $\lim_{\lambda\to\infty} \lambda R(\lambda) f(x) = f(x)$, $f \in C_b(E)$. As in the proof of Proposition 2.4 we see that $\mathcal{T}_\beta$-$\lim_{\lambda\to\infty} \lambda R(\lambda) f = f$, $f \in C_b(E)$: see e.g. (2.59). Let $f \geqslant 0$ belong to $C_b(E)$, and consider the function $U_\lambda^1(f)$ defined by

$$U_\lambda^1(f) = \sup_{K \in \mathcal{K}(E)} \inf_{g \in D(L)} \left\{ g \geqslant f \mathbf{1}_K : \lambda g - L g \geqslant 0 \right\}. \tag{4.19}$$

In fact this definition is copied from (2.42). As was shown in Proposition 2.4, we have the following equality:

$$\begin{aligned} U_\lambda^1(f) &= \sup \left\{ \mu^k \left((\lambda + \mu) I - \overline{L} \right)^{-k} f : \mu > 0,\, k \in \mathbb{N} \right\} \\ &= \sup \left\{ e^{-\lambda t} S(t) f : t \geqslant 0 \right\}. \end{aligned} \tag{4.20}$$

In fact in Proposition 2.4 the first equality in (4.20) was proved. The second equality follows from the representations:

$$\left(\mu R(\lambda + \mu)\right)^k f = \frac{\mu^k}{(k-1)!} \int_0^\infty t^{k-1} e^{-\mu t} e^{-\lambda t} S(t) f dt \quad \text{and} \tag{4.21}$$

$$e^{-\lambda t} S(t) f = \mathcal{T}_\beta\text{-}\lim_{\mu\to\infty} e^{-\mu t} \sum_{k=0}^\infty \frac{(\mu t)^k}{k!} \left(\mu R(\lambda + \mu)\right)^k f. \tag{4.22}$$

A similar argument will be used in the proof of Theorem 4.3 (iii) $\Longrightarrow$ (i): see (4.133) and (4.134). The representation in (4.20) implies that the operator L is λ-dominant. Altogether this proves the implication (i) $\Longrightarrow$ (ii) of Theorem 4.2.

(ii) $\Longrightarrow$ (i). As in Proposition 4.3 assertion (a) below, the operator L is $\mathcal{T}_\beta$-closable. Let $\overline{L}$ be its $\mathcal{T}_\beta$-closure. Then the operator $\overline{L}$ is $\mathcal{T}_\beta$-dissipative, λ-dominant, and satisfies the maximum principle. In addition $R\left(\lambda I - \overline{L}\right) = C_b(E)$, $\lambda > 0$. Consequently, the inverses $R(\lambda) = \left(\lambda I - \overline{L}\right)^{-1}$, $\lambda > 0$, exist. The formulas in (4.21) and (4.22) can be used to represent the powers of the resolvent operators, and to define the $\mathcal{T}_\beta$-continuous semigroup generated by $\overline{L}$. The λ-dominance is used in a crucial manner to prove that the semigroup represented by (4.22) is a $\mathcal{T}_\beta$-equi-continuous semigroup which consists of operators, which assign bounded continuous functions to such functions. For details the reader is referred to the proof of Theorem 4.3 implication (iii) $\Longrightarrow$ (i), where a very similar construction is carried for a time space operator $L^{(1)}$ which is the $\mathcal{T}_\beta$-closure of $D_1 + L$. In Theorem 4.3 the operator D_1 is taking derivatives with respect to time, and L generates a Feller evolution.

The proof of Theorem 4.2 is complete now. □

In the context of $\mathcal{T}_\beta$-continuous Feller semigroups we establish a generation result.

Proposition 4.2. *Let L be a $\mathcal{T}_\beta$-closed linear operator with domain and range in $C_b(E)$. Suppose that the operator L satisfies the maximum principle, and is such that $R(\lambda I - L) = C_b(E)$, $\lambda > 0$. Then the resolvent family $\left\{R(\lambda) = (\lambda I - L)^{-1} : \lambda > 0\right\}$ consists of positivity preserving operators. In addition, suppose that L possesses a $\mathcal{T}_\beta$-dense domain, and that the following limits exist: for all $(t,x) \in [0,\infty) \times E$ and for all $f \in C_b(E)$*

$$e^{-\lambda t} S(t) f(\tau, x) = \lim_{\mu\to\infty} e^{-\mu t} \sum_{k=0}^{\infty} \frac{(\mu t)^k}{k!} \left(\mu R(\lambda+\mu)\right)^k f(\tau, x), \tag{4.23}$$

and for all $f \in D(L)$ and $x \in E$

$$\lim_{\mu\to\infty} \mu \left(I - \mu R(\lambda+\mu)\right) f(x) = \lambda f(x) - L f(x). \tag{4.24}$$

Moreover, suppose that the operators $R(\lambda)$, $\lambda > 0$, are $\mathcal{T}_\beta$-continuous. Fix $f \in C_b(E)$, $f \geqslant 0$, and $\lambda > 0$. The following equalities and inequality hold true:

$$\sup_{K\in\mathcal{K}(E)} \inf_{g\in D(L)} \left\{ g \geqslant f \mathbf{1}_K : (\lambda I - L) g \geqslant 0 \right\} \tag{4.25}$$

$$= \sup_{K\in\mathcal{K}(E)} \inf_{g\in C_b(E)} \left\{ g \geqslant f \mathbf{1}_K : g \geqslant \mu R(\lambda+\mu) g, \text{ for all } \mu > 0 \right\} \tag{4.26}$$

$$\geqslant \sup \left\{ \left(\mu R(\lambda+\mu)\right)^k f : \mu \geqslant 0,\, k \in \mathbb{N} \right\} \tag{4.27}$$

$$= \sup \left\{ e^{-\lambda t} S(t) f : t \geqslant 0 \right\}. \tag{4.28}$$

If the function $(t,x) \mapsto S(t)f(x)$ *is continuous, then the function* $g = \sup\left\{e^{-\lambda t}S(t)f : t \geqslant 0\right\}$ *is continuous, realizes the infimum in (4.26), and the expressions (4.25) through (4.28) are all equal.*

The proof of the following corollary is an immediate consequence of Proposition 4.2.

Corollary 4.1. *Suppose that the operator* L *with domain and range in* $C_b(E)$ *be the* $\mathcal{T}_\beta$*-generator of a Feller semigroup* $\{S(t) : t \geqslant 0\}$. *Let* $f \geqslant 0$ *belong to* $C_b(E)$. *Then the quantities in (4.25) through (4.28) are equal.*

Let $g \in D(L)$. By assumption (4.24) we see that $\lambda g - Lg \geqslant 0$ if and only if $g \geqslant \mu R(\lambda+\mu)\, g$ for all $\mu > 0$. Hence we have

$$\begin{aligned}&\inf_{g\in D(L)} \left\{g \geqslant f\mathbf{1}_K : (\lambda I - L)\, g \geqslant 0\right\}\\ &= \inf_{g\in D(L)} \left\{g \geqslant f\mathbf{1}_K : g \geqslant \mu R(\lambda+\mu)g, \text{ for all } \mu > 0\right\}.\end{aligned} \tag{4.29}$$

It is not so clear under what conditions we have equality of (4.29) and (4.26). If $f \in D(L)$ is such that $\lambda f - Lf \geqslant 0$, then the functions in (4.25) through (4.28) are all equal to f.

Proof. [Proof of Proposition 4.2.] The representation in (4.23) shows that the term in (4.28) is dominated by the one in (4.27). The equality

$$(\mu R(\lambda+\mu))^k f = \frac{\mu^k}{(k-1)!}\int_0^\infty t^{k-1}e^{-(\lambda+\mu)t}S(t)f\,dt, \quad k \geqslant 1, \tag{4.30}$$

shows that the expression in (4.27) is less than or equal to the one in (4.28). Altogether this proves the equality of (4.27) and (4.28). If the function $g \in D(L)$ is such that $g \geqslant f\mathbf{1}_K$ and $(\lambda I - L)\, g \geqslant 0$, then $((\lambda+\mu) I - L)\, g \geqslant \mu g$, and hence

$$g \geqslant \mu R(\lambda+\mu)\, g \geqslant (\mu R(\lambda+\mu))^k\, g \quad \text{for all } k \in \mathbb{N}.$$

Consequently, the term in (4.25) dominates the second one. It also follows that the expression in (4.26) is greater than or equal to

$$\sup_{K\in\mathcal{K}(E)} \sup\left\{(\mu R(\lambda+\mu))^k (f\mathbf{1}_K) : \mu > 0,\, k \in \mathbb{N}\right\}. \tag{4.31}$$

Since the operators $(\mu R(\lambda+\mu))^k$, $\mu > 0$ and $k \in \mathbb{N}$, are $\mathcal{T}_\beta$-continuous the expression in (4.31) is equal to the quantity in (4.27). Next we will show that the expression in (4.26) is less than or equal to (4.25). Therefore we chose an arbitrary compact subset K of E. Let $g \in C_b(E)$ be a function with the following properties: $g \geqslant f\mathbf{1}_K$, and $g \geqslant \mu R(\lambda+\mu)\, g$. Then for $\eta > 0$ arbitrary small and $\alpha = \alpha_\eta > 0$ sufficiently large we have $\alpha R(\alpha)\,(g+\eta) \geqslant g\mathbf{1}_K \geqslant f\mathbf{1}_K$. Moreover, the function $g_{\alpha,\eta} := \alpha R(\alpha)\,(g+\eta)$ belongs to $D(L)$ and satisfies

$$g_{\alpha,\eta} \geqslant \mu R(\lambda+\mu)\, g_{\alpha,\eta} \quad \text{for all } \mu > 0 \tag{4.32}$$

Here we employed the fact that $D(L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$. In fact we used the fact that, uniformly on the compact subset K, $g + \eta = \lim_{\alpha\to\infty} \alpha R(\alpha)\,(g+\eta)$. From (4.32) we obtain

$$(\lambda I - L)\, g_{\alpha,\eta} = \lim_{\mu\to\infty} \mu\,(I - \mu R(\lambda+\mu))\, g_{\alpha,\eta} \geqslant 0, \tag{4.33}$$

From (4.33) we obtain the inequality:

$$\begin{aligned} &\inf_{g\in C_b(E)} \{g \geqslant f\mathbf{1}_K : g \geqslant \mu R(\lambda+\mu)\, g\} \\ &\geqslant \inf_{g\in D(L)} \{g \geqslant f\mathbf{1}_K : g \geqslant \mu R(\lambda+\mu)\, g\}. \end{aligned} \tag{4.34}$$

The inequality in (4.34) shows that the expression in (4.26) is less than or equal to the one in (4.25). Thus far we showed (4.25) = (4.26) $\geqslant$ (4.27) = 4.28). The final assertion about the fact that the (continuous) function in (4.28) realizes the equality in (4.27) being obvious, concludes the proof of Proposition 4.2. □

In the following theorem (Theorem 4.3) we use the following subspaces of the space $C_b([0,T]\times E)$:

$$\begin{aligned} C_{P,b}^{(1)} = \Big\{ & f \in C_b([0,T]\times E) : \text{all functions of the form } (\tau,x) \mapsto \\ & \int_\tau^{\tau+\rho} P(\tau,\sigma)\, f(\sigma,\cdot)(x)d\sigma,\ \rho>0, \text{ belong to } D(D_1) \Big\}; \end{aligned} \tag{4.35}$$

$$\begin{aligned} C_{P,b}^{(1)}(\lambda) = \Big\{ & f \in C_b([0,T]\times E) : \text{the function } (\tau,x) \mapsto \\ & \int_\tau^{\infty} e^{-\lambda\sigma} P(\tau,\sigma)\, f(\sigma,\cdot)(x)d\sigma, \text{ belongs to } D(D_1) \Big\}. \end{aligned} \tag{4.36}$$

Here $\lambda > 0$, and $C_{P,b}^{(1)}$ is a limiting case if $\lambda = 0$. The inclusion $C_{P,b}^{(1)} \subset \bigcap_{\lambda_0>0} C_{P,b}^{(1)}(\lambda_0)$ follows from the representation of $R(\lambda_0)$ as a Laplace transform:

$$R(\lambda_0)\, f(\tau,x) = \int_0^\infty e^{-\lambda_0\rho} S(\rho) f(\tau,x)d\rho$$

$$= \int_0^\infty e^{-\lambda_0\rho} P(\tau, \tau+\rho) f(\tau+\rho, x) d\rho$$
$$= \lambda_0 \int_0^\infty e^{-\lambda_0\rho} \int_0^\rho P(\tau, \tau+\sigma) f(\tau+\sigma, x) d\sigma\, d\rho$$
$$= \lambda_0 \int_0^\infty e^{-\lambda_0\rho} \int_\tau^{\tau+\rho} P(\tau, \sigma) f(\sigma, x) d\sigma\, d\rho \tag{4.37}$$

From (4.37) we see that if for every $\rho > 0$ the function

$$(\tau, x) \mapsto \int_0^\rho S(\sigma) f(\tau, x) d\sigma = \int_\tau^{\tau+\rho} P(\tau, \sigma \wedge T) f(\sigma \wedge T, x) d\sigma$$

belongs to $D(D_1)$, then so does the function $(\tau, x) \mapsto R(\lambda_0) f(\tau, x)$, provided that the function $\rho \mapsto e^{-\lambda_0\rho} D_1 \int_0^\rho S(\sigma) f d\sigma$ is $\mathcal{T}_\beta$-integrable in the space $C_b([0,T] \times E)$. The other inclusion, i.e. $\bigcap_{\lambda_0 > 0} C_{P,b}^{(1)}(\lambda_0) \subset C_{P,b}^{(1)}$ follows from the following inversion formula:

$$\int_\tau^{\tau+\rho} P(\tau, \sigma) f(\sigma, \cdot)(x) d\sigma = \int_0^\rho S(\sigma) f(\tau, x)\, d\sigma$$
$$= \lim_{\lambda\to\infty} \int_0^\rho e^{-\sigma\lambda} e^{\sigma\lambda^2 R(\lambda)} f(\tau, x)\, d\sigma$$
$$= \lim_{\lambda\to\infty} \sum_{k=0}^\infty \frac{1}{k!} \int_0^\rho (\sigma\lambda)^k e^{-\sigma\lambda} (\lambda R(\lambda))^k f(\tau, x)\, d\sigma$$
$$= \lim_{\lambda\to\infty} \sum_{k=0}^\infty \frac{\lambda^{k+1}}{(k+1)!} \int_0^\rho (\sigma\lambda)^{k+1} e^{-\sigma\lambda} (R(\lambda))^{k+1} f(\tau, x)\, d\sigma$$
$$= \lim_{\lambda\to\infty} \sum_{k=0}^\infty \frac{\lambda^{k+1}}{(k+1)!k!} \int_0^\rho (\sigma\lambda)^{k+1} e^{-\sigma\lambda} \int_0^\infty \rho_1^k e^{-\lambda\rho_1} S(\rho_1) f(\tau, x)\, d\rho_1\, d\sigma$$
$$= \lim_{\lambda\to\infty} \sum_{k=0}^\infty \frac{(-1)^k\lambda^{k+1}}{(k+1)!k!} \int_0^\rho (\sigma\lambda)^{k+1} e^{-\sigma\lambda} \frac{\partial^k}{(\partial\lambda)^k} \int_0^\infty e^{-\lambda\rho_1} S(\rho_1) f(\tau, x)\, d\rho_1\, d\sigma$$
$$= \lim_{\lambda\to\infty} \sum_{k=0}^\infty \frac{(-1)^k\lambda^{k+1}}{(k+1)!k!} \int_0^\rho (\sigma\lambda)^{k+1} e^{-\sigma\lambda} \frac{\partial^k}{(\partial\lambda)^k} R(\lambda) f(\tau, x) d\sigma \tag{4.38}$$

where the limits have to be taken in $\mathcal{T}_\beta$-sense. A similar limit representation is valid for $D_1 \int_0^\rho S(\rho) f d\rho(\tau, x)$, provided that the family

$$\left\{ \frac{\lambda^{k+1}}{k!} D_1 R(\lambda)^k f : \lambda > 0,\ k \in \mathbb{N} \right\}$$

is uniformly bounded. A simpler approach might be to use a complex inversion formula:

$$\int_0^\rho (\tau + \rho - \sigma) P(\tau, (\tau+\sigma) \wedge T) f((\tau+\sigma) \wedge T, \cdot)(x) d\sigma$$

$$= \int_0^{\rho} \int_0^{\rho_1} S(\sigma) f(\tau, x) d\sigma \, d\rho_1 = \frac{1}{2\pi i} \int_{\omega - i\infty}^{\omega + i\infty} e^{\rho\lambda} \frac{1}{\lambda^2} R(\lambda) f(\tau, x) d\lambda, \quad (4.39)$$

and to assume that, for $\omega > 0$, the family $\{\lambda D_1 R(\lambda) f : \Re\lambda \geqslant \omega\}$ is uniformly bounded. It is clear that the operator $R(\lambda)$, $\Re\lambda > 0$, stands for

$$R(\lambda) f(\tau, x) = \int_0^{\infty} e^{-\lambda\rho} S(\rho) f(\tau, x) d\rho \quad (4.40)$$

$$= \int_0^{\infty} e^{-\lambda\rho} P(\tau, (\tau + \rho) \wedge T) f((\tau + \rho) \wedge T, \cdot)(x) d\rho, \quad f \in C_b([0, T] \times E).$$

It is also clear that the family of operators in (4.38) is a once integrated integrated semigroup, and that the family in (4.39) is a twice integrated semigroup. In order to justify the inclusion $\bigcap_{\lambda_0 > 0} C_{P,b}^{(1)}(\lambda_0) \subset C_{P,b}^{(1)}$ in both approaches we need to know that the functions: $\lambda \mapsto R(\lambda) f$, and $\lambda \mapsto D_1 R(\lambda) f$ are real analytic. For more details on inversion formulas for vector-valued Laplace transforms and integrated semigroups see e.g. [Bobrowski (1997)], [Chojnacki (1998)], [Arendt (1987)], [Arendt *et al.* (2001)], and [Miana (2005)]. For vector valued Laplace transforms the reader is also referred to [Bäumer and Neubrander (1994)].

Theorem 4.3. *Let L be a linear operator with domain $D(L)$ and range $R(L)$ in $C_b([0, T] \times E)$. Suppose that there exists $\lambda > 0$ such that the operator $D_1 + L$ is sequentially λ-dominant in the sense of Definition 4.3. Under such a hypothesis the following assertions are equivalent:*

(i) The operator L is $\mathcal{T}_\beta$-closable, its $\mathcal{T}_\beta$-closure generates a Feller evolution, the operator $D_1 + L$ is $\mathcal{T}_\beta$-densely defined, and there exists $\lambda_0 > 0$ such that the subspace $C_{P,b}^{(1)}(\lambda_0)$ is $\mathcal{T}_\beta$-dense in $C_b([0, T] \times E)$.

(ii) The operator $D_1 + L$ is $\mathcal{T}_\beta$-closable and its $\mathcal{T}_\beta$-closure generates a $\mathcal{T}_\beta$-continuous Feller semigroup in $C_b([0, T] \times E)$.

(iii) The operator $D_1 + L$ is $\mathcal{T}_\beta$-densely defined, is positive $\mathcal{T}_\beta$-dissipative, satisfies the maximum principle, and there exists $\lambda_0 > 0$ such that the range of $\lambda_0 I - D_1 - L$ is $\mathcal{T}_\beta$-dense in $C_b([0, T] \times E)$.

Theorem 4.3 will be proved in Section 4.2 after the proof of Proposition 4.3.

Remark 4.2. Let us call the operator $D_1 + L$ power $\mathcal{T}_\beta$-dissipative if for some $\lambda_0 \geqslant 0$ and for every $k \in \mathbb{N}$ there exists a $\mathcal{T}_\beta$-dense subspace D_k of $C_b([0, T] \times E)$ such that for every $u \in H^+([0, T] \times E)$ there exists $v \in H^+([0, T] \times E)$ for which the following inequality holds:

$$\lambda^k \|uf\|_\infty \leqslant \left\| v (\lambda I - D_1 - L)^k f \right\|_\infty \quad \text{for all } f \in D_k \text{ and all } \lambda \geqslant \lambda_0. \quad (4.41)$$

If the operator $D_1 + L$ is just $\mathcal{T}_\beta$-dissipative, then an inequality of the form (4.41) holds, with a function $v \in H^+([0,T] \times E)$ which depends on k. In (4.41) the function v only depends on u (and the operator $D_1 + L$), but it neither depends on k, nor on $f \in D_k$ or $\lambda \geqslant 1$. Let the operator $L^{(1)}$ be an extension of $D_1 + L$ which generates a $\mathcal{T}_\beta$-continuous semigroup $\{S_0(t) : t \geqslant 0\}$, and suppose that $D_1 + L$ satisfies (4.41). Then this semigroup is equi-continuous in the sense that for every $u \in H^+([0,T] \times E)$ there exists $v \in H^+([0,T] \times E)$ for which the following inequality holds:

$$\|uS_0(t)f\|_\infty \leqslant \|vf\|_\infty \quad \text{for all } f \in C_b([0,\infty) \times E) \text{ and all } t \in [0,\infty). \tag{4.42}$$

A closely related inequality is the following one

$$\left\|u(\lambda R(\lambda))^k f\right\|_\infty \leqslant \|vf\|_\infty, \quad \lambda \geqslant 1, \ f \in C_b([0,T] \times E). \tag{4.43}$$

Notice that (4.43) is equivalent to (4.41) provided that the operator $L^{(1)}$ is the $\mathcal{T}_\beta$-closure of $D_1 + L$ and the ranges of $\lambda I - L^{(1)}$, $\lambda > 0$, coincide with $C_b([0,T] \times E)$. In fact the semigroup $\{S_0(t) : t \geqslant 0\}$ and the resolvent family $\{R(\lambda) : \lambda > 0\}$ are related as follows:

$$(\lambda R(\lambda))^k f = \frac{\lambda^k}{k!}\int_0^\infty t^{k-1} e^{-\lambda t} S_0(t) f \, dt, \quad \text{and} \tag{4.44}$$

$$S_0(t)f = \mathcal{T}_\beta\text{-}\lim_{\lambda\to\infty} e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} (\lambda R(\lambda))^k f. \tag{4.45}$$

The integral in (4.44) has to be interpreted in $\mathcal{T}_\beta$-sense. From (4.44) and (4.45) the equivalence of (4.42) and (4.43) easily follows. We also observe that (4.43) is equivalent to the following statement. For every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b([0,T] \times E)$ which decreases pointwise to zero it follows that

$$\inf_{n\in\mathbb{N}} \sup_{\lambda\geqslant 1, k\in\mathbb{N}} (\lambda R(\lambda))^k f_n = 0.$$

4.2 Dissipative operators and maximum principle

In the following proposition we collect some of the interrelationships which exist between the concepts of closability, dissipativeness, and maximum principle. A reformulation of assertion (f) in Proposition 4.3 can be found in Lemma 8.1 in Chapter 8.

Proposition 4.3.

(a_1) *Suppose that the operator L is dissipative and that its range is contained in the closure of its domain. Then the operator L is closable.*

(a_2) *Suppose that the operator L is $\mathcal{T}_\beta$-dissipative and that its range is contained in the $\mathcal{T}_\beta$-closure of its domain. Then the operator L is $\mathcal{T}_\beta$-closable.*

(b) *If the operator L satisfies the maximum principle, then*

$$\sup_{x\in E} \Re\left(\lambda f(x) - Lf(x)\right) \geqslant \lambda \sup_{x\in E} \Re f(x), \text{ for all } \lambda > 0, \text{ and for all } f \in D(L). \tag{4.46}$$

Conversely, if L satisfies (4.46), then the operator L satisfies the maximum principle. The inequality in (4.46) is equivalent to

$$\inf_{x\in E} \Re\left(\lambda f(x) - Lf(x)\right) \leqslant \lambda \inf_{x\in E} \Re f(x), \text{ for all } \lambda > 0, \text{ and for all } f \in D(L). \tag{4.47}$$

(c) *If the operator L satisfies the maximum principle, then L is dissipative.*

(d) *If the operator L satisfies the maximum principle, and if $f \in D(L)$ is such that $\lambda f - Lf \geqslant 0$ for some $\lambda > 0$, then $f \geqslant 0$.*

(e) *If the operator L is dissipative, then*

$$\left\|\lambda f - Lf\right\|_\infty \geqslant \Re\lambda \left\|f\right\|_\infty, \text{ for all } \lambda \text{ with } \Re\lambda > 0, \text{ and for all } f \in D(L). \tag{4.48}$$

(f) *The operator L is dissipative if and only if for every $f \in D(L)$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ such that $\lim_{n\to\infty} \left|f(x_n)\right| = \left\|f\right\|_\infty$, and $\lim_{n\to\infty} \Re\left(\overline{f(x_n)} Lf(x_n)\right) \leqslant 0$.*

(g) *If the operator L is positive $\mathcal{T}_\beta$-dissipative, then it is $\mathcal{T}_\beta$-dissipative.*

For the definition of an operator which is positive $\mathcal{T}_\beta$-dissipative or $\mathcal{T}_\beta$-dissipative, the reader is referred to Definition 4.1. The same is true for the other notions in Proposition 4.3.

Proof. (a_1) Let $(f_n)_{n\in\mathbb{N}} \subset D(L)$ be any sequence with the following properties:

$$\lim_{n\to\infty} f_n = 0, \quad \text{and} \quad g = \lim_{n\to\infty} Lf_n$$

exists in $C_b(E)$. Then we consider

$$\left\|(\lambda f_n + g_m) - \lambda^{-1} L\left(\lambda f_n + g_m\right)\right\|_\infty \geqslant \left\|\lambda f_n + g_m\right\|_\infty,$$

where $(g_m)_{m\in\mathbb{N}} \subset D(L)$ converges to g. First we let n tend to infinity, then λ, and finally m. This limiting procedure results in

$$\lim_{m\to\infty} \left\|g_m - g\right\|_\infty \geqslant \lim_{m\to\infty} \left\|g_m\right\|_\infty = \left\|g\right\|_\infty.$$

Hence $g = 0$.

(a_2) Let $(f_n)_{n\in\mathbb{N}} \subset D(L)$ be any sequence with the following properties:

$$\mathcal{T}_\beta\text{-}\lim_{n\to\infty} f_n = 0, \quad \text{and} \quad g = \mathcal{T}_\beta\text{-}\lim_{n\to\infty} Lf_n$$

exists in $C_b(E)$. Let $u \in H^+(E)$ be given and let the function v be as in (4.14). Then we consider

$$\left\| v\left((\lambda f_n + g_m) - \lambda^{-1} L(\lambda f_n + g_m) \right) \right\|_\infty \geqslant \| u(\lambda f_n + g_m) \|_\infty, \tag{4.49}$$

where $(g_m)_{m\in\mathbb{N}} \subset D(L)$ $\mathcal{T}_\beta$-converges to g. First we let n tend to infinity, then λ, and finally m. The result will be

$$\lim_{m\to\infty} \| vg_m - vg \|_\infty \geqslant \lim_{m\to\infty} \| ug_m \|_\infty = \| ug \|_\infty,$$

and hence $g = 0$.

(b) Let $f \in D(L)$. Then choose a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ as in (4.12). Then we have

$$\sup_{x\in E} \Re(\lambda f(x) - Lf(x)) \geqslant \lim_{n\to\infty} \Re(\lambda f(x_n) - Lf(x_n)) \geqslant \lambda \sup_{x\in E} \Re f(x)$$

which is the same as (4.46). Suppose that the operator L satisfies (4.46). Then for every $\lambda > 0$ we choose $x_\lambda \in E$ such that

$$\lambda \Re f(x_\lambda) - \Re Lf(x_\lambda) \geqslant \lambda \sup_{x\in E} \Re f(x) - \frac{1}{\lambda}. \tag{4.50}$$

From (4.50) we infer:

$$\Re Lf(x_\lambda) \leqslant \frac{1}{\lambda}, \text{ and} \tag{4.51}$$

$$\sup_{x\in E} \Re f(x) \leqslant \Re f(x_\lambda) + \frac{1}{\lambda^2} - \frac{1}{\lambda} \Re Lf(x_\lambda). \tag{4.52}$$

From (4.51) we see that $\limsup_{\lambda\to\infty} \Re Lf(x_\lambda) \leqslant 0$, and from (4.52) it follows that $\limsup_{\lambda\to\infty} \Re f(x_\lambda) = \sup_{x\in E} \Re f(x)$. From these observations it is easily seen that (4.46) implies the maximum principle.

The substitution $f \to -f$ shows that (4.47) is a consequence of (4.46).

(c) Let $f \neq 0$ belong to $D(L)$, choose $\alpha \in \mathbb{R}$ and a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ in such a way that $0 < \|f\|_\infty = \lim_{n\to\infty} \Re e^{i\alpha} f(x_n) = \sup_{x\in E} \Re e^{i\alpha} f(x)$, and that $\lim_{n\to\infty} \Re L\left(e^{i\alpha} f\right)(x_n) \leqslant 0$. Then

$$\begin{aligned} \|\lambda f - Lf\|_\infty &\geqslant \lim_{n\to\infty} \Re\left(e^{i\alpha}(\lambda f - Lf)(x_n)\right) \\ &= \lim_{n\to\infty} \lambda \Re\left(e^{i\alpha} f(x_n)\right) - \Re\left(e^{i\alpha} Lf\right)(x_n) \geqslant \lambda \|f\|_\infty. \end{aligned} \tag{4.53}$$

The inequality in (4.53) means that L is dissipative in the sense of Definition 4.2.

(d) Let $f \in D(L)$ be such that, for some $\lambda > 0$, $\lambda f(x) - Lf(x) \geqslant 0$ for all $x \in E$. From (4.47) in (b) we see that

$$\begin{aligned}\lambda \inf_{x\in E} \Im f(x) &= \lambda \inf_{x\in E} \Re(-if)(x) \geqslant \inf_{x\in E} \Re\left(\lambda(-if)(x) - L(-if)(x)\right) \\ &= \inf_{x\in E} \Im\left(\lambda f(x) - L(f)(x)\right) = 0. \qquad (4.54)\end{aligned}$$

From (4.54) we get $\Im f \geqslant 0$. If we apply the same argument to $-f$ instead of f we get $\Im f \leqslant 0$. Hence $\Im f \equiv 0$, and so the function f is real-valued. But then we have

$$0 \leqslant \inf_{x\in E} \left(\lambda f(x) - Lf(x)\right) \leqslant \lambda \inf_{x\in E} f(x),$$

and consequently $f \geqslant 0$.

(e) From the proof it follows that L is dissipative if and only if for every $f \in D(L)$ there exists an element x^* in $C_b\left([0,T] \times E\right)^*$ such that $\|x^*\| = 1$, such that $\langle f, x^*\rangle = \|f\|_\infty$, and such that $\Re\langle Lf, x^*\rangle \leqslant 0$. A proof of all this runs as follows. Let L be dissipative. Fix f in $D(L)$ and choose for each $\lambda > 0$ an element x^*_λ in $C_b\left([0,T] \times E\right)^*$ in such a way that $\|x^*_\lambda\| \leqslant 1$ and

$$\|\lambda f - Lf\|_\infty = \langle \lambda f - Af, x^*_\lambda\rangle. \qquad (4.55)$$

Choose an element x^* in the intersection $\bigcap_{\mu>0}$ weak*closure$\{x^*_\lambda : \lambda > \mu\}$. Since, by the theorem off Banach-Aloglu the dual unit ball of $C_b\left([0,T] \times E\right)^*$ is weak*-compact such an element x^* exists. From (4.55) it follows that

$$\begin{aligned}\Re\langle Lf, x^*_\lambda\rangle &= \lambda\Re\langle f, x^*_\lambda\rangle - \|\lambda f - Lf\|_\infty \\ &\leqslant \lambda\|f\|_\infty - \|\lambda f - Lf\|_\infty \leqslant 0, \quad \lambda > 0. \qquad (4.56)\end{aligned}$$

Here we used the fact that the operator L is supposed to be dissipative. From (4.55) we also obtain the equality

$$\langle f, x^*_\lambda\rangle = \left\|f - \lambda^{-1}Lf\right\|_\infty + \lambda^{-1}\langle Lf, x^*_\lambda\rangle, \quad \lambda > 0. \qquad (4.57)$$

Since x^* is a weak* limit point of $\{x^*_\lambda : \lambda > \mu\}$ for each $\mu > 0$ it follows from (4.56) and (4.57) that

$$\Re\langle Lf, x^*\rangle \leqslant 0, \text{ and} \qquad (4.58)$$

$$\langle f, x^*\rangle = \|f\|_\infty, \quad \|x^*\| \leqslant 1. \qquad (4.59)$$

Finally pick $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$. From (4.58) and (4.59) we infer

$$\|\lambda f - Lf\|_\infty \geqslant \Re\langle \lambda f - Lf, x^*\rangle = \Re\left(\lambda\langle f, x^*\rangle\right) - \Re\langle Lf, x^*\rangle$$

$$\geqslant \Re\left(\lambda\left\|f\right\|_{\infty}\right) - 0 = \Re\lambda\left\|f\right\|_{\infty}. \tag{4.60}$$

(f) If L is dissipative and if $f \in D(L)$, then there exists a family $(x_\lambda)_{\lambda>0} \subset E$ such that

$$\left|\lambda f\left(x_\lambda\right) - Lf\left(x_\lambda\right)\right| \geqslant \lambda\left\|f\right\|_{\infty} - \frac{\left\|Lf\right\|_{\infty}}{\lambda}. \tag{4.61}$$

From (4.61) we infer

$$\lambda\left|f\left(x_\lambda\right)\right| + \left\|Lf\right\|_{\infty} \geqslant \lambda\left\|f\right\|_{\infty} - \frac{\left\|Lf\right\|_{\infty}}{\lambda}, \tag{4.62}$$

and

$$\begin{aligned} &\lambda^2\left|f\left(x_\lambda\right)\right|^2 - 2\lambda\Re\left(\overline{f\left(x_\lambda\right)}Lf\left(x_\lambda\right)\right) + \left|Lf\left(x_\lambda\right)\right|^2 \\ &\geqslant \lambda^2\left\|f\right\|_2^2 - 2\left\|f\right\|_{\infty}\left\|Lf\right\|_{\infty} + \frac{\left\|Lf\right\|_{\infty}^2}{\lambda^2}. \end{aligned} \tag{4.63}$$

From (4.62) and (4.63) we easily infer

$$\left|f\left(x_\lambda\right)\right| \geqslant \left\|f\right\|_{\infty} - \frac{\left\|Lf\right\|_{\infty}}{\lambda} - \frac{\left\|Lf\right\|_{\infty}}{\lambda^2}, \tag{4.64}$$

and

$$\begin{aligned} &\lambda^2\left\|f\right\|_{\infty}^2 - 2\lambda\Re\left(\overline{f\left(x_\lambda\right)}Lf\left(x_\lambda\right)\right) + \left\|Lf\right\|_{\infty}^2 \\ &\geqslant \lambda^2\left\|f\right\|_{\infty}^2 - 2\left\|f\right\|_{\infty}\left\|Lf\right\|_{\infty} + \frac{\left\|Lf\right\|_{\infty}^2}{\lambda^2}. \end{aligned} \tag{4.65}$$

From (4.65) we get

$$\Re\left(\overline{f\left(x_\lambda\right)}Lf\left(x_\lambda\right)\right) \leqslant \frac{\left\|f\right\|_{\infty}\left\|Lf\right\|_{\infty}}{\lambda} + \frac{1}{2\lambda}\left(1 - \frac{1}{\lambda^2}\right)\left\|Lf\right\|_{\infty}^2. \tag{4.66}$$

From (4.66) we obtain $\limsup_{\lambda\to\infty}\Re\left(\overline{f\left(x_\lambda\right)}Lf\left(x_\lambda\right)\right) \leqslant 0$. From (4.64) we see $\lim_{\lambda\to\infty}\left|f\left(x_\lambda\right)\right| = \left\|f\right\|_{\infty}$. By passing to a countable sub-family we see that there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ such that $\lim_{n\to\infty}\left|f\left(x_n\right)\right| = \left\|f\right\|_{\infty}$ and such that the limit $\lim_{n\to\infty}\Re\left(\overline{f\left(x_n\right)}Lf\left(x_n\right)\right)$ exists and is $\leqslant 0$. The proof of the converse statement is (much) easier. Let $(x_n)_{n\in\mathbb{N}} \subset E$ be a sequence such that $\lim_{n\to\infty}\left|f\left(x_n\right)\right| = \left\|f\right\|_{\infty}$ and that the limit $\lim_{n\to\infty}\Re\left(\overline{f\left(x_n\right)}Lf\left(x_n\right)\right)$ exists and is $\leqslant 0$. Fix $f \in D(L)$. Then we have

$$\left\|\lambda f - Lf\right\|_{\infty}^2 \geqslant \lambda^2\left|f\left(x_n\right)\right|^2 - 2\lambda\Re\left(\overline{f\left(x_n\right)}Lf\left(x_n\right)\right) + \left|Lf\left(x_n\right)\right|^2$$

$$\geqslant \lambda^2 \left|f\left(x_n\right)\right|^2 - 2\lambda \Re\left(\overline{f\left(x_n\right)} L f\left(x_n\right)\right). \tag{4.67}$$

From the properties of the sequence $(x_n)_{n\in\mathbb{N}}$ and (4.67) we obtain the inequality $\|\lambda f - Lf\|_\infty \geqslant \lambda \|f\|_\infty$, $\lambda > 0$, $f \in D(L)$, which is the same as saying that L is dissipative.

(g) Let the functions u and $v \in H^+(E)$ as in assertion (g), let $f \in D(L)$, and $\lambda \geqslant \lambda_0$. Then we have

$$\|v(\lambda f - Lf)\|_\infty = \sup_{\vartheta\in[-\pi,\pi]} \sup_{x\in E} v(x) \Re\left(\lambda\left(e^{i\vartheta} f\right)(x) - L\left(e^{i\vartheta} f\right)(x)\right)$$

(L is positive $\mathcal{T}_\beta$-dissipative)

$$\geqslant \lambda \sup_{\vartheta\in[-\pi,\pi]} \sup_{x\in E} u(x) \Re\left(e^{i\vartheta} f(x)\right) = \lambda \|uf\|_\infty. \tag{4.68}$$

The inequality in (4.68) shows the dissipativity of the operator L.

Finally, this completes the proof of Proposition 4.3. □

Proof. [Proof of Theorem 4.3.] (i) $\Longrightarrow$ (ii). Let $\overline{L}$ be the $\mathcal{T}_\beta$-closure of L. Then there exists a Feller evolution $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ such that

$$\frac{d}{dt} P(\tau,t) f(t,\cdot)(x) = P(\tau,t)\left(D_1 + \overline{L}(t)\right) f(t,\cdot)(x), \tag{4.69}$$

for all functions $f \in D^{(1)}\left(\overline{L}\right)$, $0 \leqslant \tau \leqslant t \leqslant T$, $x \in E$. The functions $f \in D^{(1)}\left(\overline{L}\right)$ have the property that for every $\rho \in [0,T]$ the following $\mathcal{T}_\beta$-limits exist:

(a) $\overline{L}(\rho) f(\rho,\cdot)(x) = \mathcal{T}_\beta\text{-}\lim\limits_{h\downarrow 0} \dfrac{P(\rho,\rho+h) f(\rho,\cdot) - f(\rho,\cdot)}{h}$.

(b) $\dfrac{\partial}{\partial\rho} f(\rho,x) = \mathcal{T}_\beta\text{-}\lim\limits_{h\to 0} \dfrac{f(\rho+h,x) - f(\rho,x)}{h}$.

As indicated the limits in (a) and (b) have to be interpreted in $\mathcal{T}_\beta$-sense. Moreover, these functions as functions of the pair (ρ,x) are supposed to be continuous. The equality in (4.69) was introduced in Definition 2.8. However, the reader is also referred to Proposition 4.1, and to equality (4.2). The equality in (4.69) can also be written in integral form:

$$P(\tau,t) f(t,\cdot)(x) - f(\tau,x) = \int_\tau^t P(\tau,\rho)\left(\frac{\partial}{\partial\rho} + \overline{L}(\rho)\right) f(\rho,\cdot)(x) \tag{4.70}$$

for $f \in D^{(1)}\left(\overline{L}\right)$, $0 \leqslant \tau \leqslant t \leqslant T$, $x \in E$. The Feller evolution is $\mathcal{T}_\beta$-equicontinuous. This means that for every $u \in H^+(E)$, there exists $v \in H^+(E)$ such that for all $f \in C_b(E)$ the inequality

$$\sup_{\tau\leqslant t\leqslant T} \sup_{x\in E} |u(\tau,x) P(\tau,t) f(\cdot)(x)| \leqslant \sup_{x\in E} |v(x) f(x)| \tag{4.71}$$

holds for all $f \in C_b(E)$. As was explained in Corollary 3.1, the Feller evolution $\left\{\widetilde{P}(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\right\}$, which is the same as $\{P(\tau,t) : 0 \leqslant \tau \leqslant t \leqslant T\}$ considered as a family of operators on $C_b([0,T] \times E)$, is $\mathcal{T}_\beta$-equi-continuous as well: see Corollary 2.3. As in (4.5) we define the semigroup $\mathcal{T}_\beta$-equi-continuous semigroup $\{S(\rho) : \rho \geqslant 0\}$ by

$$S(\rho)f(t,x) = P(t,(\rho+t)\wedge T) f((\rho+t)\wedge T, \cdot)(x), \quad f \in C_b([0,T]\times E), \tag{4.72}$$

where $\rho \geqslant 0$, and $(t,x) \in [0,T] \times E$. Then the semigroup in (4.72) is $\mathcal{T}_\tau$-equi-continuous. In fact we have

$$\sup_{\tau\leqslant t\leqslant T}\sup_{x\in E}|u(\tau,x)S(t)f(\tau,x)| \leqslant \sup_{(\tau,x)\in[0,T]\times E}|v(x)f(\tau,x)| \tag{4.73}$$

where $u \in H^+([0,T]\times E)$ and $v \in H^+(E)$ are as in (4.71). Let $L^{(1)}$ be its generator, and $R(\lambda)f = \int_0^\infty e^{-\lambda\rho}S(\rho)f d\rho$, $f \in C_b([0,T]\times E)$, its resolvent. Then we will prove that $L^{(1)} = D_1 + \overline{L}$, and we will also show the following well-known equalities (compare with (3.154)):

$$\begin{aligned} \left(\lambda I - L^{(1)}\right) R(\lambda) f &= f, \quad f \in C_b([0,T]\times E), \\ R(\lambda)\left(\lambda I - L^{(1)}\right) f &= f, \quad f \in D\left(L^{(1)}\right). \end{aligned} \tag{4.74}$$

In order to understand the relationship between D_1+L and the $\mathcal{T}_\beta$-generator of the semigroup $\{S(\rho) : \rho \geqslant 0\}$ we consider, for $h > 0$, $\lambda > 0$ the operators $L^{(1)}_{\lambda,h}$ and $\vartheta_h L^{(1)}_{\lambda,h}$, which are defined by

$$\begin{aligned} L^{(1)}_{\lambda,h} f(\tau,x) &= \frac{1}{h}\left(I - e^{-\lambda h}S(h)\right) f(\tau,x) \qquad (4.75) \\ &= \frac{1}{h}\left(f(\tau,x) - e^{-\lambda h}P(\tau,(\tau+h)\wedge T) f((\tau+h)\wedge T,\cdot)(x)\right) \\ &= \frac{1}{h}\left(f(\tau,x) - e^{-\lambda h}P(\tau,(\tau+h)\wedge T) f(\tau,\cdot)(x)\right) \\ &\quad - \frac{1}{h}\left(e^{-\lambda h}P(\tau,(\tau+h)\wedge T)\left(f((\tau+h)\wedge T,\cdot) - f(\tau,\cdot)\right)(x)\right) \\ &= \left(\lambda I - L^{(1)}\right)\frac{1}{h}\int_0^h e^{-\lambda\rho}S(\rho) f d\rho(\tau,x) \end{aligned}$$

and

$$\begin{aligned} \vartheta_h L^{(1)}_{\lambda,h} f(\tau,x) &= \frac{1}{h}\left(I - e^{-\lambda h}S(h)\right) f((\tau-h)\wedge T \vee 0, x) \qquad (4.76) \\ &= \frac{1}{h}\left(f((\tau-h)\wedge T\vee 0,x) - f(\tau,x)\right) \end{aligned}$$

$$-\frac{1}{h}\left(e^{-\lambda h}P\left((\tau-h)\wedge T\vee 0,\tau\right)f\left(\tau,\cdot\right)(x)-f\left(\tau,x\right)\right).$$

The operator $\vartheta_h : C_b\left([0,T]\times E\right)\to C_b\left([0,T]\times E\right)$ is defined by

$$\vartheta_h f\left(\tau,x\right)=f\left(\left((\tau-h)\wedge T\right)\vee 0,x\right),\quad f\in C_b\left([0,T]\times E\right). \tag{4.77}$$

Since

$$L^{(1)}_{\lambda,h}R(\lambda)f=R(\lambda)L^{(1)}_{\lambda,h}f=\frac{1}{h}\int_0^h e^{-\lambda\rho}S(\rho)fd\rho,\quad f\in C_b\left([0,T]\times E\right), \tag{4.78}$$

and $L^{(1)}$ is the $\mathcal{T}_\beta$-generator of the semigroup $\{S(\rho) : \rho\geqslant 0\}$, the equalities in (4.74) follow from (4.78). Since

$$\vartheta_h L^{(1)}_{\lambda,h}R(\lambda)f\left(\tau,x\right)=L^{(1)}_{\lambda,h}R(\lambda)f\left((\tau-h)\wedge T\vee 0,x\right),$$

it also follows that

$$\mathcal{T}_\beta\text{-}\lim_{h\downarrow 0}\vartheta_h L^{(1)}_{\lambda,h}R(\lambda)f\left(\tau,x\right)=\left(\lambda I-L^{(1)}\right)R(\lambda)f\left(\tau,x\right). \tag{4.79}$$

A consequence of (4.79) and the second equality in (4.76) is that

$$\begin{aligned}
&\left(\lambda I-L^{(1)}\right)f(\tau,x)\\
&=\lim_{h\downarrow 0}\left(\frac{1}{h}\left(f\left((\tau-h)\wedge T\vee 0,x\right)-f\left(\tau,x\right)\right)\right. &(4.80)\\
&\qquad\left.-\frac{1}{h}\left(e^{-\lambda h}P\left((\tau-h)\wedge T\vee 0,\tau\right)f\left(\tau,\cdot\right)(x)-f\left(\tau,x\right)\right)\right)\\
&=\lim_{h\downarrow 0}\left(\frac{1}{h}\left(f\left(\tau,x\right)-e^{-\lambda h}P\left(\tau,(\tau+h)\wedge T\right)f\left(\tau,\cdot\right)(x)\right)\right.\\
&\qquad\left.-\frac{1}{h}\left(\left(f\left((\tau+h)\wedge T,\cdot\right)-f\left(\tau,\cdot\right)\right)(x)\right)\right). &(4.81)
\end{aligned}$$

These limits exist in the strict sense; i.e. in the $\mathcal{T}_\beta$-topology. If $f\in D\left(L^{(1)}\right)$, and if f belongs to $D\left(D_1\right)$, then (4.80) and (4.81) imply that $f\in D\left(\overline{L}\right)$, that

$$\begin{aligned}
\overline{L}(f)(\tau,x)&=\lim_{h\downarrow 0}\frac{1}{h}\left(P\left(\tau,\tau+h\right)f\left(\tau,\cdot\right)(x)-f(\tau,x)\right)\\
&=\lim_{h\downarrow 0}\frac{1}{h}\left(P\left(\tau-h,\tau\right)f\left(\tau,\cdot\right)(x)-f(\tau,x)\right), &(4.82)
\end{aligned}$$

and that

$$L^{(1)}f=\overline{L}f+D_1 f. \tag{4.83}$$

Hence, in principle, the first term on the right-hand side in (4.80) converges to the negative of the time-derivative of the function f and the second to

$\left(\lambda I - \overline{L}\right) f$. The following arguments make this more precise. We will need the fact that the subspace $C_{P,b}^{(1)}$ is $\mathcal{T}_\beta$-dense in $C_b\left([0,T]\times E\right)$. Let $f \in C_b\left([0,T]\times E\right)$. In order to prove that, under certain conditions, the operator $L^{(1)}$ is the closure of $D_1 + L$, we consider for $f \in D\left(L^{(1)}\right)$ and $0 \leqslant a \leqslant b \leqslant T$ the following equality:

$$\begin{aligned}\int_a^b \vartheta_\rho S(\rho) L^{(1)} f\left(\tau, x\right) d\rho &= \int_a^b S(\rho) L^{(1)} f\left((\tau-\rho)\vee 0, x\right) d\rho \\ &= \frac{\partial}{\partial \tau}\int_a^b \vartheta_\rho S\left(\rho\right) f(\tau, x) d\rho \\ &\quad + P\left((\tau-b)\vee 0, \tau\right) f(\tau,x) - P\left((\tau-a)\vee 0,\tau\right) f(\tau,x). \end{aligned} \tag{4.84}$$

We first prove the equality on (4.84). Therefore we write

$$\begin{aligned}&\frac{\partial}{\partial \tau}\int_a^b \vartheta_\rho S\left(\rho\right) f(\tau, x) d\rho + P\left((\tau-b)\vee 0, \tau\right) f(\tau,x) - P\left((\tau-a)\vee 0,\tau\right) f(\tau,x) \\ &= \frac{\partial}{\partial \tau}\int_{\tau-b}^{\tau-a} S\left(\tau-\rho\right) f(\rho\vee 0, x) d\rho \\ &\quad + P\left((\tau-b)\vee 0, \tau\right) f(\tau,x) - P\left((\tau-a)\vee 0,\tau\right) f(\tau,x)\end{aligned}$$

(the function f belongs to $D\left(L^{(1)}\right)$)

$$\begin{aligned}&= \int_{\tau-b}^{\tau-a} S\left(\tau-\rho\right) L^{(1)} f(\rho\vee 0, x) d\rho \\ &\quad + S\left(a\right) f\left((\tau-a)\vee 0, x\right) - S\left(b\right) f\left((\tau-b)\vee 0, x\right) \\ &\quad + P\left((\tau-b)\vee 0, \tau\right) f(\tau,x) - P\left((\tau-a)\vee 0,\tau\right) f(\tau,x) \\ &= \int_a^b S\left(\rho\right) L^{(1)} f((\tau-\rho)\vee 0, x) d\rho = \int_a^b \vartheta_\rho S\left(\rho\right) L^{(1)} f(\tau, x) d\rho. \end{aligned} \tag{4.85}$$

The equality in (4.85) shows (4.84). In the same manner the following equality can be proved for $\lambda > 0$ and $f \in D\left(L^{(1)}\right)$:

$$\begin{aligned}\lambda \int_0^\infty e^{-\lambda\rho} \vartheta_\rho S(\rho) L^{(1)} f d\rho = \lambda D_1 &\int_0^\infty e^{-\lambda\rho}\vartheta_\rho S(\rho) f d\rho \\ &+ \lambda^2 \int_0^\infty e^{-\lambda\rho}\vartheta_\rho S(\rho) f d\rho - \lambda f. \end{aligned} \tag{4.86}$$

As above let $f \in D\left(L^{(1)}\right)$. From (4.86) we infer that

$$L^{(1)} f = \mathcal{T}_\beta\text{-}\lim_{\lambda\to\infty}\left(\lambda D_1 \int_0^\infty e^{-\lambda\rho}\vartheta_\rho S(\rho) f d\rho + \lambda^2 \int_0^\infty e^{-\lambda\rho}\vartheta_\rho S(\rho) f d\rho - \lambda f\right). \tag{4.87}$$

If, in addition, f belongs to the domain of D_1, then it also belongs to $D\left(\overline{L}\right)$, and

$$\begin{aligned}\overline{L}f &= \mathcal{T}_\beta\text{-}\lim_{\lambda\to\infty}\left(\lambda^2\int_0^\infty e^{-\lambda\rho}\vartheta_\rho S(\rho)f\,d\rho-\lambda f\right)\\ &= \mathcal{T}_\beta\text{-}\lim_{\lambda\to\infty}\left(\lambda^2\int_0^\infty e^{-\lambda\rho}S(\rho)\vartheta_\rho f\,d\rho-\lambda f\right). \qquad (4.88)\end{aligned}$$

The second equality in (4.88) follows from (4.87). So far the result is not conclusive. To finish the proof of the implication (i) $\Longrightarrow$ (ii) of Theorem 4.3 we will use the hypothesis that the space $C_{P,b}^{(1)}(\lambda_0)$ is $\mathcal{T}_\beta$-dense for some $\lambda_0>0$. In addition, we will use the following identity for a function f in the domain of the time derivative D_1:

$$\begin{aligned}&\lambda L^{(1)}\int_0^\infty e^{-\lambda\rho}S(\rho)\vartheta_\rho f\,d\rho\\ &=\lambda^2\int_0^\infty e^{-\lambda\rho}S(\rho)\vartheta_\rho f\,d\rho-\lambda f+\lambda\left(\lambda I-L^{(1)}\right)\int_0^\infty e^{-\lambda\rho}S(\rho)\left(I-\vartheta_\rho\right)f\,d\rho\\ &=\lambda^2\int_0^\infty e^{-\lambda\rho}S(\rho)\vartheta_\rho f\,d\rho-\lambda f+\lambda\int_0^\infty e^{-\lambda\rho}S(\rho)\vartheta_\rho D_1 f\,d\rho. \qquad (4.89)\end{aligned}$$

However, this is not the best approach either. The following arguments will show that the $\mathcal{T}_\beta$-density of $C_{P,b}^{(1)}(\lambda_0)$ is dense in $C_0([0,T]\times E)$ entails that $D^{(1)}(L)=D(L)\bigcap D(D_1)$ is a core for the operator $L^{(1)}$. From (4.83) it follows that $D^{(1)}(L)\subset D\left(L^{(1)}\right)$. From (4.87), (4.88), and from (4.89) we also get $D\left(L^{(1)}\right)\bigcap D(D_1)=D\left(\overline{L}\right)\bigcap D(D_1)$. Fix $\lambda_0>0$ such that the space $C_{P,b}^{(1)}(\lambda_0)$ is $\mathcal{T}_\beta$-dense in $C_b([0,T]\times E)$. Since $R\left(\lambda_0 I-L^{(1)}\right)=C_{P,b}^{(1)}(\lambda_0)$, this hypothesis has as a consequence that the range of the operator $\lambda_0 I-\overline{L}-D_1$ is $\mathcal{T}_\beta$-dense in $C_b([0,T]\times E)$. The $\mathcal{T}_\beta$-dissipativity of the operator $L^{(1)}$ then implies that the subspace $D\left(\overline{L}\right)\bigcap D(D_1)$ is a core for the operator $L^{(1)}$, and consequently, the closure of the operator $L+D_1$ coincides with $L^{(1)}$. We will show all this. Since the operator $L^{(1)}$ generates a Feller semigroup, the same is true for the closure of $L+D_1$. The range of $\lambda_0 I-\overline{L}-D_1$ coincides with the subspace $C_{P,b}^{(1)}(\lambda_0)$ defined in (4.36). It is easy to see that

$$C_{P,b}^{(1)}(\lambda_0)=\left\{f\in C_b([0,T]\times E):\,R(\lambda_0)f=\int_0^\infty e^{-\lambda_0\rho}S(\rho)f\,d\rho\in D(D_1)\right\}. \qquad (4.90)$$

If $f\in C_{P,b}^{(1)}(\lambda_0)$, then $f=\left(\lambda_0 I-L^{(1)}\right)R(\lambda_0)f$ where

$$R(\lambda_0)f\in D\left(L^{(1)}\right)\bigcap D(D_1)=D\left(\overline{L}\right)\bigcap D(D_1), \qquad (4.91)$$

as was shown in (4.87) and (4.88). It follows that $f \in C_{P,b}^{(1)}(\lambda_0)$ can be written as

$$f = \left(\lambda_0 I - \overline{L} - D_1\right) R(\lambda_0) f. \tag{4.92}$$

By (i) the range of $\lambda_0 I - \overline{L} - D_1$ is $\mathcal{T}_\beta$-dense in $C_b([0,T] \times E)$. Let f belong to the $\mathcal{T}_\beta$-closure of he range of $\lambda_0 I - \overline{L} - D_1$. Then there exists a net $(g_\alpha)_{\alpha\in\mathcal{A}} \subset C_{P,b}^{(1)}(\lambda_0) \subset C_b([0,T] \times E)$ such that $f = \lim_\alpha \left(\lambda_0 I - \overline{L} - D_1\right) g_\alpha$. From (4.14) we infer that $g = \mathcal{T}_\beta\text{-}\lim_\alpha g_\alpha$. Since the operator $\mathcal{T}_\beta$-closed linear operator $L^{(1)}$ extends $\overline{L} + D_1$, it follows that $\overline{L} + D_1$ is $\mathcal{T}_\beta$-closable. Let L_0 be its $\mathcal{T}_\beta$-closure. From (4.14) it also follows that $f = (\lambda_0 I - L_0) g$. Since the range of $\lambda_0 I - \overline{L} - D_1$ is $\mathcal{T}_\beta$-dense, we see that $R(\lambda_0 I - L_0) = C_b([0,T] \times E)$. Next let $g \in D\left(L^{(1)}\right)$. Then there exists $g_0 \in D(L_0)$ such that $\left(\lambda_0 I - L^{(1)}\right) g = (\lambda_0 I - L_0) g_0$. Since $L^{(1)}$ extends L_0, and since $L^{(1)}$ is dissipative (see (4.53), it follows that $g = g_0 \in D(L_0)$. In other words, the operator L_0 coincides with $L^{(1)}$, and consequently, the operator $L + D_1$ is $\mathcal{T}_\beta$-closable, and its closure coincides with $L^{(1)}$, the $\mathcal{T}_\beta$-generator of the semigroup $\{S(\rho) : \rho \geqslant 0\}$. This proves the implication (i) $\Longrightarrow$ (ii) of Theorem 4.3.

(ii) $\Longrightarrow$ (iii). Let $L^{(2)}$ be the $\mathcal{T}_\beta$-closure of the operator $D_1 + L$. From (ii) we know that $L^{(2)}$ generates a $\mathcal{T}_\beta$-continuous semigroup $\{S_2(\rho) : \rho \geqslant 0\}$. Since $D\left(L^{(2)}\right)$ is $\mathcal{T}_\beta$-dense, it follows that $D^{(1)}(L) = D(D_1) \bigcap D(L)$ is $\mathcal{T}_\beta$-dense as well. Let $L^{(1)}$ be the generator of a $\mathcal{T}_\beta$-continuous semigroup $\{S(\rho) : \rho \geqslant 0\}$ which extends $D_1 + L$, and hence it also extends $L^{(2)}$. Since $L^{(2)}$ generates a Feller semigroup, it is dissipative, and so it satisfies (4.53). Let $g \in D\left(L^{(1)}\right)$, and choose $g_0 \in D\left(L^{(2)}\right)$ such that

$$\left(\lambda_0 I - L^{(1)}\right) g = \left(\lambda_0 I - L^{(1)}\right) g_0 = \left(\lambda_0 I - L^{(2)}\right) g_0.$$

The inequality in (4.53) implies that $g = g_0 \in D\left(L^{(2)}\right)$, and hence $D\left(L^{(2)}\right) = D\left(L^{(1)}\right)$. Moreover, $L^{(1)}$ extends $L^{(2)}$. Therefore $L^{(2)} = L^{(1)}$. It also follows that the semigroup $\{S_2(\rho) : \rho \geqslant 0\}$ is the same as $\{S(\rho) : \rho \geqslant 0\}$. In addition, there exists $\lambda_0 > 0$ such that the range of $\lambda_0 I - D_1 - L$ is $\mathcal{T}_\beta$-dense in $C_b([0,T] \times E)$. In fact this is true for all λ, $\Re\lambda > 0$. Finally, we will show that the operator $D_1 + L$ is positive $\mathcal{T}_\beta$-dissipative. Let $u \in H^+([0,T] \times E)$, and consider the functionals $f \mapsto u(\tau,x)\lambda R(\lambda) f(\tau,x)$, $\lambda \geqslant \lambda_0 > 0$, $(\tau,x) \in [0,T] \times E$. Since $L^{(1)}$ generates a $\mathcal{T}_\beta$-continuous semigroup we know that

$$\lim_{\lambda\to\infty} \left\| u\,(f - \lambda R(\lambda) f)\right\|_\infty = 0. \tag{4.93}$$

If $(f_m)_{m\in\mathbb{N}} \subset C_b([0,T]\times E)$ decreases pointwise to 0, then the sequence $(u(\tau,x)\lambda R(\lambda)f_m(\tau,x))_{m\in\mathbb{N}}$ also decreases to 0. By Dini's Lemma and (4.93) this convergence is uniform in $\lambda \geqslant \lambda_0$ and $(\tau,x)\in[0,T]\times E$, because $u \in H^+([0,T]\times E)$. From Theorem 2.3 it follows that there exists a function $v\in H^+([0,T]\times E)$ such that

$$\|u\lambda R(\lambda)f\|_\infty \leqslant \|vf\|_\infty, \quad f\in C_b([0,T]\times E), \quad \lambda\geqslant\lambda_0. \tag{4.94}$$

Since the operator $L^{(1)}$ sends real functions to real functions from (4.94), and $u\geqslant 0$, we derive for $(\sigma,y)\in[0,T]\times E$

$$\begin{aligned}\Re\left(u(\sigma,y)\lambda R(\lambda)f(\sigma,y)\right) &= u(\sigma,y)\lambda R(\lambda)\left(\Re f\right)(\sigma,y)\\ &\leqslant u(\sigma,y)\lambda R(\lambda)\left(\Re f\right)^+(\sigma,y)\\ &\leqslant \sup_{(\tau,x)\in[0,T]\times E} v(\tau,x)\left(\Re f\right)^+(\tau,x)\\ &\leqslant \sup_{(\tau,x)\in[0,T]\times E} v(\tau,x)\left(\Re f\right)(\tau,x).\end{aligned} \tag{4.95}$$

By the substitution $f=\left(\lambda I - L^{(1)}\right)g$ in (4.95) we obtain:

$$\begin{aligned}&\lambda \sup_{(\tau,x)\in[0,T]\times E} u(\tau,x)\Re g(\tau,x)\\ &\leqslant \sup_{(\tau,x)\in[0,T]\times E} v(\tau,x)\Re\left(\lambda g(\tau,x) - L^{(1)}g(\tau,x)\right).\end{aligned} \tag{4.96}$$

Since the operator $L^{(1)}$ extends D_1+L, the inequality in (4.96) displays the fact that the operator D_1+L is positive $\mathcal{T}_\beta$-dissipative.

Altogether, this shows the implication (ii) $\Longrightarrow$ (iii) of Theorem 4.3.

(iii) $\Longrightarrow$ (i). Suppose that we already know that the $\mathcal{T}_\beta$-closure of D_1+L generates a $\mathcal{T}_\beta$-continuous semigroup $\{S(\rho): \rho\geqslant 0\}$. Then we define the evolution $\{P(\tau,t): 0\leqslant\tau\leqslant t\leqslant T\}$ by

$$P(\tau,t)f(x) = S(t-\tau)\left[(s,y)\mapsto f(y)\right](\tau,x), \quad f\in C_b(E). \tag{4.97}$$

We have to prove that the family $\{P(\tau,t): 0\leqslant\tau\leqslant t\leqslant T\}$ is a Feller evolution indeed. First we show that it has the evolution property:

$$\begin{aligned}P(\tau,t_1)P(t_1,t)f(x) &= S(t_1-\tau)\left[(s,y)\mapsto P(t_1,t)f(y)\right](\tau,x)\\ &= S(t_1-\tau)\left[(s,y)\mapsto S(t-t_1)f(s,y)\right](\tau,x)\\ &= S(t_1-\tau)S(t-t_1)\left[(s,y)\mapsto f(s,y)\right](\tau,x)\\ &= S(t-\tau)\left[(s,y)\mapsto f(s,y)\right](\tau,x)\\ &= P(\tau,t)f(x).\end{aligned} \tag{4.98}$$

The equality in (4.98) exhibits the evolution property. The continuity of the function $(\tau, t, x) \mapsto P(\tau, t) f(x)$ follows from the continuity of the function $(\tau, t, x) \mapsto S(t-\tau)[(s, y) \mapsto f(y)](\tau, x)$: see (4.97).

Next we prove that the operator $D_1 + L$ is $\mathcal{T}_\beta$-closable, and that its closure generates a Feller semigroup. Since the operator $D_1 + L$ is $\mathcal{T}_\beta$-densely defined and $\mathcal{T}_\beta$-dissipative, it is $\mathcal{T}_\beta$-closable: see Proposition 4.3 assertion (a). Let $L^{(1)}$ be its $\mathcal{T}_\beta$-closure. Since there exists $\lambda_0 > 0$ such that the range of $\lambda_0 I - D_1 - L$ is $\mathcal{T}_\beta$-dense in $C_b([0, T] \times E)$, and since $D_1 + L$ is $\mathcal{T}_\beta$-dissipative, it follows that $R\left(\lambda_0 I - L^{(1)}\right) = C_b([0, T] \times E)$. Put $R(\lambda_0) = \left(\lambda_0 I - L^{(1)}\right)^{-1}$, and $R(\lambda) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n (R(\lambda_0))^{n+1}$, $|\lambda - \lambda_0| < \lambda_0$. This series converges in the uniform norm. It follows that $R\left(\lambda I - L^{(1)}\right) = C_b([0, T] \times E)$ for all $\lambda \in \mathbb{C}$ for which $|\lambda - \lambda_0| < \lambda_0$. This procedure can be repeated to obtain: $R\left(\lambda I - L^{(1)}\right) = C_b([0, T] \times E)$ for all $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$. Put

$$S_0(t) f = \mathcal{T}_\beta\text{-}\lim_{\lambda \to \infty} e^{-\lambda t} e^{t\lambda^2 R(\lambda)} f, \quad f \in C_b([0, T] \times E). \tag{4.99}$$

Of course we have to prove that the limit in (4.99) exists. For brevity we write $A(\lambda) = \lambda^2 R(\lambda) - \lambda I = L^{(1)}(\lambda R(\lambda))$, and notice that for $f \in D\left(L^{(1)}\right)$ we have $A(\lambda) f = \lambda R(\lambda) L^{(1)} f$, and that

$$A(\lambda) f = \lambda R(\lambda) L^{(1)} f = R(\lambda)\left(L^{(1)}\right)^2 f + L^{(1)} f, \quad \text{for } f \in D\left(\left(L^{(1)}\right)^2\right). \tag{4.100}$$

Let $0 < \lambda < \mu < \infty$. From Duhamel's formula we get

$$\begin{aligned} & e^{-\lambda t} e^{\lambda t(\lambda R(\lambda))} f - e^{-\mu t} e^{\mu t(\mu R(\mu))} f \\ & = e^{tA(\lambda)} f - e^{tA(\mu)} f = \int_0^t e^{sA(\lambda)} (A(\lambda) - A(\mu)) e^{(t-s)A(\mu)} f ds. \end{aligned} \tag{4.101}$$

If f belongs to $D\left((L^{(1)})^2\right)$, then $A(\lambda) f - A(\mu) f = (R(\lambda) - R(\mu))\left(L^{(1)}\right)^2 f$, and hence the equality in (4.101) can be rewritten as:

$$\begin{aligned} & e^{-\lambda t} e^{\lambda t(\lambda R(\lambda))} f - e^{-\mu t} e^{\mu t(\mu R(\mu))} f \\ & = \int_0^t e^{sA(\lambda)} (R(\lambda) - R(\mu)) e^{(t-s)A(\mu)} \left(L^{(1)}\right)^2 f ds. \end{aligned} \tag{4.102}$$

From (4.102) we infer that for the uniform norm we have:

$$\begin{aligned} & \left\| e^{-\lambda t} e^{\lambda t(\lambda R(\lambda))} f - e^{-\mu t} e^{\mu t(\mu R(\mu))} f \right\|_\infty \\ & \leqslant \int_0^t \left\| e^{sA(\lambda)} (R(\lambda) - R(\mu)) e^{(t-s)A(\mu)} \left(L^{(1)}\right)^2 f \right\|_\infty ds \end{aligned}$$

$$\leqslant \int_0^t \left\| e^{sA(\lambda)} \right\| \left\| R(\lambda) - R(\mu) \right\| \left\| e^{(t-s)A(\mu)} \right\| \left\| \left(L^{(1)} \right)^2 f \right\|_\infty ds$$
$$\leqslant t \left(\frac{1}{\lambda} + \frac{1}{\mu} \right) \left\| \left(L^{(1)} \right)^2 f \right\|_\infty . \tag{4.103}$$

From (4.103) we infer that for $f \in D\left(\left(L^{(1)}\right)^2\right)$ the limit

$$S_0(t)f(\tau, x) = \lim_{\lambda\to\infty} e^{tA(\lambda)} f(\tau, x) \tag{4.104}$$

exists uniformly in $(\tau, t, x) \in [0, T] \times [0, T] \times E$. The next step consists in showing that the limit in (4.104) exists for $f \in D\left(L^{(1)}\right)$. Let $f \in D\left(L^{(1)}\right)$, and $\lambda > \mu > 0$. Then we have for $\lambda_0 > 0$ sufficiently large:

$$\begin{aligned}
&\left\| e^{tA(\lambda)} f - e^{tA(\mu)} f \right\|_\infty \\
&\leqslant \left\| e^{tA(\lambda)} \left(f - \lambda_0 R\left(\lambda_0\right) f \right) - e^{tA(\mu)} \left(f - \lambda_0 R\left(\lambda_0\right) f \right) \right\|_\infty \\
&\quad + \left\| e^{tA(\lambda)} \left(\lambda_0 R\left(\lambda_0\right) f \right) - e^{tA(\mu)} \left(\lambda_0 R\left(\lambda_0\right) f \right) \right\|_\infty \\
&\leqslant \left(\left\| e^{tA(\lambda)} \right\| + \left\| e^{tA(\mu)} \right\| \right) \left\| R\left(\lambda_0\right) L^{(1)} f \right\|_\infty \\
&\quad + \left\| e^{tA(\lambda)} \left(\lambda_0 R\left(\lambda_0\right) f \right) - e^{tA(\mu)} \left(\lambda_0 R\left(\lambda_0\right) f \right) \right\|_\infty \\
&\leqslant \frac{2}{\lambda_0} \left\| L^{(1)} f \right\|_\infty + \left\| e^{tA(\lambda)} \left(\lambda_0 R\left(\lambda_0\right) f \right) - e^{tA(\mu)} \left(\lambda_0 R\left(\lambda_0\right) f \right) \right\|_\infty .
\end{aligned} \tag{4.105}$$

From (4.105) together with (4.104) it follows that (4.104) also holds for $f \in D\left(L^{(1)}\right)$. There remains to be shown that the limit in (4.104) also exists in $\mathcal{T}_\beta$-sense, but now for $f \in C_b([0, T] \times E)$. Since the operator $L^{(1)}$ is $\mathcal{T}_\beta$-dissipative, there exists, for $u \in H^+([0, T] \times E)$, a function $v \in H^+([0, T] \times E)$ such that for all $\lambda \geqslant \lambda_0 > 0$ the inequality in (4.14) in Definition 4.2 is satisfied, i.e.

$$\left\| v\left(\lambda f - L^{(1)} f \right) \right\|_\infty \geqslant \lambda \left\| u f \right\|_\infty, \text{ for all } \lambda \geqslant \lambda_0, \text{ and for all } f \in D\left(L^{(1)} \right). \tag{4.106}$$

From (4.106) we infer

$$\lambda \left\| u R(\lambda) f \right\|_\infty \leqslant \left\| v f \right\|_\infty, \quad f \in C_b([0, T] \times E). \tag{4.107}$$

Let $f \in C_b([0, T] \times E)$. By Hausdorff-Bernstein-Widder inversion theorem there exists a unique Borel-measurable function $(\tau, t, x) \mapsto \widetilde{S}_0(t) f(\tau, x)$ such that

$$R(\lambda) f(\tau, x) = \left(\lambda I - L^{(1)} \right)^{-1} f(\tau, x) = \int_0^\infty e^{-\lambda\rho} \widetilde{S}_0(\rho) f(\tau, x) d\rho, \quad \Re\lambda > 0. \tag{4.108}$$

For the result in (4.108) see [Widder (1946)] Theorem 16a, page 315. The resolvent property of the mapping $\lambda \mapsto R(\lambda)$, $\lambda > 0$, implies the semigroup property of the mapping $\rho \mapsto S(\rho)$. To be precise we have:

$$\begin{aligned} R(\lambda)f - R(\mu)f &= \int_0^\infty \left(e^{-\lambda\rho} - e^{-\mu\rho}\right) \widetilde{S}_0(\rho) f d\rho \\ &= (\mu - \lambda) \int_0^\infty \int_0^\rho e^{-\lambda(\rho-s)-\mu s} \widetilde{S}_0 \left(\rho - s + s\right) f \, ds \, d\rho \\ &= (\mu - \lambda) \int_0^\infty \int_s^\infty e^{-\lambda(\rho-s)-\mu s} \widetilde{S}_0 \left(\rho - s + s\right) f \, d\rho \, ds \\ &= (\mu - \lambda) \int_0^\infty \int_0^\infty e^{-\lambda\rho-\mu s} \widetilde{S}_0 \left(\rho + s\right) f \, d\rho \, ds. \end{aligned} \tag{4.109}$$

On the other hand we also have

$$\begin{aligned} R(\lambda)f - R(\mu)f &= (\mu - \lambda) \, R(\lambda) R(\mu) f \\ &= (\mu - \lambda) \int_0^\infty \int_0^\infty e^{-\lambda\rho-\mu s} \widetilde{S}_0 \left(\rho\right) \widetilde{S}_0(s) f d\rho \, ds. \end{aligned} \tag{4.110}$$

Comparing (4.109) and (4.110) shows the equality:

$$\widetilde{S}_0 \left(\rho + s\right) f = \widetilde{S}_0(\rho) \widetilde{S}_0(s) f, \quad \rho, \, s \geqslant 0, \; f \in C_b \left([0, T] \times E\right).$$

Hence the family $\left\{\widetilde{S}_0(\rho) : \rho \geqslant 0\right\}$ is a semigroup. We have to show that the function $(\tau, t, x) \mapsto \widetilde{S}_0 \left(t\right) f(\tau, x)$ is a bounded continuous function. This will be done in several steps. First we will prove the following representation for $\widetilde{S}_0(t) f$, $f \in C_b \left([0, T] \times E\right)$,

$$\widetilde{S}_0(t) f = \lim_{\lambda\to\infty} e^{-\lambda t} \sum_{k=0}^\infty \frac{(\lambda t)^k}{k!} \left(\lambda R(\lambda)\right)^k f = \lim_{\lambda\to\infty} e^{-\lambda t} e^{\lambda t(\lambda R(\lambda))} f = S_0(t) f, \tag{4.111}$$

provided that the limit in (4.111) exists, and where $S_0(t)$ is as in (4.100). Let $f \in D\left(L^{(1)}\right)$. Then the function $S_0(t) f$ is the uniform limit of functions of the form $(\tau, t, x) \mapsto e^{tA(\lambda)} f(\tau, x)$, and such functions are continuous in the variables (τ, t, x): see (4.105). Consequently, the function $S_0(t) f$ inherits this continuity property. Again let $f \in D\left(L^{(1)}\right)$. We will prove that $R(\mu) f = \int_0^\infty e^{-\mu t} S_0(t) f dt$, $\mu > 0$. Therefore we notice

$$\begin{aligned} \int_0^\infty e^{-\mu t} S_0(t) f dt &= \int_0^\infty e^{-\mu t} \lim_{\lambda\to\infty} e^{tA(\lambda)} f dt \\ &= \lim_{\lambda\to\infty} \int_0^\infty e^{-\mu t} e^{tA(\lambda)} f dt = \lim_{\lambda\to\infty} \left(\mu I - A(\lambda)\right)^{-1} f \end{aligned}$$

$$= \lim_{\lambda\to\infty} \left(\frac{\lambda}{\lambda+\mu} I - \frac{1}{\lambda+\mu} L^{(1)}\right) \left(\frac{\lambda\mu}{\lambda+\mu} I - L^{(1)}\right)^{-1} f$$
$$= \left(\mu I - L^{(1)}\right)^{-1} f. \tag{4.112}$$

From (4.108) and (4.112) we infer the equality

$$\widetilde{S}_0(t) f = S_0(t) f \quad \text{for} \quad f \in D\left(L^{(1)}\right). \tag{4.113}$$

After that we will prove that the averages of the semigroup $\{S_0(\rho) : \rho \geqslant 0\}$ is $\mathcal{T}_\beta$-equi-continuous. As a consequence, for $f \in C_b([0,T] \times E)$ the function

$$(\tau, t, x) \mapsto \frac{1}{t} \int_0^t e^{-\lambda\rho} S_0(\rho) f(\tau, x)\, d\rho,$$

is a bounded and continuous function, and the family of operators

$$\left\{ \frac{1}{t} \int_0^t e^{-\lambda\rho} S_0(\rho) d\rho : 0 \leqslant t \leqslant T \right\} \text{ is } \mathcal{T}_\beta\text{-equi-continuous.} \tag{4.114}$$

As above we write $A(\lambda) f = \lambda^2 R(\lambda) f - \lambda f$. Two very relevant equalities are:

$$R(\lambda) f = \left(\lambda I - L^{(1)}\right)^{-1} f = \int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho + e^{-\lambda t} S_0(t) \left(\lambda I - L^{(1)}\right)^{-1} f$$
$$= \int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) d\rho\, f + e^{-\lambda t} \widetilde{S}_0(t) \left(\lambda I - L^{(1)}\right)^{-1} f, \tag{4.115}$$

and

$$f = \left(\lambda I - L^{(1)}\right) \int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho + e^{-\lambda t} \widetilde{S}_0(t) f. \tag{4.116}$$

Here we wrote

$$\int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) d\rho\, f = \int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) f\, d\rho$$

to indicate that the operator $f \mapsto \int_0^t e^{-\lambda\rho} S_0(\rho) d\rho\, f$, $f \in C_b([0,T] \times E)$, is a mapping from $C_b([0,T] \times E)$ to itself, whereas it is not so clear what the target space is of the mappings $\widetilde{S}_0(\rho)$, $\rho > 0$. In order to show that the operators $\widetilde{S}_0(t)$, $t \geqslant 0$, are mappings from $C_b([0,T] \times E)$ into itself, we need the sequential λ-dominance of the operator $D_1 + L$ for some $\lambda > 0$. Moreover, it follows from this sequential λ-dominance that the semigroup $\left\{ e^{-\lambda t} \widetilde{S}_0(t) : t \geqslant 0 \right\}$ is $\mathcal{T}_\beta$-equi-continuous. Once we know all this, then the formula in (4.116) makes sense and is true.

For every measure ν on the Borel field of $[0,T] \times E$ the mapping $\rho \mapsto \int \widetilde{S}_0(\rho) f d\nu$ is a Borel measurable function on the semi-axis $[0,\infty)$. The formula in (4.115) is correct, and poses no problem provided $f \in C_b([0,T] \times E)$. In fact we have

$$\begin{aligned}
&\int_0^\infty e^{-\mu t} e^{-\lambda t} S_0(t) R(\lambda) f\, dt = R(\lambda+\mu) R(\lambda) f = \frac{1}{\mu}\left(R(\lambda) - R(\lambda+\mu)\right) f \\
&= \int_0^\infty \frac{1-e^{-\mu\rho}}{\mu} e^{-\lambda\rho} \widetilde{S}_0(\rho) f\, d\rho = \int_0^\infty \int_0^\rho e^{-\mu t} dt\, e^{-\lambda\rho} \widetilde{S}_0(\rho) f\, d\rho \\
&= \int_0^\infty e^{-\mu t} \int_t^\infty e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho\, dt, \qquad (4.117)
\end{aligned}$$

and hence

$$e^{-\lambda t} S_0(t) \left(\lambda I - L^{(1)}\right)^{-1} f = e^{-\lambda t} S_0(t) R(\lambda) f = \int_t^\infty e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho. \qquad (4.118)$$

From (4.118) we infer

$$\begin{aligned}
&\int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho + e^{-\lambda t} S_0(t) \left(\lambda I - L^{(1)}\right)^{-1} f \\
&= \int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho + \int_t^\infty e^{-\lambda\rho} \widetilde{S}_0(\rho) f\, d\rho \\
&= \int_0^\infty e^{-\lambda\rho} \widetilde{S}_0(\rho) f d\rho = \left(\lambda I - L^{(1)}\right)^{-1} f. \qquad (4.119)
\end{aligned}$$

The equality in (4.115) is the same as the one in (4.119). From the equality in (4.115) it follows that the function $(\tau, t, x) \mapsto \int_0^t e^{-\lambda\rho} \widetilde{S}_0(\rho) f(\tau, x) d\rho$ is continuous. Next let $g \in D\left(L^{(1)}\right)$ and put $f = \left(\lambda I - L^{(1)}\right) g$. From (4.115) we get:

$$g - e^{-\lambda t} S_0(t) g = \int_0^t e^{-\lambda\rho} S_0(\rho) d\rho\, f. \qquad (4.120)$$

From (4.120) we infer:

$$\begin{aligned}
&\left\|g - e^{-\lambda t} S_0(t) g\right\|_\infty = \left\|\int_0^t e^{-\lambda\rho} S_0(\rho) d\rho\, f\right\|_\infty \\
&\leqslant \int_0^t e^{-\lambda\rho} \left\|\widetilde{S}_0(\rho) f\right\|_\infty d\rho \leqslant \int_0^t e^{-\lambda\rho} d\rho \left\|\lambda g - L^{(1)} g\right\|_\infty, \qquad (4.121)
\end{aligned}$$

and hence for $\lambda = 0$ we obtain $\|g - S_0(t) g\|_\infty \leqslant t \left\|L^{(1)} g\right\|_\infty$. This inequality proves the uniform boundedness of the family $\left\{\frac{1}{t}\left(g - S_0(t) g\right) : t > 0\right\}$. Next let us discuss its convergence. Therefore we again employ (4.115), and proceed as follows. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $C_b([0,T]\times E)$ which decreases pointwise to the zero-function. Choose the sequence $\left(g_n^{\lambda}\right)_{n\in\mathbb{N}} \subset D\left(L^{(1)}\right)$ in such a way that $\lambda f_n = \lambda g_n^{\lambda} - L^{(1)} g_n^{\lambda}$. Then the sequence $\left(g_n^{\lambda}\right)_{n\in\mathbb{N}}$ decreases to zero as well. For $t > 0$ have

$$g_n^{\lambda} = e^{-\lambda t} S_0(t) g_n^{\lambda} + \lambda \int_0^t e^{-\lambda\rho} S_0(\rho) d\rho\, f_n, \tag{4.122}$$

or, what is equivalent,

$$e^{\lambda t} g_n^{\lambda} = S_0(t) g_n^{\lambda} + \lambda \int_0^t e^{\lambda t - \lambda\rho} S_0(\rho) d\rho\, f_n. \tag{4.123}$$

Since the operator $L^{(1)}$ is $\mathcal{T}_\beta$-dissipative, it follows that $\sup\limits_{\lambda,\,\lambda\geqslant T^{-1}} g_n^{\lambda}$ decreases pointwise to zero for all $T > 0$. So that, with $\lambda = t^{-1}$, the equality in (4.122) implies

$$\sup_{t,\,0<t\leqslant T} \frac{1}{t}\int_0^t S_0(\rho) d\rho\, f_n \downarrow 0, \quad \text{as } n \to \infty. \tag{4.124}$$

Consequently, for any fixed $\lambda \in \mathbb{R}$, the family of operators

$$\left\{\frac{1}{t}\int_0^t e^{-\lambda\rho} S_0(\rho) d\rho : T \geqslant t > 0\right\} \tag{4.125}$$

is $\mathcal{T}_\beta$-equi-continuous: see Corollary 2.3. Let $f \in C_b([0,T]\times E)$. We will show that

$$\mathcal{T}_\beta\text{-}\lim_{t\downarrow 0} \frac{1}{t}\int_0^t e^{-\lambda\rho} S_0(\rho) d\rho\, f = f. \tag{4.126}$$

It suffices to prove (4.126) for $\lambda > 0$. First assume that $f = R(\lambda) g$ belongs to the domain of $L^{(1)}$. Then we have

$$\begin{aligned}
\frac{1}{t}\int_0^t e^{-\lambda\rho} S_0(\rho) d\rho\, f &= \frac{1}{t}\int_0^t e^{-\lambda\rho} S_0(\rho) \int_0^\infty e^{-\lambda\sigma} S_0(\sigma)\, d\sigma\, d\rho g \\
&= \frac{1}{t}\int_0^t \int_0^\infty e^{-\lambda(\sigma+\rho)} S_0(\sigma+\rho)\, d\sigma\, g\, d\rho \\
&= \frac{1}{t}\int_0^t \int_\rho^\infty e^{-\lambda\sigma} S_0(\sigma)\, d\sigma\, g\, d\rho.
\end{aligned} \tag{4.127}$$

Since the function $\rho \mapsto \int_\rho^\infty e^{-\lambda\sigma} S_0(\sigma)\, d\sigma\, g$ is continuous for the uniform norm topology on $C_b([0,T]\times E)$, (4.127) implies

$$\|\cdot\|_\infty\text{-}\lim_{t\downarrow 0} \frac{1}{t}\int_0^t \int_\rho^\infty e^{-\lambda\sigma} S_0(\sigma) d\sigma\, f = \int_0^\infty e^{-\lambda\sigma} S_0(\sigma) d\sigma\, g = f,\ f \in D\left(L^{(1)}\right). \tag{4.128}$$

Since $D\left(L^{(1)}\right)$ is $\mathcal{T}_\beta$-dense in $C_b\left([0,T]\times E\right)$, the equi-continuity of the family in (4.125) implies that

$$\mathcal{T}_\beta\text{-}\lim_{t\downarrow 0}\frac{1}{t}\int_0^t e^{-\lambda\rho}S_0(\rho)d\rho\, f = f, \quad f\in C_b\left([0,T]\times E\right). \tag{4.129}$$

From the equality in (4.120) together with (4.129) we see that

$$\mathcal{T}_\beta\text{-}\lim_{t\downarrow 0}\frac{g-e^{-\lambda t}S_0(t)g}{t} = f = \lambda g - L^{(1)}g, \quad g\in D\left(L^{(1)}\right). \tag{4.130}$$

So far we have proved that the semigroup $\left\{\widetilde{S}_0(t) : t\geqslant 0\right\}$ maps the domain of $L^{(1)}$ to bounded continuous functions, and that the family in (4.129) consists of mappings which assign to bounded continuous again bounded continuous functions. What is not clear, is whether or not the operators $\widetilde{S}_0(t)$, $t\geqslant 0$, leave the space $C_b\left([0,T]\times E\right)$ invariant. Fix $\lambda>0$, and to every $f\in C_b\left([0,T]\times E\right)$, $f\geqslant 0$, we assign the function f^λ defined by

$$f^\lambda = \sup\left\{\left(\mu R\left(\lambda+\mu\right)\right)^k f : \mu>0,\ k\in\mathbb{N}\right\}, \tag{4.131}$$

The reader is invited to compare the function f^λ with (2.49) and other results in Proposition 2.4. The arguments which follow are in line with the proof of Proposition 2.4. The function f^λ is the smallest λ-super-median valued function which exceeds f. A closely related notion is the notion of λ-super-mean valued function. A function $g:[0,T]\times E\to[0,\infty)$ is called λ-super-median valued if $e^{-\lambda t}\widetilde{S}_0(t)g\leqslant g$ for all $t\geqslant 0$; it is called λ-super-mean valued if $\mu R\left(\lambda+\mu\right)g\leqslant g$ for all $\mu>0$. In Lemma (9.12) in [Sharpe (1988)] it is shown that, essentially speaking, these notions are equivalent. In fact the proof is not very difficult. It uses the Hausdorff-Bernstein-Widder theorem about the representation by Laplace transforms of positive Borel measures on $[0,\infty)$ of completely positive functions. The reader is also referred to Remark 4.1 and Definition 4.3.

Let $f\in C_b\left([0,T]\times E\right)$ be positive. Here we use the representation

$$e^{-\lambda t}\widetilde{S}_0(t)f = \lim_{\mu\to\infty} e^{-\mu t}\sum_{k=0}^{\infty}\frac{(\mu t)^k}{k!}\left(\mu R\left(\lambda+\mu\right)\right)^k f \leqslant f^\lambda, \tag{4.132}$$

and hence

$$\sup_{t>0} e^{-\lambda t}\widetilde{S}_0(t)f\leqslant f^\lambda. \tag{4.133}$$

Since

$$\left(\mu R\left(\lambda+\mu\right)\right)^k f = \frac{\mu^k}{(k-1)!}\int_0^\infty t^{k-1}e^{-\mu t}e^{-\lambda t}\widetilde{S}_0(t)f\,dt \tag{4.134}$$

we see by invoking (4.131) that the two expressions in (4.133) are the same. In order to finish the proof of Theorem 4.3 we need the hypothesis that the operator $D_1 + L$ is sequentially λ-dominant for some $\lambda > 0$. In fact, let the sequence $(f_n)_{n\in\mathbb{N}} \subset C_b([0,T]\times E)$ converge downward to zero, and select functions $g_n^\lambda \in D(D_1+L)$, $n\in\mathbb{N}$, with the following properties:

(1) $f_n \leqslant g_n^\lambda$;
(2) $g_n^\lambda = \sup_{K\in\mathcal{K}([0,T]\times E)} \inf_{g\in D(D_1+L)} \{g \geqslant f_n \mathbf{1}_K : (\lambda I - D_1 - L)\, g \geqslant 0\}$;
(3) $\lim_{n\to\infty} g_n^\lambda(\tau,x) = 0$ for all $(\tau,x)\in[0,T]\times E$.

In the terminology of (2.42) and Definition 4.3 the functions g_n^λ are denoted by $g_n^\lambda = U_\lambda^1(f_n)$, $n\in\mathbb{N}$. Recall that $\mathcal{K}([0,T]\times E)$ denotes the collection of all compact subsets of $[0,T]\times E$. By hypothesis, the sequence as defined in 2 satisfies 1 and 3. Let K be any compact subset of $[0,T]\times E$, and $g\in D(D_1+L)$ be such that $g \geqslant f_n\mathbf{1}_K$ and $(\lambda I - D_1 - L)\, g \geqslant 0$. Then we have

$$\left((\lambda+\mu)\, I - L^{(1)}\right) g = ((\lambda+\mu)\, I - D_1 - L)\, g \geqslant \mu g. \tag{4.135}$$

From (4.135) and $g \geqslant f_n\mathbf{1}_K$ we infer

$$g \geqslant \mu R(\lambda+\mu)\, g \geqslant (\mu R(\lambda+\mu))^k\, g \geqslant (\mu R(\lambda+\mu))^k (f_n \mathbf{1}_K), \tag{4.136}$$

and hence (4.136) together with (4.131) and (4.133) (which is in fact an equality) we see

$$\begin{aligned} g_n^\lambda \geqslant f_n^\lambda &= \sup_{K\in\mathcal{K}([0,T]\times E)} \sup\left\{e^{-\lambda t}\widetilde{S}_0(t)\,(f_n\mathbf{1}_K) : t\geqslant 0\right\} \\ &= \sup_{K\in\mathcal{K}([0,T]\times E)} \sup\left\{(\mu R(\lambda+\mu))^k (f_n\mathbf{1}_K) : \mu>0,\, k\in\mathbb{N}\right\} \\ &= \sup\left\{(\mu R(\lambda+\mu))^k f_n : \mu>0,\, k\in\mathbb{N}\right\} \\ &= \sup\left\{e^{-\lambda t}\widetilde{S}_0(t) f_n : t\geqslant 0\right\}. \end{aligned} \tag{4.137}$$

Since by hypothesis $\lim_{n\to\infty} g_n^\lambda = 0$ the inequality in (4.137) implies: $\lim_{n\to\infty} f_n^\lambda = 0$. It follows that

$$\lim_{n\to\infty} \sup\left\{e^{-\lambda t}\widetilde{S}_0(t) f_n : t\geqslant 0\right\} = 0. \tag{4.138}$$

From Corollary 2.3 it follows that the family of operators

$$\left\{(\mu R(\lambda+\mu))^k : \mu\geqslant 0,\, k\in\mathbb{N}\right\}$$

is $\mathcal{T}_\beta$-equi-continuous. Hence for every function $u \in H^+([0,T] \times E)$ there exists a function $v \in H^+([0,T] \times E)$ such that

$$\left\| u\left(\mu R(\lambda+\mu)\right)^k f \right\|_\infty \leqslant \|vf\|_\infty, \quad f \in C_b([0,T] \times E), \ \mu \geqslant 0, \ k \in \mathbb{N}. \tag{4.139}$$

Since

$$\sup\left\{ e^{-\lambda t}\widetilde{S}_0(t)f : t \geqslant 0 \right\} = \sup\left\{ \left(\mu R(\lambda+\mu)\right)^k f : \mu \geqslant 0, k \in \mathbb{N} \right\}, \quad f \geqslant 0, \tag{4.140}$$

the inequality in (4.139) yields

$$\left\| u e^{-\lambda t}\widetilde{S}_0(t)f \right\|_\infty \leqslant \|vf\|_\infty, \quad f \in C_b([0,T] \times E), \ t \geqslant 0. \tag{4.141}$$

Since $D\left(L^{(1)}\right)$ is $\mathcal{T}_\beta$-dense, and the operators $\widetilde{S}_0(t)$, $t \geqslant 0$, are mappings from $D\left(L^{(1)}\right)$ to $C_b([0,T] \times E)$ the $\mathcal{T}_\beta$-equi-continuity in (4.141) shows that the operators $\widetilde{S}_0(t)$, $t \geqslant 0$, are in fact mappings from $C_b([0,T] \times E)$ to itself, and that the family $\left\{ e^{-\lambda t}\widetilde{S}_0(t) : t \geqslant 0 \right\}$ is $\mathcal{T}_\beta$-equi-continuous.

However, all these observations conclude the proof of the implication (iii) $\Longrightarrow$ (i) of Theorem 4.3.

So, finally, the proof of Theorem 4.3 is complete. □

Remark 4.3. The equality in (4.115) shows that the function $g := R(\lambda)f$, where $f \geqslant 0$ and $f \in C_b([0,T] \times E)$ is λ-super-mean valued in the sense that an inequality of the form $e^{-\lambda t}S_0(t)g \leqslant g$ holds. Such an inequality is equivalent to $\mu R(\mu+\lambda)\, g \leqslant g$. For details on such functions and on λ-excessive functions see [Sharpe (1988)], page 17 and Lemma 9.12, page 45.

4.3 Korovkin property

The following notions and results are being used to prove Theorem 2.13. We recall the definition of Korovkin property.

Definition 4.4. Let E_0 be a subset of E The operator L is said to possess the Korovkin property on E_0 if there exists a strictly positive real number $\lambda_0 > 0$ such that for every $x_0 \in E_0$ the equality

$$\inf_{h \in D(L)} \sup_{x \in E_0} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda_0} L \right) h \right](x) \right\} \tag{4.142}$$

$$= \sup_{h \in D(L)} \inf_{x \in E_0} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda_0} L \right) h \right](x) \right\} \tag{4.143}$$

is valid for all $g \in C_b(E)$.

Let $g \in C_b(E)$ and $\lambda > 0$. The equalities

$$\begin{aligned}
&\inf_{h\in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda} L \right) h \geqslant g \quad \text{on } E_0 \right\} \\
&= \inf_{h\in D(L)} \sup_{x\in E_0} \left(h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L \right) h \right](x) \right) \\
&= \inf_{\substack{\Gamma\subset D(L)\\ \#\Gamma<\infty}} \sup_{\substack{\Phi\subset E_0\\ \#\Phi<\infty}} \min_{h\in\Gamma} \max_{x\in\Phi} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L \right) h \right](x) \right\}, \qquad (4.144)
\end{aligned}$$

show that the Korovkin property could also have been defined in terms of any of the quantities in (4.144). In fact, if L satisfies the (global) maximum principle on E_0, i.e. if for every real-valued function $f \in D(L)$ the inequality

$$\lambda \sup_{x\in E_0} f(x) \leqslant \sup_{x\in E_0} \left(\lambda f(x) - Lf(x) \right) \qquad (4.145)$$

holds for all $\lambda > 0$, then the Korovkin property (on E_0) does not depend on $\lambda_0 > 0$. In other words, if it holds for one $\lambda_0 > 0$, then it is true for all $\lambda > 0$. This is part of the contents of the following proposition. In fact the maximum principle as formulated in (4.145) is not adequate in the present context. The correct version here is the following one, which is kind of a weak maximum principle on a subset of E.

Definition 4.5. Let E_0 be a subset of E. Suppose that the operator L has the property that for every $\lambda > 0$ and for every $x_0 \in E_0$ it is true that $h(x_0) \geqslant 0$, whenever $h \in D(L)$ is such that $(\lambda I - L)h \geqslant 0$ on E_0. Then the operator L is said to satisfy the weak maximum principle on E_0.

As we proved in Proposition 2.8 the notion of "weak maximum principle" and "maximum principle" coincide, provided $\mathbf{1} \in D(L)$ and $L\mathbf{1} = 0$.

In order to be really useful, the Korovkin property on E_0 should be accompanied by the maximum principle on E_0. To be useful the global Korovkin property (see Definition 4.6) requires the global maximum principle (see (4.145)). In addition we need the fact that the constant functions belong to $D(L)$ and that $L\mathbf{1} = 0$. If we only know the global maximum principle, in the sense of (4.145), then the global Korovkin property is required.

Definition 4.6. The operator L is said to possess the *global Korovkin property* if there exists a strictly positive real number $\lambda_0 > 0$ such that for every $x_0 \in E$ the equality

$$\inf_{h\in D(L)} \sup_{x\in E} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda_0} L \right) h \right](x) \right\} \tag{4.146}$$

$$= \sup_{h\in D(L)} \inf_{x\in E} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda_0} L \right) h \right](x) \right\} \tag{4.147}$$

is valid for all $g \in C_b(E)$.

First we treat the situation of a subset of E. The global version is obtained from the one on E_0 by replacing the subset E_0 with the full state space E. Again a resolvent family is obtained. In order to prove the equalities of (4.166) through (4.175) the global maximum principle is used. In fact it is used to show the equalities

$$\inf_{h \in D(L)} \sup_{x \in E} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L \right) h \right](x) \right\} \tag{4.148}$$

$$= \inf_{h\in D(L)} \left\{ h\left(x_0\right) : \left(I - \frac{1}{\lambda} L \right) h \geqslant g \quad \text{on } E \right\} \tag{4.149}$$

$$= \sup_{h\in D(L)} \left\{ h\left(x_0\right) : \left(I - \frac{1}{\lambda} L \right) h \leqslant g \quad \text{on } E \right\} \tag{4.150}$$

$$= \sup_{h\in D(L)} \inf_{x\in E} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L \right) h \right](x) \right\}. \tag{4.151}$$

In particular, if $g = 0$, and if L satisfies the global maximum principle, then the expressions in (4.148) through (4.151) are all equal to 0. Put

$$\begin{aligned} &\lambda_0 R\left(\lambda_0\right) g\left(x_0\right) \\ &= \inf_{h\in D(L)} \sup_{x\in E_0} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda_0} L \right) h \right](x) \right\} \\ &= \sup_{h\in D(L)} \inf_{x\in E_0} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda_0} L \right) h \right](x) \right\}. \end{aligned} \tag{4.152}$$

Then $\lambda_0 R\left(\lambda_0\right)$ is a linear operator from $C_b\left(E_0\right)$ to $C_b\left(E_0\right)$. The following proposition shows that there exists a family of operators $\{R(\lambda) : 0 < \lambda < 2\lambda_0\}$ which has the resolvent property. The operator $\lambda R(\lambda)$ is obtained from (4.152) by replacing λ_0 with λ. It is clear that this procedure can be extended to the whole positive real axis. In this way we obtain a resolvent family $\{R(\lambda) : \lambda > 0\}$. The operator $R(\lambda)$ can be written in the form $R(\lambda) = (\lambda I - L_0)^{-1}$, where L_0 is a closed linear operator which extends L (in case $E_0 = E$), and which satisfies the maximum principle on E_0, and, under certain conditions, generates a Feller semigroup and a Markov process. For convenience we insert the following lemma. It is used for $E_0 = E$ and for E_0 a subset of E which is Polish with respect to the relative metric. The condition in (4.155) is closely related to the maximum principle.

Lemma 4.1. *Suppose that the constant functions belong to $D(L)$, and that $L\mathbf{1} = 0$. Fix $x_0 \in E$, $\lambda > 0$, and $g \in C_b(E_0)$. Let E_0 be any subset of E. Then the following equalities hold:*

$$\begin{aligned}\inf_{h\in D(L)}\sup_{x\in E_0}&\left\{h(x_0)+g(x)-\left(I-\frac{1}{\lambda}L\right)h(x)\right\}\\ &=\inf_{h\in D(L)}\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\geqslant g\quad \text{on } E_0\right\},\end{aligned}\tag{4.153}$$

and

$$\begin{aligned}\sup_{h\in D(L)}\inf_{x\in E_0}&\left\{h(x_0)+g(x)-\left(I-\frac{1}{\lambda}L\right)h(x)\right\}\\ &=\sup_{h\in D(L)}\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\leqslant g\quad \text{on } E_0\right\}.\end{aligned}\tag{4.154}$$

If $\displaystyle\inf_{h\in D(L)}\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\geqslant 0\quad \text{on } E_0\right\}\geqslant 0,$ *then* (4.155)

$$\sup_{x\in E_0} g(x)\geqslant \inf_{h\in D(L)}\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\geqslant g\quad \text{on } E_0\right\}\geqslant \inf_{x\in E_0} g(x),\tag{4.156}$$

and also

$$\begin{aligned}\inf_{h\in D(L)}&\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\geqslant g\quad \text{on } E_0\right\}\\ &\geqslant\sup_{h\in D(L)}\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\leqslant g\quad \text{on } E_0\right\}.\end{aligned}\tag{4.157}$$

First notice that by taking $h = 0$ in the left-hand side of (4.153) we see that the quantity in (4.153) is less than or equal to $\sup_{x\in E_0} g(x)$, and that the quantity in (4.154) is greater than or equal to $\inf_{x\in E_0} g(x)$. However, it is not excluded that (4.153) is equal to $-\infty$, and that (4.154) is equal to ∞.

Proof. Upon replacing g with $-g$ we see that the equality in (4.154) is a consequence of (4.153). We put

$$\alpha_{E_0}=\inf_{h\in D(L)}\left\{h(x_0):\ \left(I-\frac{1}{\lambda}L\right)h\geqslant g\ \text{ on } E_0\right\}\quad\text{and}$$

$$\beta_{E_0} = \inf_{h \in D(L)} \sup_{x \in E_0} \left\{ h\left(x_0\right) + g(x) - \left(I - \frac{1}{\lambda} L\right) h(x) \right\}. \tag{4.158}$$

First assume that $\beta_{E_0} \in \mathbb{R}$. Let $\varepsilon > 0$. Choose $h_\varepsilon \in D(L)$ in such a way that for $x \in E_0$ we have

$$h_\varepsilon\left(x_0\right) + g(x) - \left(I - \frac{1}{\lambda} L\right) h_\varepsilon\left(x\right) \leqslant \beta_{E_0} + \varepsilon.$$

Then

$$\begin{aligned} g(x) &\leqslant \left(I - \frac{1}{\lambda} L\right) h_\varepsilon\left(x\right) + \beta_{E_0} + \varepsilon - h_\varepsilon\left(x_0\right) \\ &= \left(I - \frac{1}{\lambda} L\right) \left(h_\varepsilon - h_\varepsilon\left(x_0\right) + \beta_{E_0} + \varepsilon\right)(x). \end{aligned} \tag{4.159}$$

The substitution $\widetilde{h}_\varepsilon = h_\varepsilon - h_\varepsilon\left(x_0\right) + \beta_{E_0} + \varepsilon$ in (4.159) yields $\alpha_{E_0} \leqslant \widetilde{h}_\varepsilon\left(x_0\right) = \beta_{E_0} + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we get $\alpha_{E_0} \leqslant \beta_{E_0}$. The same argument with $-n$ instead of $\beta_{E_0} + \varepsilon$ shows $\alpha_{E_0} = -\infty$ if $\beta_{E_0} = -\infty$. Next we assume that $\alpha_{E_0} \in \mathbb{R}$. Again let $\varepsilon > 0$ be arbitrary. Choose a function $h_\varepsilon \in D(L)$ such that $h_\varepsilon\left(x_0\right) \leqslant \alpha_{E_0} + \varepsilon$, and $\left(I - \frac{1}{\lambda} L\right) h_\varepsilon \geqslant g$ on E_0. Then we have, for $x \in E_0$,

$$h_\varepsilon\left(x_0\right) + g(x) - \left(I - \frac{1}{\lambda} L\right) h_\varepsilon(x) \leqslant h_\varepsilon\left(x_0\right) \leqslant \beta_{E_0} + \varepsilon,$$

and hence $\beta_{E_0} \leqslant \alpha_{E_0} + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we get $\beta_{E_0} \leqslant \alpha_{E_0}$. Again, the argument can be adapted if $\alpha_{E_0} = -\infty$: replace $\alpha_{E_0} + \varepsilon$ by $-n$, and let n tend to ∞. If condition (4.155) is satisfied, then with $m = \inf_{y \in E_0} g(y)$ we have

$$\begin{aligned} \alpha_{E_0} &\geqslant \inf_{h \in D(L)} \left\{ h\left(x_0\right) : \ \left(I - \frac{1}{\lambda} L\right) h \geqslant \inf_{y \in E_0} g(y) \ \text{ on } E_0 \right\} \\ &= \inf_{h \in D(L)} \left\{ h\left(x_0\right) : \ \left(I - \frac{1}{\lambda} L\right) (h - m) \geqslant 0 \ \text{ on } E_0 \right\} \geqslant m. \end{aligned} \tag{4.160}$$

The inequality in (4.160) shows the lower estimate in (4.156). The upper estimate is obtained by taking $h = \sup_{y \in E_0} g(y)$. Next we prove the inequality in (4.157). Therefore we observe that the functional $\Lambda^+_{E_0} : C_b\left(E, \mathbb{R}\right) \to \mathbb{R}$, defined by

$$\Lambda^+_{E_0}(g) = \inf_{h \in D(L)} \left\{ h\left(x_0\right) : \ \left(I - \frac{1}{\lambda} L\right) h \geqslant g \ \text{ on } E_0 \right\} \tag{4.161}$$

is sub-additive and positive homogeneous. The latter means that

$$\Lambda^+_{E_0}(g_1+g_2) \leqslant \Lambda^+_{E_0}(g_1) + \Lambda^+_{E_0}(g_2), \quad \text{and} \quad \Lambda^+_{E_0}(\alpha g) = \alpha \Lambda^+_{E_0}(g)$$

for g_1, g_2, $g \in C_b(E,\mathbb{R})$, and $\alpha \geqslant 0$. Moreover,

$$-\Lambda^+_{E_0}(-g) = \sup_{h \in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda} L \right) h \leqslant g \ \text{ on } E_0 \right\}. \tag{4.162}$$

It follows that

$$\begin{aligned} &\Lambda^+_{E_0}(g) + \Lambda^+_{E_0}(-g) \geqslant \Lambda^+_{E_0}(0) \\ &= \inf_{h \in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda} L \right) h \geqslant 0 \ \text{ on } E_0 \right\} \geqslant 0. \end{aligned} \tag{4.163}$$

The inequality in (4.157) is a consequence of (4.162) and (4.163).

This completes the proof of Lemma 4.1. □

The definition of an operator L satisfying the maximum principle on a subset E_0 can be found in Definition 4.5. Proposition 4.4 contains the basic formulas which turn the Korovkin property into a resolvent family of operators, and ultimately a Feller semigroup.

Proposition 4.4. *Let $0 < \lambda < 2\lambda_0$ and $g \in C_b(E)$ and E_0 a subset of E. Suppose the operator L satisfies the maximum principle on E_0. In addition, let the domain of L contain the constant functions, and assume $L\mathbf{1} = 0$. Let $x_0 \in E_0$. Put*

$$\begin{aligned} &\lambda R(\lambda) g(x_0) \\ &= \liminf_{n\to\infty} \inf_{h_0 \in D(L)} \sup_{x_1 \in E_0} \inf_{h_1 \in D(L)} \sup_{x_2 \in E_0} \cdots \inf_{h_n \in D(L)} \sup_{x_{n+1} \in E_0} \\ &\quad \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \end{aligned} \tag{4.164}$$

$$\begin{aligned} &= \liminf_{n\to\infty} \sup_{h_0 \in D(L)} \inf_{x_1 \in E_0} \sup_{h_1 \in D(L)} \inf_{x_2 \in E_0} \cdots \\ &\quad \sup_{h_n \in D(L)} \inf_{x_{n+1} \in E_0} \\ &\quad \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\}. \end{aligned} \tag{4.165}$$

Then the following identities are true:

$$\lambda R(\lambda) g(x_0)$$

$$= \lim_{n\to\infty} \left[\frac{\lambda}{\lambda_0} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j (\lambda_0 R(\lambda_0))^{j+1} g(x_0) \right] \tag{4.166}$$

$$= \frac{\lambda}{\lambda_0} \sum_{j=0}^{\infty} \left(1 - \frac{\lambda}{\lambda_0}\right)^j (\lambda_0 R(\lambda_0))^{j+1} g(x_0) \tag{4.167}$$

$$= \lim_{n\to\infty} \inf_{\substack{h_j \in D(L),\ (\lambda I - L) h_0 = \frac{\lambda}{\lambda_0} \sum_{j=1}^{\infty} \left(1-\frac{\lambda}{\lambda_0}\right)^{j-1} (\lambda I - L) h_j \\ j\geqslant 0}} \max_{\substack{x_j \in E_0 \\ 1\leqslant j\leqslant n+1}} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \tag{4.168}$$

$$= \inf_{h\in D(L)} \max_{x\in E_0} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L\right) h \right](x) \right\} \tag{4.169}$$

$$= \inf_{h\in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda} L\right) h \geqslant g \quad on\ E_0 \right\} \tag{4.170}$$

$$= \sup_{h\in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda} L\right) h \leqslant g \quad on\ E_0 \right\} \tag{4.171}$$

$$= \sup_{h\in D(L)} \min_{x\in E_0} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L\right) h \right](x) \right\} \tag{4.172}$$

$$= \lim_{n\to\infty} \sup_{\substack{h_j \in D(L),\ (\lambda I - L) h_0 = \frac{\lambda}{\lambda_0} \sum_{j=1}^{\infty} \left(1-\frac{\lambda}{\lambda_0}\right)^{j-1} (\lambda I - L) h_j \\ j\geqslant 0}} \min_{\substack{x_j \in E_0 \\ 1\leqslant j\leqslant n+1}} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \tag{4.173}$$

$$= \lim_{n\to\infty} \inf_{h_j\in D(L),\, 0\leqslant j\leqslant n} \max_{x_j\in E_0,\, 1\leqslant j\leqslant n+1} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \tag{4.174}$$

$$= \lim_{n\to\infty} \sup_{h_j\in D(L),\, 0\leqslant j\leqslant n} \min_{x_j\in E_0,\, 1\leqslant j\leqslant n+1} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\}. \tag{4.175}$$

Suppose that the operator possesses the global Korovkin property, and satisfies the maximum principle, as described in (4.145). Put

$$\begin{aligned}
&\lambda R(\lambda)g(x_0) \\
&= \liminf_{n\to\infty} \inf_{h_0\in D(L)} \sup_{x_1\in E} \inf_{h_1\in D(L)} \sup_{x_2\in E} \cdots \inf_{h_n\in D(L)} \sup_{x_{n+1}\in E} \\
&\quad \sum_{j=0}^{n} \left(1-\frac{\lambda}{\lambda_0}\right)^j \left\{h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0}L\right) h_j(x_{j+1})\right\} \quad (4.176) \\
&= \liminf_{n\to\infty} \sup_{h_0\in D(L)} \inf_{x_1\in E} \sup_{h_1\in D(L)} \inf_{x_2\in E} \cdots \sup_{h_n\in D(L)} \inf_{x_{n+1}\in E} \\
&\quad \sum_{j=0}^{n} \left(1-\frac{\lambda}{\lambda_0}\right)^j \left\{h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0}L\right) h_j(x_{j+1})\right\}. \quad (4.177)
\end{aligned}$$

Then the quantities in (4.166) through (4.175) are all equal to $\lambda R(\lambda)g(x_0)$, provided that the set E_0 is replaced by E.

In case we deal with the (local) Korovkin property on E_0, the convergence of

$$(\lambda I - L) h_0 = \frac{\lambda}{\lambda_0} \sum_{j=1}^{\infty} \left(1-\frac{\lambda}{\lambda_0}\right)^{j-1} (\lambda I - L) h_j \quad (4.178)$$

in (4.168) and (4.173) is supposed to be uniform on E_0. In case we deal with the global Korovkin property, and the maximum principle in (4.145), then the convergence in (4.178) should be uniform on E.

Corollary 4.2. *Suppose that the operator L possesses the Korovkin property on E_0. Then for all $\lambda > 0$ the quantities in (4.169), (4.170), (4.171), and (4.172) are equal for all $x_0 \in E_0$ and all functions $g \in C_b(E_0)$. If L possesses the global Korovkin property, then*

$$\inf_{h \in D(L)} \max_{x \in E} \left\{h(x_0) + \left[g - \left(I - \frac{1}{\lambda}L\right) h\right](x)\right\} \quad (4.179)$$

$$= \inf_{h\in D(L)} \left\{h(x_0) : \left(I - \frac{1}{\lambda}L\right) h \geqslant g \quad \text{on } E\right\} \quad (4.180)$$

$$= \sup_{h\in D(L)} \left\{h(x_0) : \left(I - \frac{1}{\lambda}L\right) h \leqslant g \quad \text{on } E\right\} \quad (4.181)$$

$$= \sup_{h \in D(L)} \min_{x \in E} \left\{ h(x_0) + \left[g - \left(I - \frac{1}{\lambda} L \right) h \right] (x) \right\}. \tag{4.182}$$

Moreover, for $\lambda > 0$ and $f \in D(L)$, the equality $R(\lambda)\,(\lambda I - L)\,f = f$ holds.

Proof. By repeating the result in Proposition 4.4 for all $\lambda_1 \in (0, 2\lambda_0)$ instead of λ_0 we get these equalities for λ in the interval $(0, 4\lambda_0)$. This procedure can be repeated once more. Induction then yields the desired result. That for $\lambda > 0$ and $f \in D(L)$, the equality $R(\lambda)\,(\lambda I - L)\,f = f$ holds can be seen by the following arguments. By definition we have

$$\begin{aligned} &\lambda R(\lambda)\,(\lambda I - L)\,f\,(x_0) \\ &= \inf \left\{ h\,(x_0) : \, (\lambda I - L)\,f \geqslant h \text{ on } E_0, \; h \in D(L) \right\} \leqslant f\,(x_0)\,. \end{aligned} \tag{4.183}$$

We also have

$$\begin{aligned} &\lambda R(\lambda)\,(\lambda I - L)\,f\,(x_0) \\ &= \sup \left\{ h\,(x_0) : \, (\lambda I - L)\,f \leqslant h \text{ on } E_0, \; h \in D(L) \right\} \geqslant f\,(x_0)\,. \end{aligned} \tag{4.184}$$

The stated equality is a consequence of (3.161) and (4.183). It also completes the proof of Corollary 4.2. □

We continue with a proof of Proposition 4.4.

Proof. [Proof of Proposition 4.4.] The equality of each term in (4.164) and (4.165) follows from the Korovkin property on E_0 as exhibited in the formulas (4.142) and (4.143) of Definition 4.4, provided that the limit in (4.164) exists. The existence of this limit, and its identification are given in (4.166) and (4.167) respectively. For this to make sense we must be sure that the partial sums of the first $n+1$ terms of the quantities in (4.164) and (4.166) are equal. In fact a rewriting of the quantity in (4.164) before taking the limit shows that the quantity in (4.174) is also equal to (4.164); i.e.

$$\begin{aligned} &\inf_{h_0 \in D(L)} \sup_{x_1 \in E_0} \inf_{h_1 \in D(L)} \sup_{x_2 \in E_0} \cdots \inf_{h_n \in D(L)} \sup_{x_{n+1} \in E_0} \left\{ \sum_{j=0}^{n} \cdots \right\} \\ &= \inf_{h_j \in D(L),\, 0 \leqslant j \leqslant n} \; \max_{x_j \in E_0,\, 1 \leqslant j \leqslant n+1} \left\{ \sum_{j=0}^{n} \cdots \right\}. \end{aligned}$$

In fact the same is true for the corresponding partial sums in (4.165) and (4.175), but with inf instead of sup, and min instead of max. For $0 < \lambda < 2\lambda_0$, we have $|\lambda_0 - \lambda| < \lambda_0$. Since

$$|\lambda_0 - \lambda| \left\| R\,(\lambda_0)\,f \right\|_\infty \leqslant \left| \frac{\lambda_0 - \lambda}{\lambda_0} \right| \left\| f \right\|_\infty, \quad f \in C_b\,(E, \mathbb{R})\,, \tag{4.185}$$

the sum in (4.167) converges uniformly. The equality of the sum of the first $n+1$ terms in (4.164) and (4.166) can be proved as follows. For $1 \leqslant k \leqslant n$ we may employ the following identities:

$$\begin{aligned}
&\inf_{h_0 \in D(L)} \sup_{x_1 \in E_0} \cdots \inf_{h_n \in D(L)} \sup_{x_{n+1} \in E_0} \\
&\sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \\
&= \inf_{h_0 \in D(L)} \sup_{x_1 \in E_0} \cdots \inf_{h_{n-k} \in D(L)} \sup_{x_{n-k+1} \in E_0} \\
&\sum_{j=0}^{n-k} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \\
&\quad + \frac{\lambda}{\lambda_0} \sum_{j=n-k+1}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left(\lambda_0 R(\lambda_0)\right)^{j-(n-k)} g(x_{n-k+1}). \qquad (4.186)
\end{aligned}$$

The equality in (4.186) can be proved by induction with respect to k, and by repeatedly employing the definition of $\lambda_0 R(\lambda_0) f$, $f \in C_b(E, \mathbb{R})$, together with its linearity. Using (4.186) with $k = n$ we get

$$\begin{aligned}
&\inf_{h_0 \in D(L)} \sup_{x_1 \in E_0} \cdots \inf_{h_n \in D(L)} \sup_{x_{n+1} \in E_0} \\
&\sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j(x_{j+1}) \right\} \\
&= \inf_{h_0 \in D(L)} \sup_{x_1 \in E_0} \left[\left\{ h_0(x_0) + \frac{\lambda}{\lambda_0} g(x_1) - \left(I - \frac{1}{\lambda_0} L\right) h_0(x_1) \right\} \right. \\
&\qquad \left. + \frac{\lambda}{\lambda_0} \sum_{j=1}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left(\lambda_0 R(\lambda_0)\right)^j g(x_1) \right] \\
&= \frac{\lambda}{\lambda_0} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left(\lambda_0 R(\lambda_0)\right)^{j+1} g(x_0). \qquad (4.187)
\end{aligned}$$

From the equality of (4.164) and (4.165), together with (4.187) we infer

$$\lambda R(\lambda) g(x_0) = \lim_{n \to \infty} \frac{\lambda}{\lambda_0} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^j \left(\lambda_0 R(\lambda_0)\right)^{j+1} g(x_0). \qquad (4.188)$$

Notice that by (4.185) the series in (4.188) converges uniformly. Consequently, the equalities of the quantities in (4.164), (4.165), (4.166), (4.167), (4.174), and (4.175) follow, and all these expressions are equal to

$\lambda R(\lambda) g\,(x_0)$. Next let $(h_j)_{j\in\mathbb{N}} \subset D(L)$ be any sequence with the following property:

$$(\lambda I - L)\, h_0 = \frac{\lambda}{\lambda_0} \sum_{j=1}^{\infty} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j-1} (\lambda I - L)\, h_j \tag{4.189}$$

where the series in (4.189) converges uniformly. Then by the maximum principle the series $\frac{\lambda}{\lambda_0} \sum_{j=1}^{\infty} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j-1} h_j$ converges uniformly as well. So it makes sense to write:

$$h_0 = \frac{\lambda}{\lambda_0} \sum_{j=1}^{n+1} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j-1} h_j' \quad \text{where} \quad h_j' = h_j, \quad 1 \leqslant j \leqslant n, \text{ and}$$

$$h_{n+1}' = \sum_{j=n+1}^{\infty} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j-n-1} h_j. \tag{4.190}$$

Again the series in (4.190) converges uniformly. From the representation of h_0 in (4.190) we infer the equalities:

$$\inf_{h_j \in D(L),\, 0 \leqslant j \leqslant n} \max_{x_j \in E_0,\, 1 \leqslant j \leqslant n+1} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j} \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g\,(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j\,(x_{j+1}) \right\} \tag{4.191}$$

$$= \inf_{\substack{h_j \in D(L),\ (\lambda I - L)\, h_0 = \frac{\lambda}{\lambda_0} \sum_{j=1}^{\infty} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j-1} (\lambda I - L)\, h_j \\ j \geqslant 0}} \max_{\substack{x_j \in E_0 \\ 1 \leqslant j \leqslant n+1}} \sum_{j=0}^{n} \left(1 - \frac{\lambda}{\lambda_0}\right)^{j} \left\{ h_j(x_j) + \frac{\lambda}{\lambda_0} g(x_{j+1}) - \left(I - \frac{1}{\lambda_0} L\right) h_j\,(x_{j+1}) \right\}. \tag{4.192}$$

Hence, the equality of (4.174) and (4.168) follows. A similar argument shows the equality of (4.175) and (4.173). Of course, here we used the equality of (4.191) and (4.192) with "inf" instead of "sup", and "max" replaced with "min" and vice versa. So we have equality of the following expressions: (4.164), (4.165), (4.166), (4.167), (4.168), (4.173), (4.174), (4.175). The proof of the fact that these quantities are also equal to (4.169), (4.170), (4.171), and (4.172) is still missing. Therefore we first show that the expression in (4.168) is greater than or equal to (4.169). In a similar manner it is shown that the expression in (4.173) is less than or equal to (4.172): in fact by applying the inequality (4.168) $\geqslant$ (4.169) to $-g$ instead of $+g$ we obtain that (4.173) is less than or equal to (4.172). From the (local) maximum principle it will follow that the expression in (4.169) is greater than or equal to (4.172). As a consequence we will obtain that, with the exception of (4.170) and (4.171), all quantities in Proposition 4.4 are equal. Proving the equality of (4.169) and (4.170), and of (4.171) and (4.172) is a separate issue. In fact the equality of (4.169) and (4.170) follows from Lemma 4.1 equality (4.153), and the equality of (4.171) and (4.172) follows from the same lemma equality (4.154).

(4.168) $\geqslant$ (4.169). Fix the subset E_0 of E, and let $(h_j)_{j\in\mathbb{N}} \subset D(L)$ with the following property: $Lh_0 = \lim_{n\to\infty} \frac{\lambda}{\lambda_0} \sum_{j=1}^{n} \left(1-\frac{\lambda}{\lambda_0}\right)^{j-1} h_j$. Here the convergence is uniform on E_0. In fact each h_j may chosen equal to h_0. In (4.168) we choose all $x_j = x \in E_0$. Then we get

$$\begin{aligned}
&\sum_{j=0}^{n}\left(1-\frac{\lambda}{\lambda_0}\right)^{j}\left\{h\left(x_j\right)+\frac{\lambda}{\lambda_0}g\left(x_{j+1}\right)-\left(I-\frac{1}{\lambda_0}L\right)h_j\left(x_{j+1}\right)\right\} \\
&= h_0\left(x_0\right)+\sum_{j=1}^{n}\left(1-\frac{\lambda}{\lambda_0}\right)^{j}h_j\left(x\right)+\frac{\lambda}{\lambda_0}\sum_{j=0}^{n}\left(1-\frac{\lambda}{\lambda_0}\right)^{j}g(x)-h_0(x) \\
&\quad -\sum_{j=1}^{n}\left(1-\frac{\lambda}{\lambda_0}\right)^{j}h_j\left(x\right)+\frac{1}{\lambda}L\left(\frac{\lambda}{\lambda_0}\sum_{j=0}^{n}\left(1-\frac{\lambda}{\lambda_0}\right)^{j}h_j\right)(x) \\
&= h_0\left(x_0\right)+\frac{\lambda}{\lambda_0}\sum_{j=0}^{n}\left(1-\frac{\lambda}{\lambda_0}\right)^{j}g(x)-\left(I-\frac{1}{\lambda}L\right)h_0\left(x\right) \\
&\quad -\frac{1}{\lambda}\left(1-\frac{\lambda}{\lambda_0}\right)^{n+1}L\left(\frac{\lambda}{\lambda_0}\sum_{j=n+1}^{\infty}\left(1-\frac{\lambda}{\lambda_0}\right)^{j-n-1}h_j\right)(x).
\end{aligned} \tag{4.193}$$

The expression in (4.193) tends to

$$h_0\left(x_0\right)+g(x)-\left(I-\frac{1}{\lambda}L\right)h_0(x) \quad \text{uniformly on } E_0, \tag{4.194}$$

and consequently, since $h_0 \in D(L)$ may be chosen arbitrarily, we see that (4.168) $\geqslant$ (4.169).

(4.173) $\leqslant$ (4.172). The proof of this inequality follows the same lines as the proof of (4.168) $\geqslant$ (4.169). In fact it follows from the latter inequality by applying it to $-g$ instead of g. The reader is invited to check the details.

(4.169) $\geqslant$ (4.172). Consider the mapping $\Lambda^+ : C_b(E,\mathbb{R}) \to [-\infty+\infty)$ defined by

$$\Lambda^+(g) = \inf_{h\in D(L)} \sup_{x\in E_0} \left\{h\left(x_0\right)+g(x)-\left(I-\frac{\lambda}{\lambda_0}L\right)h(x)\right\}. \tag{4.195}$$

where $g \in C_b(E, \mathbb{R})$. From the weak maximum principle (see Definition 4.5) and Lemma 4.1, inequality (4.156) it follows that Λ^+ attains its values in $\mathbb{R}$. In addition, the functional Λ^+ is sub-additive, and the expression in (4.172) is equal to $-\Lambda^+(-g)$. It follows that

$$\begin{aligned}&\Lambda^+(g) + \Lambda^+(-g) \geqslant \Lambda^+(0)\\ &= \inf_{h \in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda}L\right) h \geqslant 0, \quad \text{on } E_0 \right\} \geqslant 0. \end{aligned} \tag{4.196}$$

In (4.196) we used the weak maximum principle: compare with the arguments in (4.163) of the proof of inequality (4.157) in Lemma 4.1.

This establishes Proposition 4.4. □

Proposition 4.5. *Suppose that the operator L possesses the Korovkin property on E_0. Then for all $\lambda > 0$ and $f \in C_b(E)$ the quantities in (4.169), (4.170), (4.171), and (4.172) are equal for all $x_0 \in E$. These quantities are also equal to*

$$\sup_{v \in H^+(E)} \inf_{h \in D(L)} \left\{ h(x_0) : v \left(I - \frac{1}{\lambda}L\right) h \geqslant vg \right\} \tag{4.197}$$

$$= \inf_{v \in H^+(E)} \sup_{h \in D(L)} \left\{ h(x_0) : v \left(I - \frac{1}{\lambda}L\right) h \leqslant vg \right\}. \tag{4.198}$$

Recall that $H^+(E)$ stands for all functions $u \in H(E)$, $u \geqslant 0$, with the property that for every $\alpha > 0$ the level set $\{u \geqslant \alpha\}$ is a compact subset of E. Observe that for every $u \in H(E)$ there exists a function $u_0 \in H^+(E)$ such that $|u(x)| \leqslant u_0(x)$ for all $x \in E$.

Corollary 4.3. *Suppose that the operator L possesses the Korovkin property on E_0, and is positive $\mathcal{T}_\beta$-dissipative on E_0. Then the family $\{\lambda R(\lambda) : \lambda \geqslant \lambda_0\}$, as defined in Proposition 4.4, is $\mathcal{T}_\beta$-equi-continuous (on E_0) for some $\lambda_0 > 0$.*

Proof. We use the representation in (4.171):

$$\lambda R(\lambda) f(x_0) = \sup_{h \in D(L)} \left\{ h(x_0) : \left(I - \frac{1}{\lambda}L\right) h \leqslant f \quad \text{on } E_0 \right\}. \tag{4.199}$$

Let $u \in H^+(E)$ and $x_0 \in E$. Since L is supposed to be positive $\mathcal{T}_\beta$-dissipative on E_0, there exists $\lambda_0 > 0$ and $v \in H^+(E_0)$ such that

$$\lambda u(x_0) h(x_0) \leqslant \sup_{x \in E_0} v(x) (\lambda h(x) - Lh(x)) \tag{4.200}$$

for all $h \in D(L)$ which are real-valued and for all $\lambda \geqslant \lambda_0$. For the precise definition of positive $\mathcal{T}_\beta$-dissipativity (on E) see (4.15) in Definition 4.2. From (4.199) and (4.200) we infer:

$$
\begin{aligned}
&u(x_0)\,\lambda R(\lambda)f(x_0) \\
&= \sup_{h\in D(L)} \left\{u(x_0)\,h(x_0) : \lambda h - Lh \leqslant \lambda f \quad \text{on } E_0\right\} \\
&\leqslant \sup_{h\in D(L)} \left\{u(x_0)\,h(x_0) : \lambda h - Lh \leqslant \lambda f \quad \text{on } E_0\right\} \\
&\leqslant \sup_{h\in D(L)} \left\{\frac{1}{\lambda}\sup_{x\in E} v(x)\left(\lambda h(x) - Lh(x)\right) : \lambda h - Lh \leqslant \lambda f \quad \text{on } E_0\right\} \\
&\leqslant \sup_{x\in E_0} v(x)f(x).
\end{aligned}
\tag{4.201}
$$

Since by construction $\Re R(\lambda)f = R(\lambda)\Re f$, (4.201) implies:

$$
\left\|u\lambda R(\lambda)f\right\|_\infty \leqslant \left\|vf\right\|_\infty, \quad f \in C_b(E_0), \ \lambda \geqslant \lambda_0. \tag{4.202}
$$

The conclusion in Corollary 4.3 is a consequence of (4.202). □

In the following theorem we wrap up more or less everything we proved so far about an operator with the Korovkin property on a subset E_0 of E. Theorem 4.4 and the related observations were used in the proof of Theorem 2.13.

Theorem 4.4. *Let E_0 be a Polish subspace of the Polish space E. Suppose that every function $f \in C_b(E_0)$ can be extended to a bounded continuous function on E. Let L be a linear operator with domain and range in $C_b(E)$ which assigns the zero function to a constant function. Suppose that the operator L possesses the following properties:*

(1) Its domain $D(L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$.
(2) The operator L assigns real-valued functions to real-valued functions: $\Re(Lf) = L\Re f$ for all $f \in D(L)$.
(3) If $f \in D(L)$ vanishes on E_0, then Lf vanishes on E_0 as well.
(4) The operator L satisfies the maximum principle on E_0.
(5) The operator L is positive $\mathcal{T}_\beta$-dissipative on E_0.
(6) The operator L is sequentially λ-dominant on E_0 for some $\lambda > 0$.
(7) The operator L has the Korovkin property on E_0.

Let $L{\restriction}_{E_0}$ be the operator defined by $D\left(L{\restriction}_{E_0}\right) = \left\{f{\restriction}_{E_0} : f \in D(L)\right\}$, and $L{\restriction}_{E_0}\left(f{\restriction}_{E_0}\right) = Lf{\restriction}_{E_0}$, $f \in D(L)$. Then the operator $L{\restriction}_{E_0}$ possesses a unique linear extension to the generator L_0 of a Feller semigroup $\{S_0(t) : t \geqslant 0\}$ on $C_b(E_0)$.

In addition, the time-homogeneous Markov process associated to the Feller semigroup $\{S_0(t) : t \geqslant 0\}$ serves as the unique solution to the martingale problem associated with L.

Proof. *Existence.* First we prove that the restriction operator $L\restriction_{E_0}$ is well-defined and that it is $\mathcal{T}_\beta$-densely defined. The fact that it is well-defined follows from 3. In order to prove that it is $\mathcal{T}_\beta$-densely defined, we use a Hahn-Banach type argument. Let $\widetilde{\mu}$ be a bounded Borel measure on E_0 such that $\langle f\restriction_{E_0}, \widetilde{\mu}\rangle = \int_{E_0} f d\widetilde{\mu} = 0$ for all $f \in D(L)$. Define the measure μ on the Borel field of E by $\mu(B) = \widetilde{\mu}\left(B \bigcap E_0\right)$, $B \in \mathcal{E}$. Then $\langle f, \mu\rangle = 0$ for all $f \in D(L)$. Since $D(L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$, we infer $\langle f, \mu\rangle = 0$ for all $f \in C_b(E)$. Let $\widetilde{f} \in C_b(E)$. Then there exists $f \in C_b(E)$ such that $f = \widetilde{f}$ on E_0, and hence

$$\left\langle \widetilde{f}, \widetilde{\mu}\right\rangle = \langle f\restriction_{E_0}, \widetilde{\mu}\rangle = \langle f, \mu\rangle = 0. \tag{4.203}$$

From (4.203) we see that a bounded Borel measure which annihilates $D\left(L\restriction_{E_0}\right)$ also vanishes on $C_b\left(E_0\right)$. By the theorem of Hahn-Banach in combination with the fact that every element of the dual of $\left(C_b\left(E_0\right), \mathcal{T}_\beta\right)$ can be identified with a bounded Borel measure on E_0, we see that the subspace $D\left(L\restriction_{E_0}\right)$ is $\mathcal{T}_\beta$-dense in $C_b\left(E_0\right)$. Define the family of operators $\{\lambda R(\lambda) : \lambda > 0\}$ as in Proposition 4.4, By the properties 4 and 7 such definitions make sense. Moreover, the family $\{R(\lambda) : \lambda > 0\}$ possesses the resolvent property: $R(\lambda) - R(\mu) = (\mu - \lambda) R(\mu)R(\lambda)$, $\lambda > 0$, $\mu > 0$. It also follows that $R(\lambda)\left(\lambda I - D_1 - L\right) f = f$ on E_0 for $f \in D^{(1)}(L)$. This equality is an easy consequence of the inequalities in (3.161): see Corollary 4.2. Fix $\lambda > 0$ and $f \in C_b\left(E_0\right)$. If f is of the form $f = R(\lambda)g$, $g \in C_b\left(E_0\right)$, then by the resolvent property we have

$$\alpha R(\alpha)f - f = \alpha R(\alpha)R(\lambda)g - R(\lambda)g = \frac{\alpha}{\alpha - \lambda}R(\lambda)g - R(\lambda)g - \frac{\alpha R(\alpha)g}{\alpha - \lambda}. \tag{4.204}$$

Since $\left\|\alpha R(\alpha)g\right\|_\infty \leqslant \|g\|_\infty$, $g \in C_b\left(E_0\right)$, the equality in (4.204) yields

$$\|\cdot\|_\infty\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha)f - f = 0 \quad \text{for } f \text{ of the form } f = R(\lambda)g,\ g \in C_b\left(K\right).$$

Since $g = R(\lambda)\left(\lambda I - D_1 - L\right) g$ on K, $g \in D^{(1)}(L)$, it follows that

$$\lim_{\alpha\to\infty} \left\|\alpha R(\alpha)g - g\right\|_\infty = 0 \quad \text{for } g \in D^{(1)}(L) = D\left(D_1\right) \bigcap D(L). \tag{4.205}$$

As was proved in Corollary 4.3 there exists $\lambda_0 > 0$ such that the family $\{\lambda R(\lambda) : \lambda \geqslant \lambda_0\}$ is $\mathcal{T}_\beta$-equi-continuous. Hence for $u \in H^+\left(E_0\right)$ there exists $v \in H^+\left(K\right)$ that for $\alpha \geqslant \lambda_0$ we have

$$\left\|u\alpha R(\alpha)g\right\|_\infty \leqslant \|vg\|_\infty, \quad g \in C_b\left(E_0\right). \tag{4.206}$$

Fix $\varepsilon > 0$, and choose for $f \in C_b(E_0)$ and $u \in H^+(E_0)$ given the function $g \in D(L\restriction_{E_0})$ in such a way that

$$\|u(f-g)\|_\infty + \|v(f-g)\|_\infty \leqslant \frac{2}{3}\varepsilon. \tag{4.207}$$

Since $D(L\restriction_{E_0})$ is $\mathcal{T}_\beta$-dense in $C_b(E_0)$ such a choice of g. The inequality (4.207) and the identity

$$\begin{aligned}
&\alpha R(\alpha)f - f = \alpha R(\alpha)(f-g) - (f-g) + \alpha R(\alpha)g - g \quad \text{yield} \\
&\|u(\alpha R(\alpha)f - f)\|_\infty \\
&\leqslant \|u(\alpha R(\alpha)(f-g))\|_\infty + \|u(f-g)\|_\infty + \|u(\alpha R(\alpha)g - g)\|_\infty \\
&\leqslant \|v(f-g)\|_\infty + \|u(f-g)\|_\infty + \|u(\alpha R(\alpha)g - g)\|_\infty \\
&\leqslant \frac{2}{3}\varepsilon + \|u(\alpha R(\alpha)g - g)\|_\infty .
\end{aligned} \tag{4.208}$$

From (4.205) and (4.208) we infer

$$\mathcal{T}_\beta\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha)f = f, \quad f \in C_b(E_0). \tag{4.209}$$

Define the operator L_0 in $C_b(E_0)$ as follows. Its domain is given by $D(L_0) = R(\lambda)C_b(E_0)$, $\lambda > 0$. By the resolvent property the space $R(\lambda)C_b(E_0)$ does not depend on $\lambda > 0$, and so $D(L_0)$ is well-defined. The operator $L_0 : D(L_0) \to C_b(E_0)$ is defined by $L_0R(\lambda)f = \lambda R(\lambda)f - f$, $f \in C_b(E_0)$. Since $R(\lambda)f_1 = R(\lambda)f_2$, f_1, $f_2 \in C_b(E_0)$, implies $R(\lambda)(f_2 - f_1) = 0$. By the resolvent property we see that $\alpha R(\alpha)(f_2 - f_1) = 0$ for all $\alpha > 0$. From (4.209) we infer $f_2 = f_1$. In other words, the operator L_0 is well-defined. Since the operators $R(\lambda)$, $\lambda > 0$, are $\mathcal{T}_\beta$-continuous it follows that the graph of the operator L_0 is $\mathcal{T}_\beta$-closed. As in the proof of (iii) $\Longrightarrow$ (i) we have, like in (4.111),

$$\widetilde{S}_0(t)f = \lim_{\lambda\to\infty} e^{-\lambda t}\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!}(\lambda R(\lambda))^k f = \lim_{\lambda\to\infty} e^{-\lambda t}e^{\lambda t(\lambda R(\lambda))}f = S_0(t)f \tag{4.210}$$

where the operator $\widetilde{S}_0(t)$ is defined by using the Hausdorff-Bernstein-Widder Laplace inversion theorem we have (compare with (4.108))

$$R(\lambda)f(x) = (\lambda I - L_0)^{-1} f(x) = \int_0^\infty e^{-\lambda\rho}\widetilde{S}_0(\rho)f(x)d\rho, \quad \Re\lambda > 0,\ x \in E_0. \tag{4.211}$$

For a function f belonging to the space $R(\lambda)C_b(E_0)$ the equality $\widetilde{S}_0(t)f = S_0(t)f$ holds: see (4.113). Here $S_0(t)f$ is defined as the uniform limit in (4.210). Since the operator $L\restriction_{E_0}$ is sequentially λ-dominant for some $\lambda > 0$

we infer that the family of operators $\left\{(\mu R(\lambda+\mu))^k : \mu \geqslant 0, \quad k \in \mathbb{N}\right\}$ is $\mathcal{T}_\beta$-equi-continuous: see (2.44) in Proposition 2.3. Since for $f \in R(\alpha)C_b(E_0)$ we have

$$\sup\left\{e^{-\lambda t}S_0(t)f : t \geqslant 0\right\} = \sup\left\{(\mu R(\lambda+\mu))^k f : \mu \geqslant 0, \quad k \in \mathbb{N}\right\}. \tag{4.212}$$

From (4.212) combined with the $\mathcal{T}_\beta$-equi-continuity and the $\mathcal{T}_\beta$ density of $D\left(L\restriction_{E_0}\right)$ we see that, each operator $S_0(t)$ has a $\mathcal{T}_\beta$-continuous extension to all of $C_b(E_0)$. One way of achieving this is by fixing $f \in C_b(E_0)$, and considering the family $\{\alpha R(\alpha)f : \alpha \geqslant \lambda\}$. Then $\mathcal{T}_\beta$-$\lim_{\alpha\to\infty} \alpha R(\alpha)f = f$. Let $u \in H^+(E_0)$. By the $\mathcal{T}_\beta$-equi-continuity of the family $\left\{e^{-\lambda t}S_0(t) : t \geqslant 0\right\}$ we see that

$$\lim_{\alpha,\,\beta\to\infty} \sup_{t\geqslant 0} \left\| u e^{-\lambda t} S_0(t)\left(\beta R(\beta)f - \alpha R(\alpha)f\right)\right\|_\infty = 0. \tag{4.213}$$

Since the functions $(t,x) \mapsto S_0(t)\left(\alpha R(\alpha)f\right)(x)$, $\alpha \geqslant \lambda$, are continuous the same is true for the function $(t,x) \mapsto \left[S_0(t)f\right](x)$, where $S_0(t)f = \mathcal{T}_\beta\text{-}\lim_{\alpha\to\infty} \alpha R(\alpha)f$. Of course, for almost all $t \geqslant 0$ we have $S_0(t)f(x) = \widetilde{S}_0(t)t(x)$ for all $x \in E_0$. Since

$$\mathcal{T}_\beta\text{-}\lim_{t\downarrow} \frac{1}{t}\left(I - e^{-\lambda t}S_0(t)\right)R(\lambda)f = R(\lambda)f, \quad f \in C_b(E_0),$$

we see that the operator L_0 generates the semigroup $\{S_0(t) : t \geqslant 0\}$. The continuous extension of $S_0(t)$, which was originally defined on $R(\lambda)C_b(E_0)$, to $C_b(E_0)$ is again denoted by $S_0(t)$. Let $f \in D(L)$. Moreover, since

$$R(\lambda)(\lambda f - Lf) = f \quad \text{on } E_0,$$

we have $D\left(L\restriction_{E_0}\right) \subset D(L_0)$, and

$$\begin{aligned} L_0 f &= L_0 R(\lambda)(\lambda I - L)f = \lambda R(\lambda)(\lambda I - L)f - (\lambda I - L)f \\ &= \lambda f - \lambda f + Lf = Lf \end{aligned} \tag{4.214}$$

on E_0. From (4.214) we see that the operator L_0 extends the operator $L\restriction_{E_0}$.

Uniqueness of Feller semigroups. Let L_1 and L_2 be two extensions of the operator $L\restriction_{E_0}$ which generate Feller semigroups. Let $\{R_1(\lambda) : \lambda > 0\}$ and $\{R_2(\lambda) : \lambda > 0\}$ be the corresponding resolvent families. Since L_1 extends $L\restriction_{E_0}$ we obtain, for $h \in D(L)$,

$$\lambda_0 R(\lambda_0)\left(I - \frac{1}{\lambda_0}L\right)h = R(\lambda_0)(\lambda_0 I - L_1)h = h. \tag{4.215}$$

Then by the maximum principle and (4.215) we infer

$$\begin{aligned}
&\sup_{h\in D(L)}\inf_{x\in E_0}\left(h\left(x_0\right)+g(x)-\left(I-\frac{1}{\lambda_0}L\right)h\left(x\right)\right) \\
&\leqslant \sup_{h\in D(L)}\left(h\left(x_0\right)+\lambda_0 R\left(\lambda_0\right)\left(g-\left(I-\frac{1}{\lambda_0}L\right)h\right)\left(x_0\right)\right) \\
&= \sup_{h\in D(L)}\left(h\left(x_0\right)+\lambda_0 R_1\left(\lambda_0\right)g\left(x_0\right)-\lambda_0 R_1\left(\lambda_0\right)\left(I-\frac{1}{\lambda_0}L\right)h\left(x_0\right)\right) \\
&= \sup_{h\in D(L)}\left(h\left(x_0\right)+\lambda_0 R_1\left(\lambda_0\right)g\left(x_0\right)-h\left(x_0\right)\right) \\
&= \lambda_0 R_1\left(\lambda_0\right)g\left(x_0\right) \\
&\leqslant \inf_{h\in D(L)}\sup_{x\in E_0}\left(h\left(x_0\right)+g(x)-\left(I-\frac{1}{\lambda_0}L\right)h(x)\right). \qquad (4.216)
\end{aligned}$$

The same reasoning can be applied to the operator $R_2\left(\lambda_0\right)$ Since the extremities in (4.215) are equal we see that $R_1\left(\lambda_0\right)=R_2\left(\lambda_0\right)$. Hence we get $\left(\lambda_0-L_1\right)^{-1}=\left(\lambda_0-L_2\right)^{-1}$, and consequently $L_1=L_2$.

Of course the same arguments work if $E_0=E$.

Uniqueness of solutions to the martingale problem. Let L_0 be the (unique) extension of L, which generates a Feller semigroup $\{S_0(t) : t\geqslant 0\}$, and let

$$\left\{\left(\Omega,\mathcal{F},\mathbb{P}_x\right),\left(X(t),t\geqslant 0\right),\left(\vartheta_t,t\geqslant 0\right),\left(E,\mathcal{E}\right)\right\}$$

be the corresponding time-homogeneous Markov process with $\mathbb{E}_x\left[g(X(t)\right]=S_0(t)g(x)$, $g\in C_b(E)$, $x\in E$, $t\geqslant 0$. Then the family $\{\mathbb{P}_x : x\in E\}$ is a solution to the martingale problem associated to L. The proof of the uniqueness part follows a pattern similar to the proof of the uniqueness part of linear extensions of L which generate Feller semigroups. We will show that the family of probability measures $\{\mathbb{P}_x : x\in E\}$ is a solution to the martingale problem associated to the operator L. Let f be a member of $D(L)$ and put $M_f(t)=f(X(t))-f(X(0))-\int_0^t Lf(X(s))ds$. Then, for $t_2>t_1$ we have

$$\mathbb{E}_x\left[M_f\left(t_2\right)\,\middle|\,\mathcal{F}_{t_1}\right]-M_f\left(t_1\right)=\mathbb{E}_x\left[M_f\left(t_2-t_1\right)\circ\vartheta_{t_1}\,\middle|\,\mathcal{F}_{t_1}\right]$$

(Markov property)

$$=\mathbb{E}_{X(t_1)}\left[M_f\left(t_2-t_1\right)\right]. \qquad (4.217)$$

Since, in addition, by virtue of the fact that L_0, which is an extension of L, generates the semigroup $\{S_0(t) : t \geqslant 0\}$, we have

$$\begin{aligned}\mathbb{E}_z\left[M_f(t)\right] &= S_0(t)f(z) - f(z) - \int_0^t S_0(u)Lf(z)du \\ &= S_0(t)f(z) - f(z) - \int_0^t \frac{\partial}{\partial u}\left(S_0(u)f(z)\right)du \\ &= S_0(t)f(z) - f(z) - \left(S_0(t)f(z) - S_0(0)f(z)\right) = 0,\end{aligned}$$

the assertion about the existence of solutions to the martingale problem follows from (4.217). Next we prove uniqueness of solutions to the martingale problem. Its proof resembles the way we proved the uniqueness of extensions of L which generate Feller semigroups. Let $\left\{\mathbb{P}_x^{(1)} : x \in E\right\}$ and $\left\{\mathbb{P}_x^{(2)} : x \in E\right\}$ be two solutions to the martingale problem for L. Let $h \in D(L)$, and consider

$$\begin{aligned}&\lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[h\left(X(s)\right) - \left(I - \frac{1}{\lambda}L\right)h\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt \\ &= h\left(X(s)\right) - \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[\left(\lambda I - L\right)h\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt \\ &= h\left(X(s)\right) - \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[h\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt \\ &\quad + \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[Lh\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt\end{aligned}$$

(integration by parts)

$$\begin{aligned}&= h\left(X(s)\right) - \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[h\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt \\ &\quad + \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[\int_0^t Lh\left(X(\rho+s)\right)d\rho \,\middle|\, \mathcal{F}_s\right] dt \\ &= h\left(X(s)\right) - \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[h\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt \\ &\quad + \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[\int_s^{t+s} Lh\left(X(\rho)\right)d\rho \,\middle|\, \mathcal{F}_s\right] dt\end{aligned}$$

(martingale property)

$$= h\left(X(s)\right) - \lambda \int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[h\left(X(t+s)\right) \,\middle|\, \mathcal{F}_s\right] dt$$

$$+\lambda\int_0^\infty e^{-\lambda t}\mathbb{E}_x^{(j)}\left[h\left(X(t+s)\right)-h\left(X(s)\right)\,\middle|\,\mathcal{F}_s\right]dt=0. \tag{4.218}$$

Fix $x_0\in E$, $g\in C_b(E)$, and $s>0$. Then, from (4.218) it follows that, for $h\in D(L)$,

$$\begin{aligned}&\int_0^\infty e^{-\lambda t}\mathbb{E}_{x_0}^{(j)}\left[g\left(X(t+s)\right)\,\middle|\,\mathcal{F}_s\right]dt\\&=\int_0^\infty e^{-\lambda t}\mathbb{E}_{x_0}^{(j)}\left[h\left(X(s)\right)+g\left(X(t+s)\right)-\left(I-\frac{1}{\lambda}L\right)h\left(X(t+s)\right)\,\middle|\,\mathcal{F}_s\right]dt,\end{aligned}$$

and hence

$$\Lambda^-(g,X(s),\lambda)\leqslant\lambda\int_0^\infty\exp(-\lambda t)\mathbb{E}_{x_0}^{(j)}\left[g\left(X(t+s)\right)\,\middle|\,\mathcal{F}_s\right]dt\leqslant\Lambda^+\left(g,X(s),\lambda\right), \tag{4.219}$$

for $j=1,\,2$, where

$$\begin{aligned}\Lambda^+\left(g,x_0,\lambda\right)&=\inf_{\substack{\Gamma\subset D(L)\\\#\Gamma<\infty}}\sup_{\substack{\Phi\subset E_0\\\#\Phi<\infty}}\min_{h\in\Gamma}\max_{x\in\Phi\bigcup\{\triangle\}}\left\{h(x_0)+\left[g-\left(I-\frac{1}{\lambda}L\right)h\right](x)\right\}\\&=\inf_{h\in D(L)}\sup_{x\in E_0}\left\{h(x_0)+\left[g-\left(I-\frac{1}{\lambda}L\right)h\right](x)\right\},\end{aligned} \tag{4.220}$$

and

$$\begin{aligned}\Lambda^-\left(g,x_0,\lambda\right)&=\sup_{\substack{\Gamma\subset D(L)\\\#\Gamma<\infty}}\inf_{\substack{\Phi\subset E_0\\\#\Phi<\infty}}\max_{h\in\Gamma}\min_{x\in\Phi\bigcup\{\triangle\}}\left\{h(x_0)+\left[g-\left(I-\frac{1}{\lambda}L\right)h\right](x)\right\}\\&=\sup_{h\in D(L)}\inf_{x\in E_0}\left\{h(x_0)+\left[g-\left(I-\frac{1}{\lambda}L\right)h\right](x)\right\}.\end{aligned} \tag{4.221}$$

We also have

$$\begin{aligned}&\lambda\int_0^\infty e^{-\lambda t}\mathbb{E}_{X(s)}^{(j)}\left[h\left(X(0)\right)-\left(I-\frac{1}{\lambda}L\right)h\left(X(t)\right)\right]dt\\&=h\left(X(s)\right)-\lambda\int_0^\infty e^{-\lambda t}\mathbb{E}_{X(s)}^{(j)}\left[h\left(X(t)\right)\right]dt+\int_0^\infty e^{-\lambda t}\mathbb{E}_{X(s)}^{(j)}\left[Lh\left(X(t)\right)\right]dt\end{aligned}$$

(integration by parts)

$$\begin{aligned}&=h\left(X(s)\right)-\lambda\int_0^\infty e^{-\lambda t}\mathbb{E}_{X(s)}^{(j)}\left[h\left(X(t)\right)\right]dt\\&\quad+\lambda\int_0^\infty e^{-\lambda t}\mathbb{E}_{X(s)}^{(j)}\left[\int_0^t Lh\left(X(\rho)\right)d\rho\right]dt\end{aligned}$$

(martingale property)

$$= h\left(X(s)\right) - \lambda \int_0^\infty e^{-\lambda t} \mathbb{E}^{(j)}_{X(s)} \left[h\left(X(t)\right)\right] dt$$
$$+ \lambda \int_0^\infty e^{-\lambda t} \mathbb{E}^{(j)}_{X(s)} \left[h\left(X(t)\right) - h\left(X(0)\right)\right] dt = 0 \tag{4.222}$$

where in the first and final step we used $X(0) = z$ $\mathbb{P}^{(j)}_z$-almost surely. In the same spirit as we obtained (4.219) from (4.218), from (4.222) we now get

$$\Lambda^-\left(g, X(s), \lambda\right) \leqslant \lambda \int_0^\infty e^{-\lambda t} \mathbb{E}^{(j)}_{X(s)} \left[g\left(X(t)\right)\right] dt \leqslant \Lambda^+\left(g, X(s), \lambda\right), \tag{4.223}$$

for $j = 1, 2$. Since, by Proposition 4.4 (formula (4.169) and (4.172)) the identity $\Lambda^+(g, x, \lambda) = \Lambda^-(g, x, \lambda)$, is true for $g \in C_b(E)$, $x \in E$, $\lambda > 0$, we obtain, by putting $s = 0$, $\mathbb{E}^{(1)}_x\left[g(X(t))\right] = \mathbb{E}^{(2)}_x\left[g(X(t))\right]$, $t \geqslant 0$, $g \in C_b(E)$. We also obtain, $\mathbb{P}^{(1)}_x$-almost surely,

$$\mathbb{E}^{(1)}_x\left[g(X(t+s)) \,\middle|\, \mathcal{F}_s\right] = \mathbb{E}^{(1)}_{X(s)}\left[g(X(t))\right],$$

and, $\mathbb{P}^{(2)}_x$-almost surely,

$$\mathbb{E}^{(2)}_x\left[g(X(t+s)) \,\middle|\, \mathcal{F}_s\right] = \mathbb{E}^{(2)}_{X(s)}\left[g(X(t))\right], \text{ for } t,\ s \geqslant 0, \text{ and } g \in C_b(E).$$

It necessarily follows that $\mathbb{P}^{(1)}_x = \mathbb{P}^{(2)}_x$, $x \in E$. Consequently, the uniqueness of the solutions to the martingale problem for the operator L follows.

This completes the proof Theorem 4.4. □

4.4 Continuous sample paths

The following Lemma 4.2 and Proposition 4.6 give a general condition which guarantee that the sample paths are $\mathbb{P}_{\tau,x}$-almost surely continuous on their life time.

Lemma 4.2. *Let $P(\tau, x; t, B)$, $0 \leqslant \tau \leqslant t \leqslant T$, $x \in E$, $B \in \mathcal{E}$, be a sub-Markov transition function. Let $(x, y) \mapsto d(x, y)$ be a continuous metric on $E \times E$ and put $B_\varepsilon(x) = \{y \in E : d(y, x) \leqslant \varepsilon\}$. Fix $t \in (0, T]$. Then the following assertions are equivalent:*

(a) For every compact subset K of E and for every $\varepsilon > 0$ the following equality holds:

$$\lim_{s_1, s_2 \to t,\, \tau < s_1 < s_2 \leqslant t} \sup_{x \in K} \frac{P\left(s_1, x; s_2, E \backslash B_\varepsilon(x)\right)}{s_2 - s_1} = 0. \tag{4.224}$$

(b) *For every compact subset K of E and for every open subset G of E such that $G \supset K$ the following equality holds:*

$$\lim_{s_1,s_2\to t,\,\tau\leqslant s_1<s_2\leqslant t}\sup_{x\in K}\frac{P\left(s_1,x;s_2,E\backslash G\right)}{s_2-s_1}=0. \tag{4.225}$$

Proof. (a) $\Longrightarrow$ (b). Let G be an open subset of E and let K be a compact subset of G. Then there exists $\varepsilon>0$, $n\in\mathbb{N}$, and $x_j\in K$, such that

$$G\supset\bigcup_{j=1}^{n}B_{2\varepsilon}\left(x_j\right)\supset\bigcup_{j=1}^{n}\operatorname{int}\left(B_{\varepsilon}\left(x_j\right)\right)\supset K. \tag{4.226}$$

For any $x\in K$ there exists j_0, $1\leqslant j_0\leqslant n$, such that $d\left(x,x_{j_0}\right)<\varepsilon$, and hence for $y\in\operatorname{int}\left(B_\varepsilon(x)\right)$ $d\left(y,x_{j_0}\right)\leqslant d\left(y,x\right)+d\left(x,x_{j_0}\right)<2\varepsilon$. It follows that $B_\varepsilon(x)\subset G$. Consequently, for $x\in K$ and $\tau\leqslant s_1<s_2<t$ we get $P\left(s_1,x;s_2,E\backslash G\right)\leqslant P\left(s_1,x;s_2,B_\varepsilon(x)\right)$. So (b) follows from (a).

(b) $\Longrightarrow$ (a). Fix $\varepsilon>0$ and let K be any compact subset of E. Like in the proof of the implication (a) $\Longrightarrow$ (b) we again choose elements $x_j\in K$, $1\leqslant j\leqslant n$, such that $K\subset\bigcup_{j=1}^{n}\operatorname{int}\left(B_{\varepsilon/4}\left(x_j\right)\right)$. Let $x\in K\bigcap B_{\varepsilon/4}\left(x_j\right)$ and $y\in B_{\varepsilon/2}\left(x_j\right)$. Then $d(y,x)\leqslant d\left(y,x_j\right)+d\left(x_j,x\right)\leqslant\frac{1}{2}\varepsilon+\frac{1}{4}\varepsilon=\frac{3}{4}\varepsilon<\varepsilon$. Suppose that $x\in K\bigcap B_{\varepsilon/4}\left(x_j\right)$. For $\tau\leqslant s_1<s_2<t$ it follows that

$$P\left(s_1,x;s_2,E\backslash B_\varepsilon(x)\right)\leqslant P\left(s_1,x;s_2,E\backslash\operatorname{int}\left(B_{\varepsilon/2}\left(x_j\right)\right)\right),$$

and hence

$$\begin{aligned}&\sup_{x\in K}P\left(s_1,x;s_2,E\backslash B_\varepsilon(x)\right)\\&\leqslant\max_{1\leqslant j\leqslant n}\sup_{x\in K\bigcap B_{\varepsilon/4}(x_j)}P\left(s_1,x;s_2,E\backslash\operatorname{int}\left(B_{\varepsilon/2}\left(x_j\right)\right)\right).\end{aligned} \tag{4.227}$$

The inequality in (4.227) together with assumption in (b) easily implies (a).

This concludes the proof of Lemma 4.2. □

The following proposition clearly shows that in the presence of condition (4.228) the sample paths are almost surely left-continuous on their life-time. Since we may assume that they are right-continuous, the sample paths are $\mathbb{P}_{\tau,x}$-almost surely continuous.

Proposition 4.6. *Let $P\left(\tau,x;t,B\right)$ be a sub-Markov transition function and let the process $X(t)$ be as in Theorem 2.9. Fix $(\tau,x)\in[0,T]\times E$. Suppose that for every $t\in[\tau,T]$, and for every compact subset K and for every open subset G for which $G\supset K$ the equality*

$$\lim_{s_1,s_2\uparrow t,\,\tau\leqslant s_1<s_2<t}\sup_{y\in K}\frac{P\left(s_1,y;s_2,E\backslash G\right)}{s_2-s_1}=0 \tag{4.228}$$

holds. Then for every $t \in (\tau, T]$ *the equality*

$$\inf_{\varepsilon>0} \sup_{t-\varepsilon \leqslant s \leqslant t} d\left(X(s), X(t)\right) \mathbf{1}_{[X(t)\in E]} = 0, \quad \textit{holds } \mathbb{P}_{\tau,x}\textit{-almost surely.}$$

Here $d : E \times E \to [0, \infty)$ *is a continuous metric on* $E \times E$.

Proof. Put $t_{j,n} = t - \varepsilon + j2^{-n}\varepsilon$, $0 \leqslant j \leqslant 2^n$. From Proposition 3.1 with $X(s)$ instead of $\widetilde{X}(s)$ it follows that it suffices to prove that for every $\eta > 0$ the equality

$$\inf_{\varepsilon>0} \lim_{n\to\infty} \mathbb{P}_{\tau,x}\left[\max_{1\leqslant j\leqslant 2^n} d\left(X\left(t_{j-1,n}\right), X\left(t_{j,n}\right)\right) \mathbf{1}_{\{X(t_{j-1,n})\in K\}} \mathbf{1}_{\{X(t_{j,n})\in K\}} > \eta\right] = 0 \tag{4.229}$$

holds for all compact subsets K of E.

$$\begin{aligned}
&\mathbb{P}_{\tau,x}\left[\max_{1\leqslant j\leqslant 2^n} d\left(X\left(t_{j-1,n}\right), X\left(t_{j,n}\right)\right) \mathbf{1}_{\{X(t_{j-1,n})\in K\}} \mathbf{1}_{\{X(t_{j,n})\in K\}} > \eta\right] \\
&\leqslant \sum_{j=1}^{2^n} \mathbb{P}_{\tau,x}\left[d\left(X\left(t_{j-1,n}\right), X\left(t_{j,n}\right)\right) \mathbf{1}_{\{X(t_{j-1,n})\in K\}} \mathbf{1}_{\{X(t_{j,n})\in K\}} > \eta\right]
\end{aligned}$$

(Markov property)

$$\begin{aligned}
&= \sum_{j=1}^{2^n} \mathbb{E}_{\tau,x}\left[\mathbb{P}_{t_{j-1,n},X(t_{j-1,n})}\left[d\left(X\left(t_{j-1,n}\right), X\left(t_{j,n}\right)\right) \mathbf{1}_{\{X(t_{j,n})\in K\}} > \eta\right]\right. \\
&\qquad \left. \times \mathbf{1}_{\{X(t_{j-1,n})\in K\}}\right] \\
&\leqslant \sum_{j=1}^{2^n} \sup_{y\in K} \mathbb{P}_{t_{j-1,n},y}\left[d\left(y, X\left(t_{j,n}\right)\right) \mathbf{1}_{\{X(t_{j,n})\in K\}} > \eta\right] \\
&= \sum_{j=1}^{2^n} \sup_{y\in K} P\left(t_{j-1,n}, y; t_{j,n}, K \backslash B_\eta(y)\right).
\end{aligned} \tag{4.230}$$

The result in Proposition 4.6 follows from (4.230) and Lemma 4.2. □

4.5 Measurability properties of hitting times

In this section we study how fast Markov process reaches a Borel subset B of the state space E. The material is taken from Chapter 2, Section 2.10 in [Gulisashvili and van Casteren (2006)]. Fix $\tau \in [0, T]$. Throughout this section we will assume that the filtrations $(\mathcal{F}_t^\tau)_{t\in[\tau,T]}$ are right-continuous and $P_{\tau,\mu}$-complete. Right-continuity means that $\mathcal{F}_{t+}^{\tau} = \bigcap_{s\in(t,T]} \mathcal{F}_s^{\tau} = \mathcal{F}_t^{\tau}$. By definition the σ-field $\mathcal{F}_t^{\tau}$ is $P_{\tau,\mu}$-complete if $P_{\tau,\mu}$-negligible events A belong to $\mathcal{F}_t^{\tau}$. The σ-field $\overline{\mathcal{F}}_{t+}^{\tau}$ is the $\mathbb{P}_{\tau,\mu}$-completion of a σ-field $\mathcal{F}_{t+}^{\tau}$ if and only if for every $A \in \mathcal{F}_{t+}^{\tau}$ there exist events A_1 and $A_2 \in \mathcal{F}_{t+}^{\tau}$ such that $A_1 \subset A \subset A_2$ and $\mathbb{P}_{\tau,x}\left(A_2 \backslash A_1\right) = 0$. It is also assumed that we are in the context of a backward Feller evolution (or propagator) $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ in the sense of Definition 2.4 and the corresponding strong Markov process with state space E:

$$\left\{ \left(\Omega, \mathcal{F}_T^{\tau}, \mathbb{P}_{\tau,x}\right)_{(\tau,x)\in[0,T]\times E}, \left(X(t), t \in [0,T]\right), (E,\mathcal{E}) \right\}. \tag{4.231}$$

By $P(E)$ will be denoted the collection of all Borel probability measures on the space E. For $A \in \mathcal{F}_T^{\tau}$ and $\mu \in P(E)$, we put $\mathbb{P}_{\tau,\mu}(A) = \int \mathbb{P}_{\tau,x}(A) d\mu(x)$. For instance, if $\mu = \delta_x$ is the Dirac measure concentrated at $x \in E$, then $\mathbb{P}_{\tau,\delta_x} = \mathbb{P}_{\tau,x}$. Let ζ be the first time the process $X(t)$ arrives at the absorption state $\triangle$:

$$\zeta = \begin{cases} \inf\{t > 0 : \ X(t) = \triangle\} & \text{if } X(t) = \triangle \text{ for some } t \in (0,T], \\ T & \text{if } X(t) \in E \text{ for all } t \in (0,T). \end{cases}$$

Definition 4.7. Let $(X(t), \mathbb{P}_{\tau,x})$ be a Markov process on Ω with state space E and sample path space $\Omega = D\left([0,T], E^{\triangle}\right)$, and let B be a Borel subset of $E^{\triangle}$. Let $\tau \in [0,T)$, and suppose that $S : \Omega \to [\tau,\zeta]$ is a $\mathcal{F}_t^{\tau}$-stopping time. For the process $X(t)$, the entry time of the set B after time S is defined by

$$D_B^S = \begin{cases} \inf\{t : t \geqslant S, \ X(t) \in B\} & \text{on} \quad \bigcup_{\tau\leqslant t<T} \{S \leqslant t, \ X(t) \in B\}, \\ \\ \zeta \quad \text{elsewhere.} \end{cases} \tag{4.232}$$

The pseudo-hitting time of the set B after time S is defined by

$$\widetilde{D}_B^S = \begin{cases} \inf\{t : t \geqslant S, \ X(t) \in B\} & \text{on} \quad \bigcup_{\tau<t<T} \{S \leqslant t, \ X(t) \in B\}, \\ \\ \zeta \quad \text{elsewhere.} \end{cases} \tag{4.233}$$

The hitting time of the set B after time S is defined by

$$T_B^S = \begin{cases} \inf\{t : t > S, \ X(t) \in B\} & \text{on} \quad \bigcup_{\tau\leqslant t<T} \{S < t, \ X(t) \in B\}, \\ \\ \zeta \quad \text{elsewhere.} \end{cases} \tag{4.234}$$

Observe that on the event $\{S = \tau,\, X(\tau) \in B\}$ we have $D_B^S = \tau$ and $\widetilde{D}_B^S = T_B^S$. It is not hard to prove that

$$\bigcup_{t:\tau\leqslant t<T} \{S \leqslant t,\, X(t) \in B\} = \bigcup_{t:\tau\leqslant t<T} \{S \vee t \leqslant t,\, X(S \vee t) \in B\} \quad \text{and}$$
$$\bigcup_{t:\tau<t<T} \{S \leqslant t,\, X(t) \in B\} = \bigcup_{t:\tau<t<T} \{S \vee t \leqslant t,\, X(S \vee t) \in B\}.$$

We also have

$$D^S_{B\bigcup\{\triangle\}} = D_B^S \wedge \zeta, \quad \widetilde{D}^S_{B\bigcup\{\triangle\}} = \widetilde{D}_B^S \wedge \zeta, \quad \text{and} \quad T^S_{B\bigcup\{\triangle\}} = T_B^S \wedge \zeta. \tag{4.235}$$

In addition, we have $D_B^S \leqslant \widetilde{D}_B^S \leqslant T_B^S$. Next we will show that the following equalities hold:

$$T_B^S = \inf_{\varepsilon>0} \left\{ D_B^{(\varepsilon+S)\wedge\zeta} \right\} = \inf_{r\in\mathbb{Q}^+} \left\{ D_B^{(r+S)\wedge\zeta} \right\}. \tag{4.236}$$

Indeed on $\left\{T_B^S < \zeta\right\}$, the first equality in (4.236) can be obtained by using the inclusion

$$\{t \geqslant (\varepsilon + S) \wedge \zeta, X(t) \in B\} \subset \{t > S,\, X(t) \in B\}$$

and the fact that for every $t \in [\tau, T)$ and $\omega \in \{S < t,\, X(t) \in B\}$, there exists $\varepsilon > 0$ depending on ω such that $\omega \in \{(\varepsilon + S) \wedge \zeta \leqslant t, X(t) \in B\}$. Since $T_B^S \leqslant D_B^{(\varepsilon+S)\wedge\zeta}$, we see that on the event $\left\{T_B^S = \zeta\right\}$ the first equality in (4.236) also holds. The second equality in (4.236) follows from the monotonicity of the entry time D_B^S with respect to S.

Our next goal is to prove that for the Markov process in (4.231) the entry time D_B^S, the pseudo-hitting time $\widetilde{D}_B^S$, and the hitting time T_B^S are stopping times. Throughout the present section, the symbols $\mathcal{K}(E)$ and $\mathcal{O}(E)$ stand for the family of all compact subsets and the family of all open subsets of the space E, respectively.

The celebrated Choquet capacitability theorem will be used in the proof of the fact that D_B^S, $\widetilde{D}_B^S$, and T_B^S are stopping times. We will restrict ourselves to positive capacities and the pavement of the space E by compact subsets. For more general cases, we refer the reader to [Doob (2001); Meyer (1966)].

Definition 4.8. A function I from the class $\mathcal{P}(E)$ of all subsets of E into the extended real half-line $\bar{\mathbb{R}}_+$ is called a Choquet capacity if it possesses the following properties:

(i) If A_1 and A_2 in $\mathcal{P}(E)$ are such that $A_1 \subset A_2$, then $I(A_1) \leqslant I(A_2)$.

(ii) If $A_n \in \mathcal{P}(E)$, $n \geqslant 1$, and $A \in \mathcal{P}(E)$ are such that $A_n \uparrow A$, then $I(A_n) \to I(A)$ as $n \to \infty$.

(iii) If $K_n \in \mathcal{K}(E)$, $n \geqslant 1$, and $K \in \mathcal{K}(E)$ are such that $K_n \downarrow K$, then $I(K_n) \to I(K)$ as $n \to \infty$.

Definition 4.9. A function $\varphi : \mathcal{K}(E) \to [0, \infty)$ is called strongly sub-additive provided that the following conditions hold:

(i) If $K_1 \in \mathcal{K}(E)$ and $K_2 \in \mathcal{K}(E)$ are such that $K_1 \subset K_2$, then $\varphi(K_1) \leqslant \varphi(K_2)$.

(ii) If K_1 and K_2 belong to $\mathcal{K}(E)$, then

$$\varphi\left(K_1 \bigcup K_2\right) + \varphi\left(K_1 \bigcap K_2\right) \leqslant \varphi(K_1) + \varphi(K_2). \tag{4.237}$$

The following construction allows us to define a Choquet capacity starting with a strongly sub-additive function. Let φ be a strongly sub-additive function satisfying the following additional continuity condition, which could be called "exterior regularity for compact subsets":

(iii) For all $K \in \mathcal{K}(E)$ and all $\varepsilon > 0$, there exists $G \in \mathcal{O}(E)$ such that $K \subset G$ and $\varphi(K') \leqslant \varphi(K) + \varepsilon$ for all compact subsets K' of G.

For any $G \in \mathcal{O}(E)$, put

$$I^*(G) = \sup_{K \in \mathcal{K}(E); K \subset G} \varphi(K). \tag{4.238}$$

Next define a set function $I : \mathcal{P}(E) \to \bar{\mathbb{R}}_+$ by

$$I(A) = \inf_{G \in \mathcal{O}(E); A \subset G} I^*(G), \qquad A \in \mathcal{P}(E). \tag{4.239}$$

It is known that the function I is a Choquet capacity. It is clear that for any $G \in \mathcal{O}(E)$, $I(G) = I^*(G)$. Moreover, it is not hard to see that for any $K \in \mathcal{K}(E)$, $\varphi(K) = I(K)$, because of our exterior regularity assumption (iii).

Definition 4.10. Let $\varphi : \mathcal{K}(E) \to [0, \infty)$ be a strongly subadditive function satisfying condition (iii), and let I be the Choquet capacity obtained from φ (see formulas (4.238) and (4.239)). A subset B of E is said to be I-capacitable if the following equality holds:

$$I(B) = \sup\left\{\varphi(K) : K \subset B, K \in \mathcal{K}(E)\right\}. \tag{4.240}$$

Now we are ready to formulate the Choquet capacitability theorem (see, e.g., [Doob (2001); Dellacherie and Meyer (1978); Meyer (1966)]). We will also need the following version of the Choquet capacity theorem. For a discussion on capacitable subsets see e.g. [Kiselman (2000)]; see [Choquet (1986)], [Benzécri (1995)] and [Dellacherie and Meyer (1978)] as well. For a general discussion on the foundations of probability theory see e.g. [Kallenberg (2002)].

Theorem 4.5. *Let E be a Polish space, and let $\varphi : \mathcal{K}(E) \to [0,\infty)$ be a strongly subadditive function satisfying condition (iii), and let I be the Choquet capacity obtained from φ (see formulas (4.238) and (4.239)). Then every analytic subset of E, and in particular, every Borel subset of E is I-capacitable.*

The definition of analytic sets can be found in [Doob (2001); Dellacherie and Meyer (1978)]. We will only need the Choquet capacitability theorem for Borel sets which form a sub-collection of the analytic sets.

Lemma 4.3. *Let $\tau \in [0,T]$, and let $\{X(t) : t \in [\tau,T]\}$ be an adapted, right-continuous, and quasi left-continuous stochastic process on the filtered probability space $\left(X(t), \left(\overline{\mathcal{F}}_{t+}^{\tau}\right)_{t\in[\tau,T]}, \mathbb{P}_{\tau,x}\right)$. Suppose that S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time such that $\tau \leqslant S \leqslant \zeta$. Then, for any $t \in [\tau,T]$ and $\mu \in P(E)$, the following functions are strongly sub-additive on $\mathcal{K}(E)$ and satisfy condition (iii):*

$$K \mapsto \mathbb{P}_{\tau,\mu}\left[D_K^S \leqslant t\right], \quad \textit{and} \quad K \mapsto \mathbb{P}_{\tau,\mu}\left[\widetilde{D}_K^S \leqslant t\right], \quad K \in \mathcal{K}(E). \tag{4.241}$$

We wrote $\overline{\mathcal{F}}_{t+}^{\tau}$ to indicate that this σ-field is right continuous and $\mathbb{P}_{\tau,x}$-complete.

Proof. We have to check conditions (i) and (ii) in Definition 4.9 and also condition (iii) for the set functions in (4.241). Let $K_1 \in \mathcal{K}(E)$ and $K_2 \in \mathcal{K}(E)$ be such that $K_1 \subset K_2$. Then $D_{K_1}^S \geqslant D_{K_2}^S$, and hence

$$\mathbb{P}_{\tau,\mu}\left[D_{K_1}^S \leqslant t\right] \leqslant \mathbb{P}_{\tau,\mu}\left[D_{K_2}^S \leqslant t\right].$$

This proves condition (i) for the function $K \mapsto \mathbb{P}_{\tau,\mu}\left[D_K^S \leqslant t\right]$. The proof of (i) for the second mapping in (4.241) is similar.

In order to prove condition (iii) for the mapping $K \mapsto \mathbb{P}_{\tau,\mu}\left[D_K^S \leqslant t\right]$, we use assertion (a) in Lemma 4.6. More precisely, let $K \in \mathcal{K}(E)$ and $G_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$, be such as in Lemma 4.6. Then by part (a) of Lemma

4.6 below (note that part (a) of Lemma 4.6 also holds under the restrictions in Lemma 4.3), we get

$$\begin{aligned}\mathbb{P}_{\tau,\mu}\left[D^S_K \leqslant t\right] &\leqslant \inf_{G\in\mathcal{O}(E):G\supset K}\ \sup_{K'\in\mathcal{K}(E):K'\subset G} \mathbb{P}_{\tau,\mu}\left[D^S_{K'} \leqslant t\right] \\ &\leqslant \inf_{n\in\mathbb{N}}\ \sup_{K'\in\mathcal{K}(E):K'\subset G_n} \mathbb{P}_{\tau,\mu}\left[D^S_{K'} \leqslant t\right] \\ &\leqslant \inf_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D^S_{G_n} \leqslant t\right] = \mathbb{P}_{\tau,\mu}\left[D^S_K \leqslant t\right]. \end{aligned} \tag{4.242}$$

It follows from (4.242) that

$$\mathbb{P}_{\tau,\mu}\left[D^S_K \leqslant t\right] = \inf_{G\in\mathcal{O}(E):G\supset K}\ \sup_{K'\in\mathcal{K}(E),\,K'\subset G} \mathbb{P}_{\tau,\mu}\left[D^S_{K'} \leqslant t\right]. \tag{4.243}$$

Now it is clear that the equality in (4.243) implies property (iii) for the mapping $K \mapsto \mathbb{P}_{\tau,\mu}\left[D^S_K \leqslant t\right]$. The proof of (iii) for the mapping $K \mapsto \mathbb{P}_{\tau,\mu}\left[\widetilde{D}^S_K \leqslant t\right]$ is similar. Here we use part (d) in Lemma 4.6 (note that part (d) of Lemma 4.6 also holds under the restrictions in Lemma 4.3).

Next we will prove that the function $K \mapsto \mathbb{P}_{\tau,\mu}\left[D^S_K \leqslant t\right]$ satisfies condition (ii). In the proof the following simple equalities will be used: for all Borel subsets B_1 and B_2 of E,

$$D^S_{B_1\bigcup B_2} = D^S_{B_1} \wedge D^S_{B_2}, \quad \text{and} \tag{4.244}$$

$$D^S_{B_1\bigcap B_2} \geqslant D^S_{B_1} \vee D^S_{B_2}. \tag{4.245}$$

By using (4.244) and (4.245) with $K_1 \in \mathcal{K}(E)$ and $K_2 \in \mathcal{K}(E)$ instead of B_1 and B_2 respectively, we get:

$$\begin{aligned}&\left\{D^S_{K_1\bigcup K_2} \leqslant t\right\} \setminus \left\{D^S_{K_2} \leqslant t\right\} = \left(\left\{D^S_{K_1} \leqslant t\right\} \bigcup \left\{D^S_{K_2} \leqslant t\right\}\right) \setminus \left\{D^S_{K_2} \leqslant t\right\} \\ &= \left\{D^S_{K_1} \leqslant t\right\} \setminus \left\{D^S_{K_2} \leqslant t\right\} = \left\{D^S_{K_1} \leqslant t\right\} \setminus \left(\left\{D^S_{K_1} \leqslant t\right\} \bigcap \left\{D^S_{K_2} \leqslant t\right\}\right) \\ &= \left\{D^S_{K_1} \leqslant t\right\} \setminus \left\{D^S_{K_1} \vee D^S_{K_2} \leqslant t\right\} \subset \left\{D^S_{K_1} \leqslant t\right\} \setminus \left\{D^S_{K_1\bigcap K_2} \leqslant t\right\}. \end{aligned} \tag{4.246}$$

It follows from (4.246) that

$$\begin{aligned}&\mathbb{P}_{\tau,\mu}\left[D^S_{K_1\bigcup K_2} \leqslant t\right] + \mathbb{P}_{\tau,\mu}\left[D^S_{K_1\bigcap K_2} \leqslant t\right] \\ &\leqslant \mathbb{P}_{\tau,\mu}\left[D^S_{K_1} \leqslant t\right] + \mathbb{P}_{\tau,\mu}\left[D^S_{K_2} \leqslant t\right]. \end{aligned} \tag{4.247}$$

Now it is clear that (4.247) implies condition (ii) for the function $K \mapsto \mathbb{P}_{\tau,\mu}\left[D^S_K \leqslant t\right]$. The proof of condition (ii) for the second function in Lemma 4.3 is similar.

This completes the proof of Lemma 4.3. □

The next theorem states that under certain restrictions, the entry time D_B^S, the pseudo-hitting time $\widetilde{D}_B^S$, and the hitting time T_B^S are stopping times. Recall that $\overline{\mathcal{F}}_{t+}^{\tau}$ denote the completion of the σ-field

$$\mathcal{F}_{t+}^{\tau} = \bigcap_{s\in(t,T]} \sigma\left(X(\rho) : \tau \leqslant \rho \leqslant s\right)$$

with respect to the family of measures $\{\mathbb{P}_{s,x} : 0 \leqslant s \leqslant \tau,\ x \in E\}$.

Theorem 4.6. *Let $\tau \in [0,T]$, and $\{X(t) : t \in [\tau,T]\}$ be as in Lemma 4.3:*

(i) The process $X(t)$ is right-continuous and quasi left-continuous on $[0,\zeta)$.
(ii) The σ-fields $\overline{\mathcal{F}}_{t+}^{\tau}$ are $\mathbb{P}_{\tau,x}$-complete and right-continuous for $t \in [\tau,T]$ and $x \in E$.

Then for every $\tau \in [0,T)$ and every $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time $S : \Omega \to [\tau,\zeta]$, the random variables D_B^S, $\widetilde{D}_B^S$, and T_B^S are $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping times.

Proof. We will first prove Theorem 4.6 assuming that it holds for all open and all compact subsets of E. The validity of Theorem 4.6 for such sets will be established in lemmas 4.4 and 4.5 below.

Let B be a Borel subset of E, and suppose that we have already shown that for any $\varepsilon \geqslant 0$ the stochastic time $D_B^{(\varepsilon+S)\wedge\zeta}$ is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. Since

$$T_B^S = \inf_{\varepsilon>0,\varepsilon\in\mathbb{Q}^+} D_B^{(\varepsilon+S)\wedge\zeta}$$

(see (4.236)), we also obtain that T_B^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. Therefore, in order to prove that T_B^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time, it suffices to show that for every Borel subset B of E, the stochastic time D_B^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. Since the process $t \mapsto X(t)$ is continuous from the right, it suffices to prove the previous assertion with S replaced by $(\varepsilon + S) \wedge \zeta$.

Fix $t \in [\tau,T)$, $\mu \in P(E)$, and $B \in \mathcal{E}$. By Lemma 4.3 and the Choquet capacitability theorem, the set B is capacitable with respect to the capacity I associated with the strongly sub-additive function $K \mapsto \mathbb{P}_{\tau,\mu}\left[D_K^S \leqslant t\right]$. Therefore, there exists an increasing sequence $K_n \in \mathcal{K}(E)$, $n \in \mathbb{N}$, and a decreasing sequence $G_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$, such that

$$\begin{aligned} &K_n \subset K_{n+1} \subset B \subset G_{n+1} \subset G_n, \quad n \in \mathbb{N}, \quad \text{and} \\ &\sup_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{K_n}^S \leqslant t\right] = \inf_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{G_n}^S \leqslant t\right]. \end{aligned} \tag{4.248}$$

The arguments in (4.248) should be compared with those in (4.263) below. Put

$$\Lambda_1^{\tau,\mu,S}(t) = \bigcup_{n\in\mathbb{N}} \left\{D_{K_n}^S \leqslant t\right\} \quad \text{and} \quad \Lambda_2^{\tau,\mu,S}(t) = \bigcap_{n\in\mathbb{N}} \left\{D_{G_n}^S \leqslant t\right\}. \tag{4.249}$$

Then Lemma 4.4 implies $\Lambda_2^{\tau,\mu,S}(t) \in \overline{\mathcal{F}}_{t+}^{\tau}$, and Lemma 4.5 gives $\Lambda_1^{\tau,\mu,S}(t) \in \overline{\mathcal{F}}_{t+}^{\tau}$. Moreover, we have

$$\Lambda_1^{\tau,\mu,S}(t) \subset \left\{D_B^S \leqslant t\right\} \subset \Lambda_2^{\tau,\mu,S}(t), \tag{4.250}$$

and

$$\begin{aligned}\mathbb{P}_{\tau,\mu}\left[\Lambda_2^{\tau,\mu,S}(t)\right] &= \inf_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{G_n}^S \leqslant t\right] \\ &= \sup_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{K_n}^S \leqslant t\right] = \mathbb{P}_{\tau,\mu}\left[\Lambda_1^{\tau,\mu,S}(t)\right]. \end{aligned}\tag{4.251}$$

It follows from (4.250) and (4.251) that $\mathbb{P}_{\tau,\mu}\left[\Lambda_2^{\tau,\mu,S}(t)\backslash\Lambda_1^{\tau,\mu,S}(t)\right] = 0$. By using (4.250) again, we see that the event $\left\{D_B^S \leqslant t\right\}$ belongs to the σ-field $\overline{\mathcal{F}}_{t+}^{\tau}$. Therefore, the stochastic time D_B^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. As we have already observed, it also follows that the stochastic time T_B^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time.

A similar argument with D_B^S replaced by $\widetilde{D}_B^S$ shows that the stochastic times $\widetilde{D}_B^S$, $B \in \mathcal{E}$, are $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping times.

This completes the proof of Theorem 4.6. □

Next we will prove two lemmas which have already been used in the proof of Theorem 4.6.

Lemma 4.4. *Let $S : \Omega \to [\tau, \zeta]$ be an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time, and let $G \in \mathcal{O}(E)$. Then the stochastic times D_G^S, $\widetilde{D}_G^S$, and T_G^S are $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping times.*

Proof. It is not hard to see that

$$\begin{aligned}\left\{D_G^S \leqslant t < \zeta\right\} &= \bigcap_{m\in\mathbb{N}} \left\{D_G^S < t + \frac{1}{m}\right\} \bigcap \{t < \zeta\} \\ &= \bigcap_{m\in\mathbb{N}} \bigcup_{\tau\leqslant\rho<t+\frac{1}{m},\rho\in\mathbb{Q}^+} \{S \leqslant \rho,\, X(\rho) \in G\}. \end{aligned}\tag{4.252}$$

We also have

$$\left\{D_G^S \leqslant t\right\} = \left\{D_G^S \leqslant t < \zeta\right\} \bigcup \{\zeta \leqslant t\} = \left\{D_G^S \leqslant t < \zeta\right\} \bigcup \{X(t) = \triangle\}. \tag{4.253}$$

The event on the right-hand side of (4.252) belongs to $\overline{\mathcal{F}}_{t+}^{\tau}$, and hence from (4.252) and (4.253) the stochastic time D_G^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. The fact that $\widetilde{D}_G^S$ is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time follows from

$$\left\{\widetilde{D}_G^S \leqslant t < \zeta\right\} = \bigcap_{m\in\mathbb{N}} \left\{\widetilde{D}_G^S < t + \frac{1}{m}\right\} \bigcap \{t < \zeta\}$$

$$= \bigcap_{m \in \mathbb{N}} \bigcup_{\rho \in (\tau, t+\frac{1}{m}) \bigcap \mathbb{Q}^+} \{S \leqslant \rho, \, X(\rho) \in G\}$$

together with

$$\left\{\widetilde{D}_G^S \leqslant t\right\} = \left\{\widetilde{D}_G^S \leqslant t < \zeta\right\} \bigcup \{X(t) = \triangle\}. \tag{4.254}$$

The equality (4.236) with G instead of B implies that T_G^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time, and completes the proof of Lemma 4.4. □

Lemma 4.5. *Let $S : \Omega \to [\tau, \zeta]$ be an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time, and let $K \in \mathcal{K}\left(E^{\triangle}\right)$. Then the stochastic times D_K^S, $\widetilde{D}_K^S$ and T_K^S are $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping times.*

Proof. First let K be a compact subset of E, and let G_n, $n \in \mathbb{N}$, be a sequence of open subsets of E with the following properties: $K \subset \overline{G}_{n+1} \subset G_n$ and $\bigcap_{n \in \mathbb{N}} G_n = K$. Then every stochastic time $D_{G_n}^S$ is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time (see Lemma 4.4), and for every $\mu \in P(E)$ the sequence of stochastic times $D_{G_n}^S$, $n \in \mathbb{N}$, increases $\mathbb{P}_{\tau,\mu}$-almost surely to D_K^S. This implies that the stochastic time T_K^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. The equality (4.236) with K instead of B then shows that T_K^S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time. Next we will show the $\mathbb{P}_{\tau,\mu}$-almost sure convergence of the sequence $D_{G_n}^S$, $n \in \mathbb{N}$. Put $D_K = \sup\limits_{n \in \mathbb{N}} D_{G_n}^S$. Since $D_{G_n}^S \leqslant D_{G_{n+1}}^S \leqslant D_K^S$, it follows that $D_K \leqslant D_K^S$. By Lemma 4.4, the stochastic times $D_{G_n}^S$, $n \in \mathbb{N}$, are $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping times. It follows from the quasi-continuity from the left of the process $X(t)$, $t \in [0, \zeta)$, that

$$\lim_{n \to \infty} X\left(D_{G_n}^S\right) = X\left(D_K\right) \quad \mathbb{P}_{\tau,\mu}\text{-a.s.}$$

Therefore,

$$X\left(D_K\right) \in \bigcap_n \overline{G}_n = K \quad \mathbb{P}_{\tau,\mu}\text{-a.s.}$$

Since $D_K^S \geqslant S$, we have $D_K^S \leqslant D_K$ $\mathbb{P}_{\tau,\mu}$-almost surely, and hence $D_K^S = D_K$ $\mathbb{P}_{\tau,\mu}$-almost surely. This establishes the $\mathbb{P}_{\tau,\mu}$-almost sure convergence of the sequence $D_{G_n}^S$, $n \in \mathbb{N}$, to D_K^S.

In order to finish the proof of Lemma 4.5, we will establish that for every $\mu \in P(E)$, the sequence of stochastic times $\widetilde{D}_{G_n}^S$ increases $\mathbb{P}_{\tau,\mu}$-almost surely to $\widetilde{D}_K^S$. Put $\widetilde{D}_K = \sup\limits_{n \in \mathbb{N}} \widetilde{D}_{G_n}^S$. Since

$$\widetilde{D}_{G_n}^S \leqslant \widetilde{D}_{G_{n+1}}^S \leqslant \widetilde{D}_K^S,$$

it follows that $\widetilde{D}_K \leqslant \widetilde{D}_K^S$. By using the fact that the process $X(t)$, $t \in [0,\zeta)$, is quasi-continuous from the left, we get

$$\lim_{n\to\infty} X\left(\widetilde{D}_{G_n}^S\right) = X\left(\widetilde{D}_K\right) \quad \mathbb{P}_{\tau,\mu}\text{-a.s.}$$

Therefore

$$X\left(\widetilde{D}_K\right) \in \bigcap_n \overline{G}_n = K \quad \mathbb{P}_{\tau,\mu}\text{-a.s.}$$

Since $\widetilde{D}_K^S \geqslant S$, we have $\widetilde{D}_K^S \leqslant \widetilde{D}_K$ $\mathbb{P}_{\tau,\mu}$-almost surely, and hence $\widetilde{D}_K^S = \widetilde{D}_K$ $\mathbb{P}_{\tau,\mu}$-almost surely. This equality shows that the stochastic time $\widetilde{D}_K^S$ is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time.

We still have to consider the case that $\triangle \in K$. For this we use the equalities $D_{K_0 \bigcup\{\triangle\}}^S = D_{K_0}^S \wedge \zeta$, and $\widetilde{D}_{K_0 \bigcup\{\triangle\}}^S = \widetilde{D}_{K_0}^S \wedge \zeta$ together with the fact that a compact subset K of $E^{\triangle}$ is a compact subset of E or is of the form $K = K_0 \bigcup \{\triangle\}$ where $K_0 \subset E$ is compact. Observe that on the event $\{\zeta \geqslant \tau\}$ ζ is an $\mathcal{F}_t^{\tau}$-stopping time.

This completes the proof of Lemma 4.5. □

Let us return to the study of standard Markov processes. It was established in Theorem 2.9 that if P is a transition sub-probability function such that the backward free Feller propagator $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ associated with P is a strongly continuous (backward) Feller propagator, then there exists a standard Markov process as in (4.268) with $(\tau, x; t, B) \mapsto P(\tau, x; t, B)$ as its transition function. Let $\tau \in [0,T]$, and let $(X(t), \mathcal{F}_t^{\tau}, \mathbb{P}_{\tau,x})$ be a Markov process. Suppose that S is an $\overline{\mathcal{F}}_{t+}^{\tau}$-stopping time such that $\tau \leqslant S \leqslant \zeta$. Fix a measure $\mu \in P(E)$, and denote by $\overline{\mathcal{F}}_T^{S,\vee}$ the completion of the σ-field $\mathcal{F}_T^{S,\vee} = \sigma(S \vee \rho, X(S \vee \rho) : 0 \leqslant \rho \leqslant T)$ with respect to the measure μ. The measure μ is used throughout Lemma 4.6 below. The next theorem provides additional examples of families of stopping times which can be used in the formulation of the strong Markov property with respect to families of measures.

Theorem 4.7. *Let $(X(t), \mathcal{F}_t^{\tau}, \mathbb{P}_{\tau,x})$ be a standard Markov process as in (4.268), and let $B \in \mathcal{E}^{\triangle}$. Then the stopping times D_B^S, $\widetilde{D}_B^S$, and T_B^S are measurable with respect to the σ-field $\overline{\mathcal{F}}_T^{S,\vee}$.*

Proof. Since the stopping time S attains its values in the interval $[\tau,\zeta]$ we see that $\{\zeta \leqslant \rho\} = \{\zeta \leqslant \rho \vee S\} = \{X(\rho \vee S) = \triangle\}$ for all $\rho \in [\tau, T]$. This shows that ζ is measurable with respect to $\overline{\mathcal{F}}_T^{S,\vee}$. By (4.235) we see $D_{B\bigcup\{\triangle\}}^S = D_B^S \wedge \zeta$, $\widetilde{D}_{B\bigcup\{\triangle\}}^S = \widetilde{D}_B^S \wedge \zeta$, and $T_{B\bigcup\{\triangle\}}^S = T_B^S \wedge \zeta$, and hence we see that it suffices to prove that the stochastic times D_B^S, $\widetilde{D}_B^S$, and T_B^S are $\overline{\mathcal{F}}_T^{S,\vee}$-measurable, whenever B is a Borel subset of E. $\square$

The proof of Theorem 4.7 is based on the following lemma. The same result with the same proof is also true with $E^{\triangle}$ instead of E.

Lemma 4.6. *Let $K \in \mathcal{K}(E)$ and $\tau \in [0,T)$. Suppose that $G_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$, is a sequence such that $K \subset \overline{G}_{n+1} \subset G_n$ and $\bigcap_{n\in\mathbb{N}} G_n = K$. Then the following assertions hold:*

(a) For every $\mu \in P(E)$, the sequence of stopping times $D_{G_n}^S$ increases and tends to D_K^S $\mathbb{P}_{\tau,\mu}$-almost surely.
(b) For every $t \in [\tau,T]$, the events $\left\{D_{G_n}^S \leqslant t\right\}$, $n \in \mathbb{N}$, are $\mathcal{F}_T^{S,\vee}$-measurable, and the event $\left\{D_K^S \leqslant t\right\}$ is $\overline{\mathcal{F}}_T^{S,\vee}$-measurable.
(c) For every $t \in [\tau,T]$, the events $\left\{T_{G_n}^S \leqslant t\right\}$, $n \in \mathbb{N}$, are $\mathcal{F}_T^{S,\vee}$-measurable, and the event $\left\{T_K^S \leqslant t\right\}$ is $\overline{\mathcal{F}}_T^{S,\vee}$-measurable.
(d) For every $\mu \in P(E)$, the sequence of stopping times $\widetilde{D}_{G_n}^S$ increases and tends to $\widetilde{D}_K^S$ $\mathbb{P}_{\tau,\mu}$-almost surely.
(e) For every $t \in [\tau,T]$, the events $\left\{\widetilde{D}_{G_n}^S \leqslant t\right\}$, $n \in \mathbb{N}$, are $\mathcal{F}_T^{S,\vee}$-measurable, and the event $\left\{\widetilde{D}_K^S \leqslant t\right\}$ is $\overline{\mathcal{F}}_T^{S,\vee}$-measurable.

Proof. (a). Fix $\mu \in P(E)$, and let $K \in \mathcal{K}(E)$ and $G_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$, be as in assertion (a) in the formulation of Lemma 4.6. Put $D_K = \sup_{n\in\mathbb{N}} D_{G_n}^S$. Since $S \leqslant D_{G_n}^S \leqslant D_{G_{n+1}}^S \leqslant D_K^S$, we always have $S \leqslant D_K \leqslant D_K^S$. Moreover, D_K is a stopping time. By using the quasi-continuity from the left of the process $t \mapsto X(t)$ on $[\tau,\zeta)$ with respect to the measure $\mathbb{P}_{\tau,\mu}$, we see that

$$\lim_{n\to\infty} X\left(D_{G_n}^S\right) = X\left(D_K\right) \quad \mathbb{P}_{\tau,\mu}\text{-almost surely on } \{D_K < \zeta\}.$$

Therefore,

$$X\left(D_K\right) \in \bigcap_{n\in\mathbb{N}} \overline{G}_n = K \quad \mathbb{P}_{\tau,\mu}\text{-almost surely on } \{D_K < \zeta\}. \tag{4.255}$$

Now by the definition of D_K^S we have $D_K^S \geqslant S$, and (4.255) implies $D_K^S \leqslant D_K$ $\mathbb{P}_{\tau,\mu}$-almost surely on $\{D_K < \zeta\}$, and hence $D_K^S = D_K$ $\mathbb{P}_{\tau,\mu}$-almost surely. In the final step we used the inequality $D_K \leqslant D_K^S$ which is always true.

(b). Fix $t \in [\tau,T)$ and $n \in \mathbb{N}$. By the right-continuity of paths on $[0,\zeta)$ we have

$$\left\{D_{G_n}^S \leqslant t < \zeta\right\} = \bigcap_{m\in\mathbb{N}} \left\{D_{G_n}^S < t + \frac{1}{m}\right\} \bigcap \{t < \zeta\}$$

$$= \bigcap_{m\in\mathbb{N}} \bigcup_{\rho\in[\tau,t+\frac{1}{m})} \{S \leqslant \rho,\, X(\rho)\in G_n\}$$
$$= \bigcap_{m\in\mathbb{N}} \bigcup_{\rho\in[\tau,t+\frac{1}{m})\bigcap\mathbb{Q}^+} \{S\vee\rho \leqslant \rho,\, X\left(S\vee\rho\right)\in G_n\}. \tag{4.256}$$

It follows that

$$\left\{D^S_{G_n} \leqslant t < \zeta\right\} \in \mathcal{F}^{S,\vee}_T,\ 0\leqslant t\leqslant T.$$

By using assertion (a), we see that the events

$$\left\{D^S_K \leqslant t < \zeta\right\} \quad \text{and} \quad \bigcap_{n\in\mathbb{N}} \left\{D^S_{G_n} \leqslant t < \zeta\right\}$$

coincide $\mathbb{P}_{\tau,\mu}$-almost surely. It follows that $\left\{D^S_K \leqslant t < \zeta\right\} \in \overline{\mathcal{F}}^{S,\vee}_T$. It also follows that the event $\left\{D^S_K < \zeta\right\}$ belongs to $\overline{\mathcal{F}}^{S,\vee}_T$. In addition we notice the equalities

$$\left\{D^S_K \leqslant t\right\} = \left\{D^S_K \leqslant t < \zeta\right\} \bigcup \left\{D^S_K \leqslant t,\, \zeta\leqslant t\right\}$$

($D^S_K \leqslant \zeta$ and $S\leqslant\zeta$)

$$= \left\{D^S_K \leqslant t < \zeta\right\} \bigcup \{\zeta \leqslant S\vee t\}$$
$$= \left\{D^S_K \leqslant t < \zeta\right\} \bigcup \{X\left(S\vee t\right) = \triangle\}. \tag{4.257}$$

From (4.257) we see that events of the form $\left\{D^S_K \leqslant t\right\}$, $t\in[\tau,T]$, belong to $\overline{\mathcal{F}}^{S,\vee}_T$. Consequently the stopping time D^S_K is $\overline{\mathcal{F}}^{S,\vee}_T$-measurable. This proves assertion (b).

(c). Since the sets G_n are open and the process $X(t)$ is right-continuous, the hitting times $T^S_{G_n}$ and the entry times $D^S_{G_n}$ coincide. Hence, the first part of assertion (c) follows from assertion (b). In order to prove the second part of (c), we reason as follows. By assertion (b), for every $r\in\mathbb{Q}^+$, the stopping time $D^{(r+S)\wedge\zeta}_K$ is $\overline{\mathcal{F}}^{(r+S)\wedge\zeta,\vee}_T$-measurable. Our next goal is to prove that for every $\varepsilon > 0$,

$$\mathcal{F}^{(\varepsilon+S)\wedge\zeta,\vee}_T \subset \mathcal{F}^{S,\vee}_T. \tag{4.258}$$

Fix $\varepsilon > 0$, and $\rho\in[\tau,\zeta]$, and put $S_1 = ((\varepsilon+S)\wedge\zeta)\vee\rho$. Observe that for $\rho,\, t\in[0,T]$, we have the following equality of events:

$$\{S_1\leqslant t\} = \{((\varepsilon+S)\wedge\zeta)\vee\rho\leqslant t\}$$
$$= \{((\varepsilon+S)\vee\rho)\wedge(\zeta\vee\rho)\leqslant t\}$$
$$= \{S\vee(\rho-\varepsilon)\leqslant t-\varepsilon,\, \rho\leqslant t\} \bigcup \{\zeta\leqslant S\vee t,\, \rho\leqslant t\}$$

$$= \{S \vee (\rho - \varepsilon) \leqslant t - \varepsilon,\, \rho \leqslant t\} \bigcup \{X(S \vee t) = \triangle,\, \rho \leqslant t\}. \tag{4.259}$$

Therefore, the stopping time $S_1 = ((\varepsilon + S) \wedge \zeta) \vee \rho$ is $\mathcal{F}_T^{S,\vee}$-measurable. Since the process $t \mapsto X(t)$ is right-continuous, it follows from Proposition 4.7 that $X(S_1)$ is $\mathcal{F}_T^{S,\vee}$-measurable. This implies inclusion (4.258). Hence,

$$\overline{\mathcal{F}}_T^{(\varepsilon+S)\wedge\zeta,\vee} \subset \overline{\mathcal{F}}_T^{S,\vee}, \tag{4.260}$$

and we see that the for every $\varepsilon > 0$ the stopping time $D_K^{(\varepsilon+S)\wedge\zeta}$ is $\overline{\mathcal{F}}_T^{S,\vee}$-measurable. Since the family $D_K^{(\varepsilon+S)\wedge\zeta}$, $\varepsilon > 0$, decreases to T_K^S, the hitting time T_K^S is $\overline{\mathcal{F}}_T^{S,\vee}$-measurable as well.

(d). Fix $\mu \in P(E)$, and let $K \in \mathcal{K}(E)$ and $G_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$, be as in assertion (a). Put $\widetilde{D}_K = \sup_{n\in\mathbb{N}} \widetilde{D}_{G_n}^S$. Since

$$\widetilde{D}_{G_n}^S \leqslant \widetilde{D}_{G_{n+1}}^S \leqslant \widetilde{D}_K^S,$$

we have $\widetilde{D}_K \leqslant \widetilde{D}_K^S$. It follows from the quasi-continuity from the left of the process $X(t)$ on $[0, \zeta)$ that

$$\lim_{n\to\infty} X\left(\widetilde{D}_{G_n}^S\right) = X\left(\widetilde{D}_K\right) \ \mathbb{P}_{\tau,\mu}\text{-almost surely on } \left\{\widetilde{D}_K < \zeta\right\}.$$

Therefore,

$$X\left(\widetilde{D}_K\right) \in \bigcap_n \overline{G}_n = K \ \mathbb{P}_{\tau,\mu}\text{-almost surely on } \left\{\widetilde{D}_K < \zeta\right\}.$$

Now $\widetilde{D}_K^S \geqslant S$ implies that $\widetilde{D}_K^S \leqslant \widetilde{D}_K$ $\mathbb{P}_{\tau,\mu}$-almost surely on $\left\{\widetilde{D}_K < \zeta\right\}$, and hence $\widetilde{D}_K^S = \widetilde{D}_K$ $\mathbb{P}_{\tau,\mu}$-almost surely on $\left\{\widetilde{D}_K < \zeta\right\}$. As in (a) we get $\widetilde{D}_K^S = \widetilde{D}_K$ $\mathbb{P}_{\tau,\mu}$-almost surely.

(e). Fix $t \in [\tau, T)$ and $n \in \mathbb{N}$. By the right-continuity of paths,

$$\begin{aligned}\left\{\widetilde{D}_{G_n}^S \leqslant t < \zeta\right\} &= \bigcap_{m\in\mathbb{N}} \left\{D_{G_n}^S < t + \frac{1}{m}\right\} \bigcap \{t < \zeta\} \\ &= \bigcap_{m\in\mathbb{N}} \bigcup_{\rho\in(\tau, t+\frac{1}{m})} \{S \leqslant \rho,\, X(\rho) \in G_n\} \\ &= \bigcap_{m\in\mathbb{N}} \bigcup_{\rho\in(\tau, t+\frac{1}{m})\bigcap\mathbb{Q}^+} \{S \vee \rho \leqslant \rho,\, X(S \vee \rho) \in G_n\}.\end{aligned} \tag{4.261}$$

It follows that $\left\{\widetilde{D}_{G_n}^S \leqslant t < \zeta\right\} \in \mathcal{F}_T^{S,\vee}$. By using assertion (d), we see that the events $\left\{\widetilde{D}_K^S \leqslant t < \zeta\right\}$ and $\bigcap_{n\in\mathbb{N}} \left\{\widetilde{D}_{G_n}^S \leqslant t < \zeta\right\}$ coincide $\mathbb{P}_{\tau,\mu}$-almost surely. Therefore, $\left\{\widetilde{D}_K^S \leqslant t < \zeta\right\} \in \overline{\mathcal{F}}_T^{S,\vee}$. As in (4.257) we have

$$\left\{\widetilde{D}_K^S \leqslant t\right\} = \left\{\widetilde{D}_K^S \leqslant t < \zeta\right\} \bigcup \left\{\widetilde{D}_K^S \leqslant t,\, \zeta \leqslant t\right\}$$

$$= \left\{ \widetilde{D}_K^S \leqslant t < \zeta \right\} \bigcup \left\{ X\left(S \vee t\right) = \triangle \right\}. \tag{4.262}$$

This proves assertion (e); therefore the proof of Lemma 4.6 is complete. □

Proof. [Proof of Theorem 4.7: continuation] Let us return to the proof of Theorem 4.7. We will first prove that for any Borel set B, the entry time D_B^S is measurable with respect to the σ-field $\overline{\mathcal{F}}_T^{S,\vee}$. Then the same assertion holds for the hitting time T_B^S. Indeed, if D_B^S is $\overline{\mathcal{F}}_T^{S,\vee}$-measurable for all stopping times S, then for every $\varepsilon > 0$, the stopping time $D_B^{(\varepsilon+S)\wedge\zeta}$ is measurable with respect to the σ-field $\overline{\mathcal{F}}_T^{(\varepsilon+S)\wedge\zeta,\vee}$. By using (4.260), we obtain the $\overline{\mathcal{F}}_T^{S,\vee}$-measurability of $D_B^{(\varepsilon+S)\wedge\zeta}$. Now (4.236) implies the $\overline{\mathcal{F}}_T^{S,\vee}$-measurability of T_B^S.

Fix $t \in [\tau, T)$, $\mu \in P(E)$, and $B \in \mathcal{E}$. By Lemma 4.3, the set B is capacitable with respect to the capacity $K \mapsto \mathbb{P}_{\tau,\mu}\left[D_K^S \leqslant t\right]$. Notice that the following argument was also employed in the proof of Theorem 4.6. Therefore, there exists an increasing sequence $K_n \in \mathcal{K}(E)$, $n \in \mathbb{N}$, and a decreasing sequence $G_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$, such that

$$\begin{aligned} &K_n \subset K_{n+1} \subset B \subset G_{n+1} \subset G_n, \quad n \in \mathbb{N}, \quad \text{and} \\ &\sup_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{K_n}^S \leqslant t\right] = \inf_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{G_n}^S \leqslant t\right]. \end{aligned} \tag{4.263}$$

Next we put

$$\Lambda_1^{\tau,\mu,S}(t) = \bigcup_{n\in\mathbb{N}} \left\{D_{K_n}^S \leqslant t\right\} \quad \text{and} \quad \Lambda_2^{\tau,\mu,S}(t) = \bigcap_{n\in\mathbb{N}} \left\{D_{G_n}^S \leqslant t\right\}. \tag{4.264}$$

The equalities in (4.249) which are the same as those in (4.264) show that the events $\Lambda_1^{\tau,\mu,S}(t)$ and $\Lambda_2^{\tau,\mu,S}(t)$ are $\overline{\mathcal{F}}_T^{S,\vee}$-measurable. Moreover, we have

$$\Lambda_1^{\tau,\mu,S}(t) \subset \left\{D_B^S \leqslant t\right\} \subset \Lambda_2^{\tau,\mu,S}(t), \tag{4.265}$$

and

$$\begin{aligned} \mathbb{P}_{\tau,\mu}\left[\Lambda_2^{\tau,\mu,S}(t)\right] &= \inf_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{G_n}^S \leqslant t\right] \\ &= \sup_{n\in\mathbb{N}} \mathbb{P}_{\tau,\mu}\left[D_{K_n}^S \leqslant t\right] = \mathbb{P}_{\tau,\mu}\left[\Lambda_1^{\tau,\mu,S}(t)\right]. \end{aligned} \tag{4.266}$$

Now (4.265) and (4.266) give $\mathbb{P}_{\tau,\mu}\left[\Lambda_2^{\tau,\mu,S}(t)\backslash\Lambda_1^{\tau,\mu,S}(t)\right] = 0$. By using (4.265), we see that the event $\left\{D_B^S \leqslant t\right\}$ is measurable with respect to the σ-field $\overline{\mathcal{F}}_T^{S,\vee}$. This establishes the $\overline{\mathcal{F}}_T^{S,\vee}$-measurability of the entry time D_B^S and the hitting time T_B^S. The proof of Theorem 4.7 for the pseudo-hitting time $\widetilde{D}_B^S$ is similar to that for the entry time D_B^S.

The proof of Theorem 4.7 is thus completed. □

Definition 4.11. Fix $\tau \in [0,T]$, and let $S_1 : \Omega \to [\tau, T]$ be an $\left(\mathcal{F}_t^\tau\right)_{t\in[\tau,T]}$-stopping time. A stopping time $S_2 : \Omega \to [\tau, T]$ is called terminal after S_1 if $S_2 \geqslant S_1$, and if S_2 is $\overline{\mathcal{F}}_T^{S_1,\vee}$-measurable.

The following corollary shows that entry and hitting times of Borel subsets which are comparable are terminal after each other.

Corollary 4.4. *Let $(X(t), \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x})$ be a standard process, and let A and B be Borel subsets of E with $B \subset A$. Then the entry time D_B^τ is measurable with respect to the σ-field $\overline{\mathcal{F}}_T^{D_A^\tau,\vee}$. Moreover, the hitting time T_B^τ is measurable with respect to the σ-field $\overline{\mathcal{F}}_T^{T_A^\tau,\vee}$.*

Proof. By Theorem 4.7, it suffices to show that the equalities

$$D_B^{D_A^\tau} = D_B^\tau \text{ and } \widetilde{D}_B^{T_A^\tau} = T_B^\tau \tag{4.267}$$

hold $\mathbb{P}_{\tau,\mu}$-almost surely for all $\mu \in P(E)$. The first equality in (4.267) follows from

$$\bigcup_{\tau\leqslant s<T} \{D_A^\tau \leqslant s,\, X(s) \in B\} = \bigcup_{\tau\leqslant s<T} \{X(s) \in B\},$$

while the second equality in (4.267) can be obtained from

$$\bigcup_{\tau< s<T} \{T_A^\tau \leqslant s,\, X(s) \in B\} = \bigcup_{\tau< s<T} \{X(s) \in B\}.$$

This proves Corollary 4.4. □

It follows from Corollary 4.4 that the families $\{D_A^\tau : A \in \mathcal{E}\}$ and $\{T_A^\tau : A \in \mathcal{E}\}$ can be used in the definition of the strong Markov property in the case of standard processes. The next theorem states that the strong Markov property holds for entry times and hitting times of comparable Borel subsets.

Theorem 4.8. *Let $(X(t), \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x})$ be a standard process, and fix $\tau \in [0,T]$. Let A and B be Borel subsets of E such that $B \subset A$, and let $f : [\tau, T] \times E^\triangle \to \mathbb{R}$ be a bounded Borel function. Then the following equalities hold $\mathbb{P}_{\tau,x}$-almost surely:*

$$\mathbb{E}_{\tau,x}\left[f\left(D_B^\tau, X\left(D_B^\tau\right)\right) \bigm| \mathcal{F}_{D_A^\tau}^\tau\right] = \mathbb{E}_{D_A^\tau, X\left(D_A^\tau\right)}\left[f\left(D_B^\tau, X\left(D_B^\tau\right)\right)\right] \qquad \textit{and}$$

$$\mathbb{E}_{\tau,x}\left[f\left(T_B^\tau, X\left(T_B^\tau\right)\right) \bigm| \mathcal{F}_{T_A^\tau}^\tau\right] = \mathbb{E}_{T_A^\tau, X\left(T_A^\tau\right)}\left[f\left(T_B^\tau, X\left(T_B^\tau\right)\right)\right]$$

The first one holds $\mathbb{P}_{\tau,x}$-almost surely on $\left\{D_A^S < \zeta\right\}$, and the second on $\left\{T_A^S < \zeta\right\}$.

Proof. Theorem 4.8 follows from Corollary 4.4 and Remark 4.4. □

Definition 4.12. The quadruple

$$\left\{\left(\Omega, \left(\overline{\mathcal{F}}_{t+}^{\tau}\right)_{t\in[\tau,T]}, \mathbb{P}_{\tau,x}\right), (X(t), t \in [0,T]), (\vee_t, t \in [0,T]), (E, \mathcal{E})\right\} \tag{4.268}$$

is called a standard Markov process if it possesses the following properties:

(1) The process $X(t)$ is adapted to the filtration $\left(\overline{\mathcal{F}}_{t+}^{\tau}\right)_{t\in[\tau,T]}$, right-continuous and possesses left limits in E on its life time.
(2) The σ-fields $\overline{\mathcal{F}}_{t+}^{\tau}$, $t \in [\tau,T]$, are right continuous and $\mathbb{P}_{\tau,x}$-complete.
(3) The process $(X(t) : t \in [0,T])$ is strong Markov with respect to the measures $\{\mathbb{P}_{\tau,x} : (\tau,x) \in [0,T] \times E\}$
(4) The process $(X(t) : t \in [0,T])$ is quasi left-continuous on $[0,\zeta)$.
(5) The equalities $X(t) \circ \vee_s = X(t \vee s)$ hold $\mathbb{P}_{\tau,x}$-almost surely for all $(\tau,x) \in [0,T] \times E$ and for s, $t \in [\tau,T]$.

If $\Omega = D\left([0,T], E^{\triangle}\right)$ and $X(t)(\omega) = \omega(t)$, $t \in [0,T]$, $\omega \in \Omega$, then parts of the items (1) and (2) are automatically satisfied. For brevity we often write $(X(t), \mathbb{P}_{\tau,x})$ instead of (4.268).

The following proposition gives an alternative way to describe stopping times which are terminal after another stopping time: see Definition 4.11.

Proposition 4.7. *Let $S_1 : \Omega \to [\tau,T]$ be an $\mathcal{F}_t^{\tau}$-stopping time, and let the stopping $S_2 : \Omega \to [\tau,T]$ be such that $S_2 \geqslant S_1$, and such that for every $t \in [\tau,T]$ the event $\{S_2 > t\}$ restricted to the event $\{S_1 < t\} = \{S_1 \vee t < t\}$ only depends on $\mathcal{F}_T^t$. Then S_2 is $\mathcal{F}_T^{S_1,\vee}$-measurable. If the paths of the process X are right-continuous, the state variable $X(S_2)$ is $\mathcal{F}_T^{S_1,\vee}$-measurable as well. It follows that the space-time variable $(S_2, X(S_2))$ is $\mathcal{F}_T^{S_1,\vee}$-measurable. Similar results are true if the σ-fields $\mathcal{F}_T^t$ and $\mathcal{F}_T^{S_1,\vee}$ by their $\mathbb{P}_{\tau,\mu}$-completions for some probability measure μ on $\mathcal{E}$.*

Proof. Suppose that for every $t \in [\tau,T]$ the random variable S_2 is such that on $\{S_1 < t\} = \{S_1 \vee t < t\}$ the event $\{S_2 > t\}$ only depends on $\mathcal{F}_T^t$. Then on $\{S_1 < t\}$ the event $\{S_2 > t\}$ only depends on the σ-field generated by the state variables $\left\{X(\rho)\restriction_{\{S_1\vee t<t\}} : \rho \geqslant t\right\} = \left\{X(\rho \vee S_1)\restriction_{\{S_1\vee t<t\}} : \rho \geqslant t\right\}$. Consequently, the event $\{S_2 > t > S_1\}$ is $\mathcal{F}_T^{S_1,\vee}$-measurable. Since $S_2 = S_1 + \int_{\tau}^{T} \mathbf{1}_{\{S_2>t>S_1\}}dt$, we see that S_2 is $\mathcal{F}_T^{S_1,\vee}$-measurable. This argument can be adapted if we only know that for every $t \in [\tau, T]$ on the event $\{S_1 < t\}$ the event $\{S_2 > t\}$ only depends on the $\mathbb{P}_{\tau,\mu}$-completion of the σ-field generated by the state variables $\left\{X(\rho)\restriction_{\{S_1 \vee t < t\}} : \rho \geqslant t\right\}$ for some probability measure μ on $\mathcal{E}$.

If the process $X(t)$ is right-continuous, and if S_2 is a stopping time which is terminal after the stopping time $S_1 : \Omega \to [0, T]$, then the space-time variable $(S_2, X(S_2))$ is $\overline{\mathcal{F}}_T^{S_1,\vee}$-measurable. This result follows from the equality in (3.46) with S_2 instead of S:

$$S_{2,n}(t) = \tau + \frac{t-\tau}{2^n}\left\lceil \frac{2^n (S_2 - \tau)}{t-\tau} \right\rceil. \tag{4.269}$$

Then notice that the stopping times $S_{2,n}(t)$, $n \in \mathbb{N}$, $t \in (\tau, T]$, are $\mathcal{F}_T^{S_1,\vee}$-measurable, provided that S_2 has this property. Moreover, we have $S_2 \leqslant S_{2,n+1}(t) \leqslant S_{2,n}(t) \leqslant S_2 + 2^{-n}(t-\tau)$. It follows that the state variables $X(S_{2,n}(t))$, $n \in \mathbb{N}$, $t \in (\tau, T]$, are $\mathcal{F}_T^{S_1,\vee}$-measurable, and that the same is true for $X(S_2) = \lim_{n\to\infty} X(S_{2,n}(t))$.

This completes the proof of Proposition 4.7. □

Remark 4.4. If in Theorem 4.9 for the sample path space Ω we take the Skorohod space $\Omega = D\left([0,T], E^{\triangle}\right)$, $X(t)(\omega) = \omega(t)$, $\omega \in \Omega$, $t \in [0, T]$, then the process $t \mapsto X(t)$, $t \in [0, T]$, is right-continuous, has left limits in E on its life time, and is quasi-left continuous on its life time as well.

Theorem 4.9. *Let, like in Lemma 4.3,*

$$\left\{\left(\Omega, \overline{\mathcal{F}}_T^{\tau}, \mathbb{P}_{\tau,x}\right), (X(t), t \in [0,T]), (\vee_t, t \in [\tau, T]), (E, \mathcal{E})\right\}$$

be a standard Markov process with right-continuous paths, which has left limits on its life time, and is quasi-continuous from the left on its life time. For fixed $(\tau, x) \in [0, T] \times E$, the σ-field $\overline{\mathcal{F}}_T^{S,\vee}$ is the completion of the σ-field $\mathcal{F}_T^{S,\vee} = \sigma(S \vee \rho, X(S \vee \rho) : 0 \leqslant \rho \leqslant T)$ with respect to the measure $\mathbb{P}_{\tau,x}$. Then, if (S_1, S_2) is a pair of stopping times such that S_2 is $\overline{\mathcal{F}}_T^{S_1,\vee}$-measurable and $\tau \leqslant S_1 \leqslant S_2 \leqslant T$, then for all bounded Borel functions f on $[\tau, T] \times E^{\triangle}$, the equality

$$\mathbb{E}_{\tau,x}\left[f(S_2, X(S_2)) \mid \overline{\mathcal{F}}_{S_1}^{\tau}\right] = \mathbb{E}_{S_1, X(S_1)}\left[f(S_2, X(S_2))\right] \tag{4.270}$$

holds $\mathbb{P}_{\tau,x}$-almost surely on $\{S_1 < \zeta\}$.

First notice that the conditions on S_1 and S_2 are such that S_2 is terminal after S_1: see Definition 4.11. Also observe that the Markov process in (4.268) is quasi-continuous from the left on its life time $[0, \zeta)$: compare with Theorem 2.9. Let A and B be Borel subsets of E such that $B \subset A$.

In (4.270) we may put $S_1 = D_A^{\tau}$ together with $S_2 = D_B^{\tau}$, or $S_1 = T_A^{\tau}$ and $S_2 = T_B^{\tau}$: see Theorem 4.7 and Corollary 4.4.

Proof. The result in Theorem 4.9 is a consequence of the strong Markov property as exhibited in Theorem 2.9. □

Remark 4.5. This remark is concerned with the concept of λ-dominance. Without the sequential λ-dominance of the operator $D_1 + L$ the second formula, i.e. the formula in (4.116), poses a difficulty as far as it is not clear that the function $e^{-\lambda t}\widetilde{S}_0(t)f$ belongs to $C_b([0,T]\times E)$ indeed. For the moment suppose that the function $f \in C_b([0,T]\times E)$ is such that $\widetilde{S}_0(t)f \in C_b([0,T]\times E)$. Then equality (4.115) yields:

$$\begin{aligned}\left(\lambda I - L^{(1)}\right)^{-1} f &= \int_0^t e^{-\lambda\rho}\widetilde{S}_0(\rho)f\,d\rho + e^{-\lambda t}S_0(t)\left(\lambda I - L^{(1)}\right)^{-1} f \\ &= \int_0^t e^{-\lambda\rho}\widetilde{S}_0(\rho)f\,d\rho + e^{-\lambda t}\left(\lambda I - L^{(1)}\right)^{-1}\widetilde{S}_0(t)f.\end{aligned} \tag{4.271}$$

Consequently, the function $\int_0^t e^{-\lambda\rho}\widetilde{S}_0(\rho)f\,d\rho$ belongs to $D\left(L^{(1)}\right)$ and the equality in (4.116) follows from (4.271). For the relation between λ-dominant operators, λ-super-mean, and λ-supermedian functions see Remark 4.1.

Moreover, we have

$$C_{P,b}^{(1)} = \bigcap_{\lambda_0>0} C_{P,b}^{(1)}(\lambda_0) = \bigcap_{\lambda_0>0}\left\{\left(\lambda_0 I - \overline{L} - D_1\right)g :\, g \in D\left(\overline{L}\right)\bigcap D\left(D_1\right)\right\}. \tag{4.272}$$

The second equality in (4.272) follows from (4.91) and (4.92).

4.5.1 *Some related remarks*

In subsection 3.1.6 we already discussed to some length topics related to Korovkin families and convergence properties of measures. Here we will say something about the maximum principle, the martingale problem, and stopping time arguments.

We notice that we have used the following version of the Choquet capacity theorem for Borel subsets instead of analytic subsets. As is well-known Borel subsets are analytic.

Theorem 4.10. *In a Polish space every analytic set is capacitable.*

For more general versions of our Choquet capacibility theory and capacitable subsets see e.g. [Kiselman (2000)], and, of course, Choquet [Choquet (1986)], Benzecri [Benzécri (1995)] and Dellacherie and Meyer [Dellacherie and Meyer (1978)] as well. For a general discussion on the foundations of probability theory see e.g. [Kallenberg (2002)]. In [Gulisashvili and van Casteren (2006)] the authors also made a thorough investigation of measurability properties of stopping times. However, in that case the underlying state space was locally compact. In [van Casteren (1992)] the author makes an extensive study of the maximum principle of an unbounded operator with domain and range in the space of continuous functions which vanish at infinity where the state space is locally compact. As indicated in Chapter 1 an operator L for which the martingale problem is well-posed need possess a unique extension which is the generator of a Dynkin-Feller semigroup. As indicated by Kolokoltsov in [Kolokoltsov (2004b)] there exist relatively easy counter-examples: see comments after Theorems 2.9 through 2.13 in §2.3. For the time-homogeneous case see, e.g., [Ethier and Kurtz (1986)] or [Ikeda and Watanabe (1998)]. In fact [Ethier and Kurtz (1986)] contains a general result on operators with domain and range in $C_0(E)$ and which have unique linear extensions generating a Feller-Dynkin semigroup. The martingale problem goes back to Stroock and Varadhan (see [Stroock and Varadhan (1979)]). It found numerous applications in various fields of mathematics. We refer the reader to [Liggett (2005)], [Kolokoltsov (2004b)], and [Kolokoltsov (2004a)] for more information about and applications of the martingale problem. In [Eberle (1999)] the reader may find singular diffusion equations which possess or which do not possess unique solutions. Consequently, for (singular) diffusion equations without unique solutions the martingale problem is not uniquely solvable. Another valuable source of information is [Jacob (2001, 2002, 2005)]. Other relevant references are papers by Hoh [Hoh (1994, 1995b,a, 2000)]. Some of Hoh's work is also employed in Jacob's books. In fact most of these references discuss the relations between pseudo-differential operators (of order less than or equal to 2), the corresponding martingale problem, and being the generator of a Feller-Dynkin semigroup.

PART 3

Backward Stochastic Differential Equations

This page is intentionally left blank

Chapter 5

Feynman-Kac formulas, backward stochastic differential equations and Markov processes

In this chapter we explain the notion of stochastic backward differential equations and its relationship with classical (backward) parabolic differential equations of second order. The chapter contains a mixture of stochastic processes like Markov processes and martingale theory and semi-linear partial differential equations of parabolic type. Some emphasis is put on the fact that the whole theory generalizes Feynman-Kac formulas: see e.g. Remark 5.4 and formula (5.33). A new method of proof of the existence of solutions is given: see equality (5.83) and Proposition 5.7.

In the literature functions with the monotonicity property are also called one-sided Lipschitz functions. In fact Theorem 5.2, with $f(t,x,\cdot,\cdot)$ Lipschitz continuous in both variables, will be superseded by Theorem 5.4 in the Lipschitz case and by Theorem 5.5 in case of monotonicity in the second variable and Lipschitz continuity in the third variable. The proof of Theorem 5.2 is part of the results in Section 5.3. Theorem 5.7 contains a corresponding result for a Markov family of probability measures. Its proof is omitted, it follows the same lines as the proof of Theorem 5.5.

All the existence arguments are based on rather precise quantitative estimates.

Unless specified otherwise all (local) martingales in this chapter and in Chapters 6 and 7 are almost surely continuous. As a consequence for such martingales we have a standard Itô calculus and stochastic integrals relative to local martingales are again local martingales. For details on this see e.g. [Williams (1991)].

5.1 Introduction

This introduction serves as a motivation for the present chapter and also for Chapter 6. Backward stochastic differential equations, in short BSDE's, have been well studied during the last ten years or so. They were introduced by Pardoux and Peng [Pardoux and Peng (1990)], who proved existence and uniqueness of adapted solutions, under suitable square-integrability assumptions on the coefficients and on the terminal condition. They provide probabilistic formulas for solution of systems of semi-linear partial differential equations, both of parabolic and elliptic type. The interest for this kind of stochastic equations has increased steadily, this is due to the strong connections of these equations with mathematical finance and the fact that they gave a generalization of the well known Feynman-Kac formula to semi-linear partial differential equations. In the present chpter we will concentrate on the relationship between time-dependent strong Markov processes and abstract backward stochastic differential equations. The equations are phrased in terms of a martingale problem, rather than a stochastic differential equation. They could be called weak backward stochastic differential equations. Emphasis is put on existence and uniqueness of solutions. The paper [Van Casteren (2009)] deals with the same subject, but it concentrates on comparison theorems and viscosity solutions. The proof of the existence result is based on a theorem which is related to a homotopy argument as pointed out by the authors of [Crouzeix *et al.* (1983)]. It is more direct than the usual approach, which uses, among other things, regularizing by convolution products. It also gives rather precise quantitative estimates. In [Van Casteren (2010)] the author extends the results on BSDE's to the Hilbert space setting.

For examples of strong solutions which are driven by Brownian motion the reader is referred to e.g. section 2 in [Pardoux (1998a)]. If the coefficients $x \mapsto b(s,x)$ and $x \mapsto \sigma(s,x)$ of the underlying (forward) stochastic differential equation are linear in x, then the corresponding forward-backward stochastic differential equation is related to option pricing in financial mathematics. The backward stochastic differential equation may serve as a model for a hedging strategy. For more details on this interpretation see e.g. [El Karoui and Quenez (1997)], pp. 198–199. A rather recent book on financial mathematics in terms of martingale theory is the one by Delbaen and Schachermeyer [Delbaen and Schacher mayer (2006)]. E. Pardoux and S. Zhang [Pardoux and Zhang (1998)] use BSDE's to give a probabilistic formula for the solution of a system of parabolic or elliptic semi-linear partial differential equation with Neumann boundary condition. For recent results on forward-backward stochastic differential equations using a martingale approach the reader is referred to [Ma *et al.* (2008)]. In [Boufoussi and van Casteren (2004b)] the authors also put BSDE's at work to prove a result on a Neumann type boundary problem.

In this chapter we want to consider the situation where the family of operators $L(s)$, $0 \leqslant s \leqslant T$, generates a time-inhomogeneous Markov process

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), \left(E, \mathcal{E}\right)\right\} \tag{5.1}$$

in the sense that

$$\frac{d}{ds}\mathbb{E}_{\tau,x}\left[f\left(X(s)\right)\right] = \mathbb{E}_{\tau,x}\left[L(s) f\left(X(s)\right)\right], \quad f \in D\left(L(s)\right), \ \tau \leqslant s \leqslant T.$$

We consider the operators $L(s)$ as operators on (a subspace of) the space of bounded continuous functions on E, i.e. on $C_b(E)$ equipped with the supremum norm: $\|f\|_\infty = \sup_{x\in E}|f(x)|$, $f \in C_b(E)$, and the strict topology $\mathcal{T}_\beta$. With the operators $L(s)$ we associate the squared gradient operator Γ_1 defined by

$$\begin{aligned}&\Gamma_1\left(f,g\right)\left(\tau,x\right)\\&= \mathcal{T}_\beta\text{-}\lim_{s\downarrow\tau}\frac{1}{s-\tau}\mathbb{E}_{\tau,x}\left[\left(f\left(X(s)\right) - f\left(X(\tau)\right)\right)\left(g\left(X(s)\right) - g\left(X(\tau)\right)\right)\right],\end{aligned} \tag{5.2}$$

for f, $g \in D\left(\Gamma_1\right)$. Here $D\left(\Gamma_1\right)$ is the domain of the operator Γ_1. It consists of those functions $f \in C_b(E) = C_b\left(E, \mathbb{C}\right)$ with the property that the strict limit

$$\mathcal{T}_\beta\text{-}\lim_{s\downarrow\tau}\frac{1}{s-\tau}\mathbb{E}_{\tau,x}\left[\left(f\left(X(s)\right) - f\left(X(\tau)\right)\right)\left(\overline{f\left(X(s)\right)} - \overline{f\left(X(\tau)\right)}\right)\right] \tag{5.3}$$

exists. We will assume that $D\left(\Gamma_1\right)$ contains an algebra of functions in $C_b\left([0,T]\times E\right)$ which is closed under complex conjugation, and which is $\mathcal{T}_\beta$-dense. These squared gradient operators are also called energy operators: see e.g. Barlow, Bass and Kumagai [Barlow *et al.* (2005)]. We assume that every operator $L(s)$, $0 \leqslant s \leqslant T$, generates a diffusion in the sense of the following definition. In the sequel it is assumed that the family of operators $\{L(s) : 0 \leqslant s \leqslant T\}$ possesses the property that the space of functions $u : [0,T]\times E \to \mathbb{R}$ with the property that the function $(s,x) \mapsto \frac{\partial u}{\partial s}(s,x) + L(s)u\left(s,\cdot\right)(x)$ belongs to $C_b\left([0,T]\times E\right) := C_b\left([0,T]\times E; \mathbb{C}\right)$ is $\mathcal{T}_\beta$-dense in the space $C_b\left([0,T]\times E\right)$. This subspace of functions is denoted by $D(L)$, and the operator L is defined by $Lu(s,x) = L(s)u\left(s,\cdot\right)(x)$, $u \in D(L)$. It is also assumed that the family $\mathcal{A}$ is a core for the operator L. We assume that the operator L, or that the family of operators

$\{L(s) : 0 \leqslant s \leqslant T\}$, generates a diffusion in the sense of the following definition. It is assumed that the constant function $\mathbf{1}$ belongs to $D(L(s))$, $s \in [0,T]$, and that $L(s)\mathbf{1} = 0$.

Definition 5.1. A family of operators $\{L(s) : 0 \leqslant s \leqslant T\}$ is said to generate a *diffusion* if for every C^∞-function $\Phi : \mathbb{R}^n \to \mathbb{R}$, and every pair $(s,x) \in [0,T] \times E$ the following identity is valid

$$\begin{aligned} & L(s)\left(\Phi\left(f_1,\dots,f_n\right)(s,\cdot)\right)(x) \\ &= \sum_{j=1}^n \frac{\partial \Phi}{\partial x_j}\left(f_1(s,x),\dots,f_n(s,x)\right) L(s) f_j(s,x) \\ &\quad + \frac{1}{2}\sum_{j,k=1}^n \frac{\partial^2 \Phi}{\partial x_j \partial x_k}\left(f_1(s,x),\dots,f_n(s,x)\right)\Gamma_1\left(f_j,f_k\right)(s,x) \end{aligned} \tag{5.4}$$

for all functions $f_1,\dots,f_n$ in an algebra of functions $\mathcal{A}$, contained in the domain of the operator L, which forms a core for L.

Generators of diffusions for single operators are described in Bakry's lecture notes [Bakry (1994)]. For more information on the squared gradient operator see e.g. [Bakry and Ledoux (2006)] and [Bakry (2006)] as well. Put $\Phi(f,g) = fg$. Then (5.4) implies

$$L(s)(fg)(s,\cdot)(x) = L(s)f(s,\cdot)(x)g(s,x) + f(s,x)L(s)g(s,\cdot)(x) + \Gamma_1(f,g)(s,x),$$

provided that the three functions f, g and fg belong to $\mathcal{A}$. Instead of using the full strength of (5.4), i.e. with a general function Φ, we just need it for the product $(f,g) \mapsto fg$: see Proposition 5.4.

Remark 5.1. Let m be a reference measure on the Borel field $\mathcal{E}$ of E, and let $p \in [1,\infty]$. If we consider the operators $L(s)$, $0 \leqslant s \leqslant T$, in $L^p(E,\mathcal{E},m)$-space, then we also need some conditions on the algebra $\mathcal{A}$ of "core" type in the space $L^p(E,\mathcal{E},m)$. For details the reader is referred to [Bakry (1994)].

By definition the gradient of a function $u \in D(\Gamma_1)$ in the direction of (the gradient of) $v \in D(\Gamma_1)$ is the function $(\tau,x) \mapsto \Gamma_1(u,v)(\tau,x)$. For given $(\tau,x) \in [0,T] \times E$ the functional $v \mapsto \Gamma_1(u,v)(\tau,x)$ is linear: its action is denoted by $\nabla_u^L(\tau,x)$. Hence, for $(\tau,x) \in [0,T] \times E$ fixed, we can consider $\nabla_u^L(\tau,x)$ as an element in the dual of $D(\Gamma_1)$. The pair

$$(\tau,x) \mapsto \left(u(\tau,x), \nabla_u^L(\tau,x)\right)$$

may be called an element in the phase space of the family $L(s)$, $0 \leqslant s \leqslant T$, (see Jan Prüss [Prüss (2002)]), and the process $s \mapsto$

$\left(u\left(s, X(s)\right), \nabla_u^L\left(s, X(s)\right)\right)$ will be called an element of the stochastic phase space. Next let $f : [0,T] \times E \times \mathbb{R} \times D\left(\Gamma_1\right)^* \to \mathbb{R}$ be a "reasonable" function, and consider, for $0 \leqslant s_1 < s_2 \leqslant T$ the expression:

$$\begin{aligned} & u\left(s_2, X\left(s_2\right)\right) - u\left(s_1, X\left(s_1\right)\right) + \int_{s_1}^{s_2} f\left(s, X(s), u\left(s, X(s)\right), \nabla_u^L\left(s, X(s)\right)\right) ds \\ & \quad - u(s_2, X(s_2)) + u(s_1, X(s_1)) + \int_{s_1}^{s_2} \left(L(s)u(s, X(s)) + \frac{\partial u}{\partial s}(s, X(s))\right) ds \end{aligned} \tag{5.5}$$

$$\begin{aligned} & = u(s_2, X(s_2)) - u(s_1, X(s_1)) + \int_{s_1}^{s_2} f(s, X(s), u(s, X(s)), \nabla_u^L(s, X(s)))ds \\ & \quad - M_u\left(s_2\right) + M_u\left(s_1\right), \end{aligned} \tag{5.6}$$

where

$$\begin{aligned} & M_u\left(s_2\right) - M_u\left(s_1\right) \\ & = u\left(s_2, X\left(s_2\right)\right) - u\left(s_1, X\left(s_1\right)\right) - \int_{s_1}^{s_2} \left(L(s)u\left(s, X(s)\right) + \frac{\partial u}{\partial s}\left(s, X(s)\right)\right) ds \\ & = \int_{s_1}^{s_2} dM_u(s). \end{aligned} \tag{5.7}$$

Details on the properties of the function f will be given in the theorems 5.2, 5.3, 5.4, 5.5, and 5.7.

The following definition also occurs in Definition 2.6, where the reader will find more details about Definitions 5.2 and 5.3. It also explains the relationship with transition probabilities and Feller propagators.

Definition 5.2. The process

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), \left(E, \mathcal{E}\right)\right\} \tag{5.8}$$

is called a time-inhomogeneous Markov process if

$$\mathbb{E}_{\tau,x}\left[f(X(t)) \,\middle|\, \mathcal{F}_s^\tau\right] = \mathbb{E}_{s,X(s)}\left[f(X(t))\right], \quad \mathbb{P}_{\tau,x}\text{-almost surely.} \tag{5.9}$$

Here f is a bounded Borel measurable function defined on the state space E and $\tau \leqslant s \leqslant t \leqslant T$.

Suppose that the process $X(t)$ in (5.8) has paths which are right-continuous and have left limits in E. Then it can be shown that the Markov property for fixed times carries over to stopping times in the sense that (5.9) may be replaced with

$$\mathbb{E}_{\tau,x}\left[Y \,\middle|\, \mathcal{F}_S^\tau\right] = \mathbb{E}_{S,X(S)}\left[Y\right], \quad \mathbb{P}_{\tau,x}\text{-almost surely.} \tag{5.10}$$

Here $S : E \to [\tau, T]$ is an $\mathcal{F}_t^\tau$-adapted stopping time and Y is a bounded random variable which is measurable with respect to the future (or terminal) σ-field after S, i.e. the one generated by $\{X(t \vee S) : \tau \leqslant t \leqslant T\}$. For this type of result the reader is referred to Chapter 2 in [Gulisashvili and van Casteren (2006)] and to Theorem 2.9. Markov processes for which (5.10) holds are called strong Markov processes. For more details on hitting times see §4.5.

The following definition is, essentially speaking, the same as Definition 2.8. Its relationship with Feller propagators or evolutions (see Chapter 2, Definition 2.7) is explained in Proposition 4.1 in Chapter 4. The derivatives and the operators $L(s)$, $s \in [0, T]$, have to be taken with respect to the strict topology: see Section 2.1.

Definition 5.3. The family of operators $L(s)$, $0 \leqslant s \leqslant T$, is said to generate a time-inhomogeneous Markov process

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), (E, \mathcal{E})\right\} \tag{5.11}$$

if for all functions $u \in D(L)$, for all $x \in E$, and for all pairs (τ, s) with $0 \leqslant \tau \leqslant s \leqslant T$ the following equality holds:

$$\frac{d}{ds}\mathbb{E}_{\tau,x}\left[u\left(s, X(s)\right)\right] = \mathbb{E}_{\tau,x}\left[\frac{\partial u}{\partial s}(s, X(s)) + L(s)u\left(s, \cdot\right)\left(X(s)\right)\right]. \tag{5.12}$$

Next we show that under rather general conditions the process $s \mapsto M_u(s) - M_u(t)$, $t \leqslant s \leqslant T$, as defined in (5.6) is a $\mathbb{P}_{t,x}$-martingale. In the following proposition we write $\mathcal{F}_s^t$, $s \in [t, T]$, for the σ-field generated by $X(\rho)$, $\rho \in [t, s]$. The proof of the following proposition could be based on Theorem 2.11 in Chapter 2. For convenience we provide a direct proof based on the Markov property.

Proposition 5.1. *Fix $t \in [\tau, T)$. Let the function $u : [t, T] \times E \to \mathbb{R}$ be such that $(s, x) \mapsto \dfrac{\partial u}{\partial s}(s, x) + L(s)u\left(s, \cdot\right)(x)$ belongs to $C_b\left([t, T] \times E\right) := C_b\left([t, T] \times E; \mathbb{C}\right)$. Then the process $s \mapsto M_u(s) - M_u(t)$ is adapted to the filtration of σ-fields $(\mathcal{F}_s^t)_{s\in[t,T]}$.*

Proof. Suppose that $T \geqslant s_2 > s_1 \geqslant t$. In order to check the martingale property of the process $M_u(s) - M_u(t)$, $s \in [t, T]$, it suffices to prove that

$$\mathbb{E}_{t,x}\left[M_u\left(s_2\right) - M_u\left(s_1\right) \,\middle|\, \mathcal{F}_{s_1}^t\right] = 0. \tag{5.13}$$

In order to prove (5.13) we notice that by the time-inhomogeneous Markov property:

$$\mathbb{E}_{t,x}\left[M_u\left(s_2\right) - M_u\left(s_1\right) \,\middle|\, \mathcal{F}_{s_1}^t\right] = \mathbb{E}_{s_1,X(s_1)}\left[M_u\left(s_2\right) - M_u\left(s_1\right)\right]$$

$$
\begin{aligned}
&= \mathbb{E}_{s_1,X(s_1)}\Bigg[u\left(s_2,X\left(s_2\right)\right)-u\left(s_1,X\left(s_1\right)\right)\\
&\quad -\int_{s_1}^{s_2}\left(L(s)u\left(s,X(s)\right)+\frac{\partial u}{\partial s}\left(s,X(s)\right)\right)ds\Bigg]\\
&= \mathbb{E}_{s_1,X(s_1)}\left[u\left(s_2,X\left(s_2\right)\right)-u\left(s_1,X\left(s_1\right)\right)\right]\\
&\quad -\int_{s_1}^{s_2}\mathbb{E}_{s_1,X(s_1)}\left[\left(L(s)u\left(s,X(s)\right)+\frac{\partial u}{\partial s}\left(s,X(s)\right)\right)\right]ds\\
&= \mathbb{E}_{s_1,X(s_1)}[u(s_2,X(s_2))-u(s_1,X(s_1))]-\int_{s_1}^{s_2}\frac{d}{ds}\mathbb{E}_{s_1,X(s_1)}[u(s,X(s))]ds\\
&= \mathbb{E}_{s_1,X(s_1)}\left[u\left(s_2,X\left(s_2\right)\right)-u\left(s_1,X\left(s_1\right)\right)\right]\\
&\quad -\mathbb{E}_{s_1,X(s_1)}\left[u\left(s_2,X\left(s_2\right)\right)-u\left(s_1,X\left(s_1\right)\right)\right]=0. \qquad (5.14)
\end{aligned}
$$

The equality in (5.14) establishes the result in Proposition 5.1. □

As explained in Definition 5.1 it is assumed that the subspace $D(L)$ contains an algebra of functions which forms a core for the operator L.

Proposition 5.2. *Let the family of operators $L(s)$, $0 \leqslant s \leqslant T$, generate a time-inhomogeneous Markov process*

$$\left\{\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x}\right),\left(X(t):T\geqslant t\geqslant 0\right),\left(E,\mathcal{E}\right)\right\} \tag{5.15}$$

in the sense of Definition 5.3: see equality (5.12). Then the process $X(t)$ has a modification which is right-continuous and has left limits on its life time.

For the definition of life time see e.g. Theorem 2.9. The life time ζ is defined by

$$\zeta=\begin{cases}\inf\{s>0:\ X(s)=\triangle\} & \text{on the event } \{X(s)=\triangle \text{ for some } s\in(0,T)\},\\ \zeta=T, & \text{if } X(s)\in E \text{ for all } s\in(0,T).\end{cases} \tag{5.16}$$

In view of Proposition 5.2 we will assume that our Markov process has left limits on its life time and is continuous from the right. The following proof is a correct outline of a proof of Proposition 5.2. If E is just a Polish space it needs a considerable adaptation. Suppose that E is Polish, and first assume that the process $t \mapsto X(t)$ is conservative, i.e. assume that $\mathbb{P}_{\tau,x}\left[X(t)\in E\right]=1$. Then, by an important intermediate result (see Proposition 3.1 in Chapter 3 and the arguments leading to it) we see that the orbits $\{X(\rho):\tau\leqslant\rho\leqslant T\}$ are $\mathbb{P}_{\tau,x}$-almost surely relatively compact in E. In case that the process $t \mapsto X(t)$ is not conservative, i.e. if, for some fixed $t \in [\tau, T]$, an inequality of the form $\mathbb{P}_{\tau,x}[X(t) \in E] < 1$ holds, then a similar result is still valid. In fact on the event $\{X(t) \in E\}$ the orbit $\{X(\rho) : \tau \leqslant \rho \leqslant t\}$ is $\mathbb{P}_{\tau,x}$-almost surely relatively compact: see Proposition 3.2 in Chapter 3. All details can be found in the proof of Theorem 2.9: see Subsection 3.1.1 in Chapter 3.

Proof. As indicated earlier the argument here works in case the space E is locally compact. However, the result is true for a Polish space E: see Theorem 2.9.

Let the function $u : [0, T] \times E \to \mathbb{R}$ belong to the space $D(L)$. Then the process $s \mapsto M_u(s) - M_u(t)$, $t \leqslant s \leqslant T$, is a $\mathbb{P}_{t,x}$-martingale. Let $D[0, T]$ be the set of numbers of the form $k2^{-n}T$, $k = 0, 1, 2, \ldots, 2^n$. By a classical martingale convergence theorem (see e.g. Chapter II in [Revuz and Yor (1999)]) it follows that the following limit $\lim_{s\uparrow t,\, s\in D[0,T]} u(s, X(s))$ exists $\mathbb{P}_{\tau,x}$-almost surely for all $0 \leqslant \tau < t \leqslant T$ and for all $x \in E$. In the same reference it is also shown that the limit $\lim_{s\downarrow t,\, s\in D[0,T]} u(s, X(s))$ exists $\mathbb{P}_{\tau,x}$-almost surely for all $0 \leqslant \tau \leqslant t < T$ and for all $x \in E$. Since the locally compact space $[0, T] \times E$ is second countable it follows that the exceptional sets may be chosen to be independent of $(\tau, x) \in [0, T] \times E$, of $t \in [\tau, T]$, and of the function $u \in D(L)$. Since by hypothesis the subspace $D(L)$ is $\mathcal{T}_\beta$-dense in $C_b([0, T] \times E)$ it follows that the left-hand limit at t of the process $s \mapsto X(s)$, $s \in D[0, T] \bigcap [\tau, t]$, exists $\mathbb{P}_{\tau,x}$-almost surely for all $(t, x) \in (\tau, T] \times E$. It also follows that the right-hand limit at t of the process $s \mapsto X(s)$, $s \in D[0, T] \bigcap (t, T]$, exists $\mathbb{P}_{\tau,x}$-almost surely for all $(t, x) \in [\tau, T) \times E$. Then we modify $X(t)$ by replacing it with $X(t+) = \lim_{s\downarrow t,\, s\in D[0,T] \bigcap (\tau,T]} X(s)$, $t \in [0, T)$, and $X(T+) = X(T)$. It also follows that the process $t \mapsto X(t+)$ has left limits in E.

This completes the proof of Proposition 5.2. □

The hypotheses in the following Proposition 5.3 are the same as those in Proposition 5.2. The functions u and v belong to $D^{(1)}(L) = D(D_1) \bigcap D(L)$: see Definition 2.7.

Proposition 5.3. *Let the continuous function $u : [0, T] \times E \to \mathbb{R}$ be such that for every $s \in [t, T]$ the function $x \mapsto u(s, x)$ belongs to $D(L(s))$ and suppose that the function $(s, x) \mapsto [L(s)u(s, \cdot)](x)$ is bounded and continuous. In addition suppose that the function $s \mapsto u(s, x)$ is continuously differentiable for all $x \in E$. Then the process $s \mapsto M_u(s) - M_u(t)$ is a $\mathcal{F}_s^t$-martingale with respect to the probability $\mathbb{P}_{t,x}$. If v is another such function, then the (right) derivative of the quadratic covariation process of* *the martingales M_u and M_v is given by:*

$$\frac{d}{dt}\langle M_u, M_v\rangle(t) = \Gamma_1(u,v)(t, X(t)).$$

In fact the following identity holds as well:

$$\begin{aligned} &M_u(t)M_v(t) - M_u(0)M_v(0) \\ &= \int_0^t M_u(s)dM_v(s) + \int_0^t M_v(s)dM_u(s) + \int_0^t \Gamma_1(u,v)(s,X(s))\,ds. \end{aligned} \tag{5.17}$$

Here $\mathcal{F}_s^t$, $s \in [t,T]$, is the σ-field generated by the state variables $X(\rho)$, $t \leqslant \rho \leqslant s$. Instead of $\mathcal{F}_s^0$ we usually write $\mathcal{F}_s$, $s \in [0,T]$. The formula in (5.17) is known as the integration by parts formula for stochastic integrals.

Proof. We outline a proof of the equality in (5.17). So let the functions u and v be as in Proposition 5.3. Then we have

$$\begin{aligned} &M_u(t)M_v(t) - M_u(0)M_v(0) \\ &= \sum_{k=0}^{2^n-1} M_u\left(k2^{-n}t\right)\left(M_v\left((k+1)2^{-n}t\right) - M_v\left(k2^{-n}t\right)\right) \\ &\quad + \sum_{k=0}^{2^n-1} \left(M_u\left((k+1)2^{-n}t\right) - M_u\left(k2^{-n}t\right)\right) M_v\left(k2^{-n}t\right) \\ &+ \sum_{k=0}^{2^n-1} (M_u((k+1)2^{-n}t) - M_u(k2^{-n}t))(M_v((k+1)2^{-n}t) - M_v(k2^{-n}t)). \end{aligned} \tag{5.18}$$

The first term on the right-hand side of (5.18) converges to $\int_0^t M_u(s)dM_v(s)$, the second term converges to $\int_0^t M_v(s)dM_u(s)$. Using the identity in (5.7) for the function u and a similar identity for v we see that the third term on the right-hand side of (5.18) converges to $\int_0^t \Gamma_1(u,v)(s,X(s))\,ds$.

The observation that for every $\tau \in [0,T]$ the process

$$t \mapsto M_u(t)M_v(t) - M_u(\tau)M_v(\tau) - \int_\tau^t \Gamma_1(u(s,\cdot), v(s,\cdot))(X(s))\,ds, \tag{5.19}$$

$\tau \leqslant t \leqslant T$, is a $\mathbb{P}_{\tau,x}$-martingale relative to the filtration $(\mathcal{F}_t^\tau)_{t\in[\tau,T]}$, then completes the proof Proposition 5.3. □

Remark 5.2. The quadratic variation process of the (local) martingale $s \mapsto M_u(s)$ is given by the process $s \mapsto \Gamma_1(u(s,\cdot), u(s,\cdot))(s, X(s))$, and therefore

$$\mathbb{E}_{s_1,x}\left[\left|\int_{s_1}^{s_2} dM_u(s)\right|^2\right] = \mathbb{E}_{s_1,x}\left[\int_{s_1}^{s_2} \Gamma_1(u(s,\cdot), u(s,\cdot))(X(s))\,ds\right] < \infty$$

under appropriate conditions on the function u. Very informally we may think of the following representation for the martingale difference:

$$M_u(s_2) - M_u(s_1) = \int_{s_1}^{s_2} \nabla_u^L(s, X(s))\, dW(s). \tag{5.20}$$

Here we still have to give a meaning to the stochastic integral in the right-hand side of (5.20). If E is an infinite-dimensional Banach space, then $W(t)$ should be some kind of a cylindrical Brownian motion. It is closely related to a formula which occurs in Malliavin calculus: see [Nualart (1995)] (Proposition 3.2.1) and [Nualart (1998)].

Remark 5.3. It is perhaps worthwhile to observe that for Brownian motion $(W(s), \mathbb{P}_x)$ the martingale difference $M_u(s_2) - M_u(s_1)$, $s_1 \leqslant s_2 \leqslant T$, is given by a stochastic integral:

$$M_u(s_2) - M_u(s_1) = \int_{s_1}^{s_2} \nabla u(\tau, W(\tau))\, dW(\tau).$$

Its increment of the quadratic variation process is given by

$$\langle M_u, M_u\rangle(s_2) - \langle M_u, M_u\rangle(s_1) = \int_{s_1}^{s_2} \left|\nabla u(\tau, W(\tau))\right|^2 d\tau.$$

Next suppose that the function u solves the equation:

$$f\left(s, x, u(s,x), \nabla_u^L(s,x)\right) + L(s)u(s,x) + \frac{\partial}{\partial s}u(s,x) = 0. \tag{5.21}$$

If moreover, $u(T,x) = \varphi(T,x)$, $x \in E$, is given, then we have

$$\begin{aligned} u(t, X(t)) = \varphi(T, X(T)) + \int_t^T f\left(s, X(s), u(s, X(s)), \nabla_u^L(s, X(s))\right) ds \\ - \int_t^T dM_u(s), \end{aligned} \tag{5.22}$$

with $M_u(s)$ as in (5.7). From (5.22) we get

$$\begin{aligned} u(t,x) &= \mathbb{E}_{t,x}\left[u(t, X(t))\right] \\ &= \mathbb{E}_{t,x}\left[\varphi(T, X(T))\right] + \int_t^T \mathbb{E}_{t,x}\left[f\left(s, X(s), u(s, X(s)), \nabla_u^L(s, X(s))\right)\right] ds. \end{aligned} \tag{5.23}$$

Theorem 5.1. *Let $u : [0,T] \times E \to \mathbb{R}$ be a continuous function with the property that for every $(t,x) \in [0,T] \times E$ the function $s \mapsto \mathbb{E}_{t,x}\left[u(s, X(s))\right]$ is differentiable and that*

$$\frac{d}{ds}\mathbb{E}_{t,x}\left[u(s, X(s))\right] = \mathbb{E}_{t,x}\left[L(s)u(s, X(s)) + \frac{\partial}{\partial s}u(s, X(s))\right], \quad t < s < T.$$

Then the following assertions are equivalent:

(a) The function u satisfies the following differential equation:

$$L(t)u(t,x)+\frac{\partial}{\partial t}u(t,x)+f\left(t,x,u(t,x),\nabla_u^L(t,x)\right)=0. \tag{5.24}$$

(b) The function u satisfies the following type of Feynman-Kac integral equation:

$$u(t,x)=\mathbb{E}_{t,x}\left[u(T,X(T))+\int_t^T f(\tau,X(\tau),u(\tau,X(\tau)),\nabla_u^L(\tau,X(\tau)))d\tau\right]. \tag{5.25}$$

(c) For every $t\in[0,T]$ the process

$$s\mapsto u(s,X(s))-u(t,X(t))+\int_t^s f\left(\tau,X(\tau),u(\tau,X(\tau)),\nabla_u^L(\tau,X(\tau))\right)d\tau$$

is an $\mathcal{F}_s^t$-martingale with respect to $\mathbb{P}_{t,x}$ on the interval $[t,T]$.

(d) For every $s\in[0,T]$ the process

$$t\mapsto u(T,X(T))-u(t,X(t))+\int_t^T f\left(\tau,X(\tau),u(\tau,X(\tau)),\nabla_u^L(\tau,X(\tau))\right)d\tau$$

is an $\mathcal{F}_T^t$-backward martingale with respect to $\mathbb{P}_{s,x}$ on the interval $[s,T]$.

Remark 5.4. Suppose that the function u is a solution to the following terminal value problem:

$$\begin{cases} L(s)u(s,\cdot)(x)+\frac{\partial}{\partial s}u(s,x)+f\left(s,x,u(s,x),\nabla_u^L(s,x)\right)=0; \\ u(T,x)=\varphi(T,x). \end{cases} \tag{5.26}$$

Then the pair $\left(u(s,X(s)),\nabla_u^L(s,X(s))\right)$ can be considered as a weak solution to a backward stochastic differential equation. More precisely, for every $s\in[0,T]$ the process

$$t\mapsto u(T,X(T))-u(t,X(t))+\int_t^T f\left(\tau,X(\tau),u(\tau,X(\tau)),\nabla_u^L(\tau,X(\tau))\right)d\tau$$

is an $\mathcal{F}_T^t$-backward martingale relative to $\mathbb{P}_{s,x}$ on the interval $[s,T]$. The symbol $\nabla_u^L v(s,x)$ stands for the functional $v\mapsto\nabla_u^L v(s,x)=\Gamma_1(u,v)(s,x)$, where Γ_1 is the squared gradient operator:

$$\begin{aligned}&\Gamma_1(u,v)(s,x) \\ &=\mathcal{T}_\beta\text{-}\lim_{t\downarrow s}\frac{1}{t-s}\mathbb{E}_{s,x}\left[(u(s,X(t))-u(s,X(s)))(v(s,X(t))-v(s,X(s)))\right].\end{aligned} \tag{5.27}$$

Possible choices for the function f are

$$f\left(s,x,y,\nabla_u^L\right) = -V(s,x)y \quad \text{and} \tag{5.28}$$

$$f\left(s,x,y,\nabla_u^L\right) = \frac{1}{2}\left|\nabla_u^L(s,x)\right|^2 - V(s,x) = \frac{1}{2}\Gamma_1(u,u)(s,x) - V(s,x). \tag{5.29}$$

The choice in (5.28) turns equation (5.26) into the following heat equation:

$$\begin{cases} \frac{\partial}{\partial s}u(s,x) + L(s)u(s,\cdot)(x) - V(s,x)u(s,x) = 0; \\ u(T,x) = \varphi(T,x). \end{cases} \tag{5.30}$$

The function $v(s,x)$ defined by the Feynman-Kac formula

$$v(s,x) = \mathbb{E}_{s,x}\left[e^{-\int_s^T V(\rho,X(\rho))d\rho}\varphi(T,X(T))\right] \tag{5.31}$$

is a candidate solution to equation (5.30).
The choice in (5.29) turns equation (5.26) into the following Hamilton-Jacobi-Bellman equation:

$$\begin{cases} \frac{\partial}{\partial s}u(s,x) + L(s)u(s,X(s)) - \frac{1}{2}\Gamma_1(u,u)(s,x) + V(s,x) = 0; \\ u(T,x) = -\log\varphi(T,x), \end{cases} \tag{5.32}$$

where $-\log\varphi(T,x)$ replaces $\varphi(T,x)$. The function S_L defined by the genuine non-linear Feynman-Kac formula

$$S_L(s,x) = -\log\mathbb{E}_{s,x}\left[e^{-\int_s^T V(\rho,X(\rho))d\rho}\varphi(T,X(T))\right] \tag{5.33}$$

is a candidate solution to (5.32). Often these "candidate solutions" are viscosity solutions. However, this was the main topic in [Van Casteren (2009)] and is the main topic in Chapter 6.

Remark 5.5. Let $u(s,x)$ satisfy one of the equivalent conditions in Theorem 5.1. Put $Y(\tau) = u(\tau,X(\tau))$, and let $M(s)$ be the martingale determined by $M(t) = Y(t) = u(t,X(t))$ and by

$$M(s) - M(t) = Y(s) + \int_t^s f\left(\tau,X(\tau),Y(\tau),\nabla_u^L(\tau,X(\tau))\right)d\tau.$$

Then the expression $\nabla_u^L(\tau,X(\tau))$ only depends on the martingale part M of the process $s \mapsto Y(s)$. This entitles us to write $Z_M(\tau)$ instead of $\nabla_u^L(\tau,X(\tau))$. The interpretation of $Z_M(\tau)$ is then the linear functional $N \mapsto \frac{d}{d\tau}\langle M,N\rangle(\tau)$, where N is a $\mathbb{P}_{t,x}$-martingale in $\mathcal{M}^2\left(\Omega,\mathcal{F}_T^t,\mathbb{P}_{t,x}\right)$. Here a process N belongs to $\mathcal{M}^2\left(\Omega,\mathcal{F}_T^t,\mathbb{P}_{t,x}\right)$ whenever N is martingale in

$L^2(\Omega, \mathcal{F}_T^t, \mathbb{P}_{t,x})$ which is $\mathbb{P}_{t,x}$-almost surely continuous. In Definition 5.7 below it will be explained why these functionals exist. Their existence is guaranteed by Lebesgue's differentiation theorem. For a discussion on this theorem see e.g. [Stein and Shakarchi (2005)]. Notice that the functional $Z_M(\tau)$ is known as soon as the martingale $M \in \mathcal{M}^2(\Omega, \mathcal{F}_T^t, \mathbb{P}_{t,x})$ is known. From our definitions it also follows that

$$M(T) = Y(T) + \int_t^T f(\tau, X(\tau), Y(\tau), Z_M(\tau))\, d\tau, \quad \mathbb{P}_{t,x}\text{-almost surely}$$

provided that $Y(t) = M(t)$.

Remark 5.6. Let the notation be as in Remark 5.5. Then the variables $Y(t)$ and $Z_M(t)$ only depend on the space-time variable $(t, X(t))$, and as a consequence the martingale increments $M(t_2) - M(t_1)$, $0 \leqslant t_1 < t_2 \leqslant T$, only depend on $\mathcal{F}_{t_2}^{t_1} = \sigma(X(s) : t_1 \leqslant s \leqslant t_2)$. In Section 5.2 we give Lipschitz type conditions on the function f in order that the BSDE

$$Y(t) = Y(T) + \int_t^T f(s, X(s), Y(s), Z_M(s))\, ds + M(t) - M(T), \quad \tau \leqslant t \leqslant T, \tag{5.34}$$

possesses a unique pair of solutions

$$(Y, M) \in L^2(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}) \times \mathcal{M}^2(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}).$$

Here $\mathcal{M}^2(\Omega, \mathcal{F}_T^t, \mathbb{P}_{t,x})$ stands for the space of all $\mathbb{P}_{t,x}$-almost sure continuous $(\mathcal{F}_s^t)_{s\in[t,T]}$-martingales in $L^2(\Omega, \mathcal{F}_T^t, \mathbb{P}_{t,x})$. Of course instead of writing "BSDE" it would be better to write "BSIE" for Backward Stochastic Integral Equation. However, since in the literature on backward stochastic differential equations people write "BSDE" even if they mean integral equations we also stick to this terminology. Suppose that the $\sigma(X(T))$-measurable variable $Y(T) \in L^2(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})$ is given. In fact we will prove that the solution (Y, M) of the equation in (5.34) belongs to the space $\mathcal{S}^2\left(\Omega, \mathcal{F}_T^t, \mathbb{P}_{t,x}; \mathbb{R}^k\right) \times \mathcal{M}^2\left(\Omega, \mathcal{F}_T^t, \mathbb{P}_{t,x}; \mathbb{R}^k\right)$. For more details see the definitions 5.4 and 5.8, and Theorem 5.7.

Remark 5.7. Let M and N be two martingales in $\mathcal{M}^2[0,T]$. Then, for $0 \leqslant s < t \leqslant T$,

$$\begin{aligned} &\left|\langle M, N\rangle(t) - \langle M, N\rangle(s)\right|^2 \\ &\leqslant \left(\langle M, M\rangle(t) - \langle M, M\rangle(s)\right)\left(\langle N, N\rangle(t) - \langle N, N\rangle(s)\right), \end{aligned}$$

and consequently

$$\left|\frac{d}{ds}\langle M, N\rangle(s)\right|^2 \leqslant \frac{d}{ds}\langle M, M\rangle(s) \frac{d}{ds}\langle N, N\rangle(s).$$

Hence, the inequality

$$\int_0^T \left| \frac{d}{ds} \langle M, N \rangle (s) \right| ds \leqslant \int_0^T \left(\frac{d}{ds} \langle M, M \rangle (s) \right)^{1/2} \left(\frac{d}{ds} \langle N, N \rangle (s) \right)^{1/2} ds \tag{5.35}$$

follows. The inequality in (5.35) says that the quantity $\int_0^T \left| \frac{d}{ds} \langle M, N \rangle (s) \right| ds$ is dominated by the Hellinger integral $\mathcal{H}(M, N)$ defined by the right-hand side of (5.35).

Remark 5.8. BSDEs, which can be of quadratic order, are also used in the context of concave utility functions (and their Legendre-Fenchel transforms, which are the so-called cost functions): see e.g. [Delbaen *et al.* (2009)]. Such functions are used in the theory of risk management. For more results on quadratic BSDEs see [Reveillac (2009)], and [Imkeller *et al.* (2009)]. In the latter paper the authors also describe the role of Malliavin calculus in the representations for solutions to BSDEs. Suppose that the underlying filter space is standard d-dimensional Brownian motion. Then the Malliavin derivative of the solutions represents the integrand of the martingale part of the solution (written as a Skorohod integrals, which turns out to be an Itô integral).

For a proof of Theorem 5.1 we refer the reader to [Van Casteren (2009)]. We insert a proof here as well.

Proof. [Proof of Theorem 5.1.] For brevity, and only in this proof, we write

$$F(\tau, X(\tau)) = f\left(\tau, X(\tau), u(\tau, X(\tau)), \nabla_u^L(\tau, X(\tau))\right).$$

(a) $\Longrightarrow$ (b). The equality in (b) is the same as the one in (5.23) which is a consequence of (5.21).

(b) $\Longrightarrow$ (a). We calculate the expression

$$\frac{\partial}{\partial s}\mathbb{E}_{t,x}\left[u(s, X(s)) + \int_t^s f\left(\tau, X(\tau), u(\tau, X(\tau)), \nabla_u^L(\tau, X(\tau))\right) d\tau\right].$$

First of all it is equal to

$$\mathbb{E}_{t,x}\left[\frac{\partial}{\partial s}u(s, X(s)) + L(s)u(s, X(s)) + F(s, X(s))\right]. \tag{5.36}$$

Next we also have by (5.25) in (b):

$$\frac{\partial}{\partial s}\mathbb{E}_{t,x}\left[u(s, X(s)) + \int_t^s f\left(\tau, X(\tau), u(\tau, X(\tau)), \nabla_u^L(\tau, X(\tau))\right) d\tau\right]$$

$$= \frac{\partial}{\partial s}\mathbb{E}_{t,x}\left[\mathbb{E}_{s,X(s)}\left[u\left(T,X(T)\right)+\int_s^T F\left(\tau,X(\tau)\right)\,d\tau\right]+\int_t^s F\left(\tau,X(\tau)\right)\,d\tau\right]$$

(Markov property)

$$= \frac{\partial}{\partial s}\mathbb{E}_{t,x}\left[\mathbb{E}_{t,x}\left[u\left(T,X(T)\right)+\int_s^T F\left(\tau,X(\tau)\right)\,d\tau \mid \mathcal{F}_s^t\right]+\int_t^s F\left(\tau,X(\tau)\right)\,d\tau\right]$$

$$= \frac{\partial}{\partial s}\mathbb{E}_{t,x}\left[\mathbb{E}_{t,x}\left[u\left(T,X(T)\right)+\int_t^T F\left(\tau,X(\tau)\right)\,d\tau \mid \mathcal{F}_s^t\right]\right]$$

$$= \frac{\partial}{\partial s}\mathbb{E}_{t,x}\left[u\left(T,X(T)\right)+\int_t^T F\left(\tau,X(\tau)\right)\,d\tau\right] = 0. \tag{5.37}$$

From (5.37) and (5.36) we get

$$\mathbb{E}_{t,x}\left[\frac{\partial}{\partial s}u(s,X(s))+L(s)u(s,X(s))+f(s,X(s),u(s,X(s)),\nabla_u^L(s,X(s)))\right]$$
$$= 0, \quad s > t. \tag{5.38}$$

Passing to the limit for $s \downarrow t$ in (5.38) we obtain:

$$\mathbb{E}_{t,x}\left[\frac{\partial}{\partial t}u\left(t,X(t)\right)+L(t)u\left(t,X(t)\right)+f\left(t,X\left(t\right),u\left(t,X\left(t\right)\right),\nabla_u^L\left(t,X\left(t\right)\right)\right)\right]$$
$$= 0. \tag{5.39}$$

Since $X(t) = x$ $\mathbb{P}_{t,x}$-almost surely, from (5.39) we obtain equality (5.24) in assertion (a).

(a) $\Longrightarrow$ (c). If the function u satisfies the differential equation in (a), then from the equality in (5.5) we see that

$$0 = u\left(s,X\left(s\right)\right)-u\left(t,X\left(t\right)\right)+\int_t^s f\left(\tau,X(\tau),u\left(\tau,X(\tau)\right),\nabla_u^L\left(\tau,X(\tau)\right)\right)d\tau$$
$$-u\left(s,X\left(s\right)\right)+u\left(t,X\left(t\right)\right)+\int_t^s\left(L(\tau)u\left(\tau,X(\tau)\right)+\frac{\partial u}{\partial \tau}\left(\tau,X(\tau)\right)\right)d\tau \tag{5.40}$$

$$= u\left(s,X\left(s\right)\right)-u\left(t,X\left(t\right)\right)+\int_t^s f\left(\tau,X(\tau),u\left(\tau,X(\tau)\right),\nabla_u^L\left(\tau,X(\tau)\right)\right)d\tau$$
$$-M_u\left(s\right)+M_u\left(t\right), \tag{5.41}$$

where, as in (5.7),

$$M_u\left(s\right)-M_u\left(t\right)$$

$$= u(s, X(s)) - u(t, X(t)) - \int_t^s \left(L(\tau)u(\tau, X(\tau)) + \frac{\partial u}{\partial \tau}(\tau, X(\tau)) \right) d\tau$$
$$= \int_t^s dM_u(\tau). \tag{5.42}$$

Since the expression in (5.41) vanishes (by assumption (a)) we see that the process in (c) is the same as the martingale $s \mapsto M_u(s) - M_u(t)$, $s \geqslant t$. This proves the implication (a) $\Longrightarrow$ (c).

The implication (c) $\Longrightarrow$ (b) is a direct consequence of assertion (c) and the fact that $X(t) = x$ $\mathbb{P}_{t,x}$-almost surely.

The equivalence of the assertions (a) and (d) is proved in the same manner as the equivalence of (a) and (c). Here we employ the fact that the process $t \mapsto M_u(T) - M_u(t)$ is an $\mathcal{F}_T^t$-backward martingale on the interval $[s, T]$ with respect to the probability $\mathbb{P}_{s,x}$.

This completes the proof of Theorem 5.1 □

Remark 5.9. Instead of considering $\nabla_u^L(s, x)$ we will also consider the bilinear mapping $Z(s)$ which associates with a pair of local semi-martingales (Y_1, Y_2) a process which is to be considered as the right derivative of the covariation process: $\langle Y_1, Y_2 \rangle(s)$. We write

$$Z_{Y_1}(s)(Y_2) = Z(s)(Y_1, Y_2) = \frac{d}{ds}\langle Y_1, Y_2 \rangle(s).$$

The function f (i.e. the generator of the backward differential equation) will then be of the form: $f(s, X(s), Y(s), Z_Y(s))$; the deterministic phase $\big(u(s,x), \nabla_u^L(s,x)\big)$ is replaced with the stochastic phase $(Y(s), Z_Y(s))$. We should find an appropriate stochastic phase $s \mapsto (Y(s), Z_Y(s))$, which we identify with the process $s \mapsto (Y(s), M_Y(s))$ in the stochastic phase space $\mathcal{S}^2 \times \mathcal{M}^2$, such that

$$Y(t) = Y(T) + \int_t^T f(s, X(s), Y(s), Z_Y(s))\, ds - \int_t^T dM_Y(s), \tag{5.43}$$

where the quadratic variation of the martingale $M_Y(s)$ is given by

$$d\langle M_Y, M_Y \rangle(s) = Z_Y(s)(Y)\, ds = Z(s)(Y, Y)\, ds = d\langle Y, Y \rangle(s).$$

This stochastic phase space $\mathcal{S}^2 \times \mathcal{M}^2$ plays a role in stochastic analysis very similar to the role played by the first Sobolev space $H^{1,2}$ in the theory of deterministic partial differential equations. For a formal definition of the functional $Z_M(s)$ the reader is referred to Definition 5.7.

Remark 5.10. In case we deal with strong solutions driven by standard Brownian motion the martingale difference $M_Y(s_2) - M_Y(s_1)$ can be written as $\int_{s_1}^{s_2} Z_Y(s)dW(s)$, provided that the martingale $M_Y(s)$ belongs to the space $\mathcal{M}^2\left(\Omega, \mathcal{G}_T^0, \mathbb{P}\right)$. Here $\mathcal{G}_T^0$ is the σ-field generated by $W(s)$, $0 \leqslant s \leqslant T$. If $Y(s) = u(s, X(s))$, then this stochastic integral satisfies:

$$\begin{aligned}\int_{s_1}^{s_2} Z_Y(s)dW(s) &= u(s_2, X(s_2)) - u(s_1, X(s_1)) \\ &\quad - \int_{s_1}^{s_2} \left(L(s) + \frac{\partial}{\partial s}\right) u(s, X(s))\, ds. \end{aligned} \tag{5.44}$$

Such stochastic integrals are for example defined if the process $X(t)$ is a solution to a stochastic differential equation (in Itô sense):

$$X(s) = X(t) + \int_t^s b(\tau, X(\tau))\, d\tau + \int_t^s \sigma(\tau, X(\tau))\, dW(\tau), \quad t \leqslant s \leqslant T. \tag{5.45}$$

Here the matrix $\left(\sigma_{jk}(\tau, x)\right)_{j,k=1}^d$ is chosen in such a way that

$$a_{jk}(\tau, x) = \sum_{\ell=1}^d \sigma_{j\ell}(\tau, x)\, \sigma_{k\ell}(\tau, x) = \left(\sigma(\tau, x)\sigma^*(\tau, x)\right)_{jk}.$$

The process $W(\tau)$ is Brownian motion or Wiener process. It is assumed that operator $L(\tau)$ has the form

$$L(\tau)u(x) = b(\tau, x) \cdot \nabla u(x) + \frac{1}{2} \sum_{j,k=1}^d a_{jk}(\tau, x) \frac{\partial^2}{\partial x_j x_k} u(x). \tag{5.46}$$

Then from Itô's formula together with (5.44), (5.45) and (5.46) it follows that the process $Z_Y(s)$ has to be identified with $\sigma(s, X(s))^* \nabla u(s, \cdot)(X(s))$. For more details see e.g. [Pardoux and Peng (1990)] and [Pardoux (1998a)]. The equality in (5.44) is a consequence of a martingale representation theorem: see e.g. Proposition 3.2 in [Revuz and Yor (1999)].

Remark 5.11. Backward doubly stochastic differential equations (BDSDEs) could have been included in the present chapter: see Boufoussi, Mrhardy and Van Casteren [Boufoussi *et al.* (2007)]. In our notation a BDSDE may be written in the form:

$$\begin{aligned} Y(t) - Y(T) &= \int_t^T f\left(s, X(s), Y(s), N \mapsto \frac{d}{ds}\langle M, N\rangle(s)\right) ds \\ &\quad + \int_t^T g\left(s, X(s), Y(s), N \mapsto \frac{d}{ds}\langle M, N\rangle(s)\right) d\overleftarrow{B}(s) \end{aligned}$$

$$+ M(t) - M(T). \tag{5.47}$$

Here the expression

$$\int_t^T g\left(s, X(s), Y(s), N \mapsto \frac{d}{ds}\langle M, N\rangle(s)\right) d\overleftarrow{B}(s)$$

represents a backward Itô integral. The symbol $\langle M, N\rangle$ stands for the quadratic covariation process of the (local) martingales M and N; it is assumed that this process is absolutely continuous with respect to Lebesgue measure. Moreover,

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), (E, \mathcal{E})\right\}$$

is a Markov process generated by a family of operators $L(s)$, $0 \leqslant s \leqslant T$, and $\mathcal{F}_t^\tau = \sigma\left\{X(s) : \tau \leqslant s \leqslant t\right\}$. The process $X(t)$ could be the (unique) weak or strong solution to a (forward) stochastic differential equation (SDE):

$$X(t) = x + \int_\tau^t b\left(s, X(s)\right) ds + \int_\tau^t \sigma\left(s, X(s)\right) dW(s). \tag{5.48}$$

Here the coefficients b and σ have certain continuity or measurability properties, and $\mathbb{P}_{\tau,x}$ is the distribution of the process $X(t)$ defined as being the unique weak solution to the equation in (5.48). We want to find a pair $(Y, M) \in \mathcal{S}^2\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}\right) \times \mathcal{M}^2\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}\right)$ which satisfies (5.47). For applications of BDSDEs to viscosity solutions of stochastic partial differential equations the reader is referred to e.g. [Buckdahn and Ma (2001a,b); N'zi and Owo (2009); Pardoux and Peng (1994)].

Next we give some definitions. Fix $(\tau, x) \in [0, T] \times E$. In the definitions 5.4 and 5.5 the probability measure $\mathbb{P}_{\tau,x}$ is defined on the σ-field $\mathcal{F}_T^\tau$. In Definition 5.8 we return to these notions. The following definition and implicit results described therein show that, under certain conditions, by enlarging the sample space a family of processes may be reduced to just one process without losing the $\mathcal{S}^2$-property.

Definition 5.4. Fix $(\tau, x) \in [0, T] \times E$. An $\mathbb{R}^k$-valued process Y is said to belong to the space $\mathcal{S}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ if $Y(t)$ is $\mathcal{F}_t^\tau$-measurable ($\tau \leqslant t \leqslant T$) and if $\mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} |Y(t)|^2\right] < \infty$. It is assumed that $Y(s) = Y(\tau)$, $\mathbb{P}_{\tau,x}$-almost surely, for $s \in [0, \tau]$. The process $Y(s)$, $s \in [0, T]$, is said to belong to the space $\mathcal{S}^2_{\text{unif}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ if

$$\sup_{(\tau,x)\in[0,T]\times E} \mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} |Y(t)|^2\right] < \infty,$$

and it belongs to $\mathcal{S}^2_{\mathrm{loc,unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ provided that

$$\sup_{(\tau,x)\in[0,T]\times K} \mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} |Y(t)|^2\right] < \infty$$

for all compact subsets K of E.

If the σ-field $\mathcal{F}^\tau_t$ and $\mathbb{P}_{\tau,x}$ are clear from the context we write $\mathcal{S}^2\left([0,T], \mathbb{R}^k\right)$ or sometimes just $\mathcal{S}^2$.

Definition 5.5. Let the process M be such that the process $t \mapsto M(t) - M(\tau)$, $t \in [\tau, T]$, is a $\mathbb{P}_{\tau,x}$-almost surely continuous martingale with the property that the random variable $M(T) - M(\tau)$ belongs to $L^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$. Then M is said to belong to the space $\mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$. By the Burkholder-Davis-Gundy inequality (see inequality (5.89) below) it follows that

$$\mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} |M(t) - M(\tau)|^2\right]$$

is finite if and only if $M(T) - M(\tau)$ belongs to the space $L^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$. Here an $\mathcal{F}^\tau_t$-adapted process $M(\cdot) - M(\tau)$ is called a $\mathbb{P}_{\tau,x}$-martingale provided that $\mathbb{E}_{\tau,x}\left[|M(t) - M(\tau)|\right] < \infty$ and $\mathbb{E}_{\tau,x}\left[M(t) - M(\tau) \,\middle|\, \mathcal{F}^\tau_s\right] = M(s) - M(\tau)$, $\mathbb{P}_{\tau,x}$-almost surely, for $T \geqslant t \geqslant s \geqslant \tau$. The $\mathbb{P}_{\tau,x}$-almost sure continuous martingale difference $s \mapsto M(s) - M(\tau)$, $s \in [\tau, T]$, is said to belong to the space $\mathcal{M}^2_{\mathrm{unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ if

$$\sup_{(\tau,x)\in[0,T]\times E} \mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} |M(t) - M(\tau)|^2\right] < \infty,$$

and it belongs to $\mathcal{M}^2_{\mathrm{loc,unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ provided that

$$\sup_{(\tau,x)\in[0,T]\times K} \mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} |M(t) - M(\tau)|^2\right] < \infty$$

for all compact subsets K of E.

There is also need for a localized notion.

Definition 5.6. Let $M \in \mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$. Then M is said to be absolutely continuous if the deterministic function $t \mapsto \mathbb{E}_{\tau,x}\left[|M(t)|^2\right]$ is absolutely continuous. The attribute AC is used to indicate that a martingale $M \in \mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ is absolutely continuous: $M \in \mathcal{M}^2_{\mathrm{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$. For the other spaces a similar notation is used: $\mathcal{M}^2_{\mathrm{AC,unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$.

From the Burkholder-Davis-Gundy inequality (see inequality (5.89) below) it follows that the process $M(s) - M(0)$ belongs to $\mathcal{M}^2_{\text{unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ if and only if

$$\begin{aligned}&\sup_{(\tau,x)\in[0,T]\times E} \mathbb{E}_{\tau,x}\left[\left|M(T) - M(\tau)\right|^2\right] \\ &= \sup_{(\tau,x)\in[0,T]\times E} \mathbb{E}_{\tau,x}\left[\langle M, M\rangle(T) - \langle M, M\rangle(\tau)\right] < \infty.\end{aligned}$$

Here $\langle M, M\rangle$ stands for the quadratic variation process of the process $t \mapsto M(t) - M(0)$.

The notions in the definitions 5.4 and 5.5 will exclusively be used in case the family of measures $\{\mathbb{P}_{\tau,x} : (\tau, x) \in [0,T] \times E\}$ constitute the distributions of a Markov process which was defined in Definition 5.2.

In order to formalize our theory we insert a definition of the fiber spaces $\mathcal{M}^{2,s}_{\text{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$, $\tau \leqslant s \leqslant T$. As mentioned in Definition 5.5 the space $\mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ consists of those L^2-martingales which are $\mathbb{P}_{\tau,x}$-almost surely continuous.

Definition 5.7. By definition the space $\mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ consists of those continuous martingales M with values in the space $\mathbb{R}^k$ which belong to $L^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$. The symbol $\mathcal{M}^{2,s}_{\text{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ consists of those functionals

$$Z(s) : \mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right) \to \mathbb{R}$$

for which there exists a martingale $M \in \mathcal{M}^2_{\text{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ such that for all $N \in \mathcal{M}^2_{\text{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ the equality

$$Z(s)(N) = \frac{d}{dt}\langle M, N\rangle(t)\Big|_{t=s+} = \lim_{h\downarrow 0} \frac{\langle M, N\rangle(s+h) - \langle M, N\rangle(s)}{h} \tag{5.49}$$

holds. Usually the notation $\frac{d}{ds}\langle M, N\rangle(s)$ is employed for the right-derivative as indicated in (5.49). The notation $Z(s) = Z_M(s)$ is used and the $\mathcal{M}^{2,s}_{\text{AC}}$-norm of $Z_M(s)$ is defined by

$$\left\|Z_M(s)\right\|_{\mathcal{M}^{2,s}_{\text{AC}}} = \left(\mathbb{E}_{\tau,x}\left[\frac{d}{ds}\langle M, M\rangle(s)\right]\right)^{1/2}.$$

Elements of the space $\mathcal{M}^{2,s}_{\text{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ are denoted by $Z_M(s)$, where the martingale M belongs to $\mathcal{M}^2\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$. From the Kunita-Watanabe inequality (see [Ikeda and Watanabe (1998)]) it follows that

$$\left|\frac{d}{ds}\langle M, N\rangle(s)\right|^2 \leqslant \left|\frac{d}{ds}\langle M, M\rangle(s)\right| \left|\frac{d}{ds}\langle N, N\rangle(s)\right|,$$

where M, $N \in \mathcal{M}^2_{\mathrm{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$, and so it makes sense to define the following inner-product on the space $\mathcal{M}^{2,s}_{\mathrm{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$:

$$\left\langle Z_M(s), Z_N(s)\right\rangle_{\mathcal{M}^{2,s}_{\mathrm{AC}}} = \mathbb{E}_{\tau,x}\left[\frac{d}{ds}\left\langle M, N\right\rangle(s)\right], \quad M,\, N \in \mathcal{M}^2_{\mathrm{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right), \tag{5.50}$$

and the $\mathcal{M}^{2,s}_{\mathrm{AC}}$-norm of $Z_M(s)$ is defined by

$$\left\|Z_M(s)\right\|_{\mathcal{M}^{2,s}_{\mathrm{AC}}} = \left(\mathbb{E}_{\tau,x}\left[\frac{d}{ds}\left\langle M, M\right\rangle(s)\right]\right)^{1/2}.$$

Relative to this inner-product and norm the space $\mathcal{M}^{2,s}_{\mathrm{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right)$ is a pre-Hilbert space. Completion turns it into a Hilbert space.

Lemma 5.1 below gives some more information on these fiber spaces.

Example 5.1. Again let the Markov process, with right-continuous sample paths and with left limits,

$$\left\{\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), \left(E, \mathcal{E}\right)\right\} \tag{5.51}$$

be generated by the family of operators $\{L(s) : 0 \leqslant s \leqslant t\}$: see definitions 5.2, equality (5.9), and 5.3, equality (5.11). Suppose that the squared gradient operators $\Gamma_1(s)$, $0 \leqslant s \leqslant T$, exist: see equality (5.2). Let the function $u \in C_b\left([0,T] \times E\right)$ belong to the domain of the operator $L = L(s)$, $0 \leqslant s \leqslant T$. Put

$$M_{u,\tau}(s) = u\left(s, X(s)\right) - u\left(\tau, X(\tau)\right) - \int_\tau^t \left(\frac{d}{d\rho} + L(\rho)\right) u\left(\rho, X(\rho)\right)\, d\rho.$$

Then the process $s \mapsto M_{u,\tau}(s)$ is a $\mathbb{P}_{\tau,x}$-martingale and, since

$$\left\langle M_{u,\tau}, M_{u,\tau}\right\rangle(s) - \left\langle M_{u,\tau}, M_{u,\tau}\right\rangle(\tau) = \int_\tau^s \Gamma_1\left(u,u\right)\left(\rho, X(\rho)\right)\, d\rho$$

the martingale $M_{u.\tau}$ belongs to the space $\mathcal{M}^2_{\mathrm{AC}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}\right)$.

Next we define the family of operators $\left\{Q\left(t_1, t_2\right) :\, 0 \leqslant t_1 \leqslant t_2 \leqslant T\right\}$ by

$$Q\left(t_1, t_2\right) f(x) = \mathbb{E}_{t_1,x}\left[f\left(X\left(t_2\right)\right)\right], \quad f \in C_b\left(E\right),\ 0 \leqslant t_1 \leqslant t_2 \leqslant T. \tag{5.52}$$

Fix $\varphi \in D(L)$. Since the process $t \mapsto M_\varphi(t) - M_\varphi(s)$, $t \in [s,T]$, is a $\mathbb{P}_{s,x}$-martingale with respect to the filtration $\left(\mathcal{F}^s_t\right)_{t\in[s,T]}$, and $X(t) = x$ $\mathbb{P}_{t,x}$-almost surely, the following equality follows:

$$\int_s^t \mathbb{E}_{s,x}\left[L(\rho)\varphi\left(\rho, \cdot\right)\left(X(\rho)\right)\right] d\rho + \mathbb{E}_{t,x}\left[\varphi\left(t, X(t)\right)\right] - \mathbb{E}_{s,x}\left[\varphi\left(t, X(t)\right)\right]$$

$$= \varphi(t,x) - \varphi(s,x) - \int_s^t \mathbb{E}_{s,x}\left[\frac{\partial\varphi}{\partial\rho}(\rho, X(\rho))\right] d\rho. \tag{5.53}$$

The fact that a process of the form $t \mapsto M_\varphi(t) - M_\varphi(s)$, $t \in [s,T]$, is a $\mathbb{P}_{s,x}$-martingale follows from Proposition 5.1. In terms of the family of operators

$$\{Q(t_1,t_2) : 0 \leqslant t_1 \leqslant t_2 \leqslant T\}$$

the equality in (5.53) can be rewritten as

$$\int_s^t Q(s,\rho) L(\rho)\varphi(\rho,\cdot)(x)\, d\rho + Q(t,t)\varphi(t,\cdot)(x) - Q(s,t)\varphi(t,\cdot)(x)$$
$$= \varphi(t,x) - \varphi(s,x) - \int_s^t Q(s,\rho) \frac{\partial\varphi}{\partial\rho}(\rho,\cdot)(x) d\rho. \tag{5.54}$$

From (5.54) we infer that

$$L(s)\varphi(s,\cdot)(x) = -\lim_{t\downarrow s} \frac{Q(t,t)\varphi(t,\cdot)(x) - Q(s,t)\varphi(t,\cdot)(x)}{t-s}$$

and that

$$L(t)\varphi(t,\cdot)(x) = \lim_{s\uparrow t} \frac{Q(t,t)\varphi(t,\cdot)(x) - Q(s,t)\varphi(t,\cdot)(x)}{t-s}. \tag{5.55}$$

Equality (5.54) also yields the following result. If $\varphi \in D(L)$ is such that

$$L(\rho)\varphi(\rho,\cdot)(y) = -\frac{\partial\varphi}{\partial\rho}(\rho,y),$$

then

$$\varphi(s,x) = Q(\rho,t)\varphi(t,\cdot)(x) = \mathbb{E}_{s,x}[\varphi(t,X(t))]. \tag{5.56}$$

Since $0 \leqslant s \leqslant t \leqslant T$ are arbitrary from (5.56) we see

$$Q(s,t')\varphi(t',\cdot)(x) = Q(s,t)Q(t,t')\varphi(t',\cdot)(x) \quad 0 \leqslant s \leqslant t \leqslant t' \leqslant T,\ x \in E. \tag{5.57}$$

If in (5.57) we (may) choose the function $\varphi(t',y)$ arbitrarily, then the family $Q(s,t)$, $0 \leqslant s \leqslant t \leqslant T$, is automatically a propagator in the space $C_b(E)$ in the sense that $Q(s,t)Q(t,t') = Q(s,t')$, $0 \leqslant s \leqslant t \leqslant t' \leqslant T$. For details on propagators or evolution families see [Gulisashvili and van Casteren (2006)].

Remark 5.12. In the sequel we want to discuss solutions to equations of the form:

$$\frac{\partial}{\partial t} u(t,x) + L(t)u(t,\cdot)(x) + f\left(t,x,u(t,x), \nabla_u^L(t,x)\right) = 0. \tag{5.58}$$

For a preliminary discussion on this topic see Theorem 5.1. Under certain hypotheses on the function f we will give existence and uniqueness results. Let m be (equivalent to) the Lebesgue measure in $\mathbb{R}^d$. In a concrete situation where every operator $L(t)$ is a genuine diffusion operator in $L^2\left(\mathbb{R}^d, m\right)$ we consider the following Backward Stochastic Differential equation

$$\begin{aligned} u\left(s, X(s)\right) = Y\left(T, X(T)\right) + \int_s^T f\left(\rho, X(\rho), u\left(\rho, X(\rho)\right), \nabla_u^L\left(\rho, X(\rho)\right)\right) d\rho \\ - \int_s^T \nabla_u^L\left(\rho, X(\rho)\right) dW\left(\rho\right). \end{aligned} \tag{5.59}$$

Here we suppose that the process $t \mapsto X(t)$ is a solution to a genuine stochastic differential equation driven by Brownian motion and with one-dimensional distribution $u(t,x)$ satisfying $L(t)u\left(t,\cdot\right)(x) = \frac{\partial u}{\partial t}(t,x)$. In fact in that case we will not consider the equation in (5.59), but we will try to find an ordered pair (Y,Z) such that

$$Y(s) = Y(T) + \int_s^T f\left(\rho, X(\rho), Y\left(\rho\right), Z\left(\rho\right)\right) d\rho - \int_s^T \left\langle Z\left(\rho\right), dW\left(\rho\right)\right\rangle. \tag{5.60}$$

If the pair (Y,Z) satisfies (5.60), then $u\left(s,x\right) = \mathbb{E}_{s,x}\left[Y(s)\right]$ satisfies (5.58). Moreover $Z(s) = \nabla_u^L\left(s, X(s)\right) = \nabla_u^L\left(s,x\right)$, $\mathbb{P}_{s,x}$-almost surely. For more details see section 2 in [Pardoux (1998a)].

Remark 5.13. Some remarks follow:

(a) In section 5.2 weak solutions to BSDEs are studied.
(b) In section 7 of [Van Casteren (2009)] and in section 2 of [Pardoux (1998a)] strong solutions to BSDEs are discussed: these results are due to Pardoux and collaborators.
(c) BSDEs go back to Nelson [Nelson (1967)]. In this context Bismut is also mentioned see e.g. [Bismut (1973, 1981b)].
(d) If $L(s)u(s,x) = \frac{1}{2}\sum_{j,k=1}^d a_{j,k}(s,x)\frac{\partial^2 u}{\partial x_j x_k}(s,x) + \sum_{j=1}^d b_j(s,x)\frac{\partial u}{\partial x_j}(s,x)$, then

$$\Gamma_1\left(u,v\right)\left(s,x\right) = \sum_{j,k=1}^d a_{j,k}(s,x)\frac{\partial u}{\partial x_j}(s,x)\frac{\partial v}{\partial x_k}(s,x).$$

As a corollary to theorems 5.1 and 5.5 we have the following result.

Corollary 5.1. *Suppose that the function u solves the following*

$$\begin{cases} \dfrac{\partial u}{\partial s}(s,y) + L(s)u(s,\cdot)(y) + f\left(s,y,u(s,y),\nabla_u^L(s,y)\right) = 0; \\ u\left(T,X(T)\right) = \xi \in L^2\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x}\right). \end{cases} \tag{5.61}$$

Let the pair (Y,M) be a solution to

$$Y(t) = \xi + \int_t^T f\left(s,X(s),Y(s),Z_M(s)\right) ds + M(t) - M(T), \tag{5.62}$$

with $M(\tau) = 0$. Then

$$(Y(t),M(t)) = \left(u\left(t,X(t)\right),M_u(t)\right),$$

where

$$M_u(t) = u(t,X(t)) - u(\tau,X(\tau)) - \int_\tau^t L(s)u(s,\cdot)(X(s))ds - \int_\tau^t \frac{\partial u}{\partial s}(s,X(s))ds.$$

Notice that the processes $s \mapsto \nabla_u^L\left(s,X(s)\right)$ and $s \mapsto Z_{M_u}(s)$ may be identified and that $Z_{M_u}(s)$ only depends on $(s,X(s))$. The decomposition

$$\begin{aligned} u\left(t,X(t)\right) - u\left(\tau,X(\tau)\right) = &\int_\tau^t \left(\frac{\partial u}{\partial s}\left(s,X(s)\right) + L(s)u\left(s,\cdot\right)\left(X(s)\right)\right) ds \\ &+ M_u(t) - M_u(\tau) \end{aligned} \tag{5.63}$$

splits the process $t \mapsto u\left(t,X(t)\right) - u\left(\tau,X(\tau)\right)$ into a part which is bounded variation (i.e. the part which is absolutely continuous with respect to Lebesgue measure on $[\tau,T]$) and a $\mathbb{P}_{\tau,x}$-martingale part $M_u(t) - M_u(\tau)$ (which in fact is a martingale difference part).

If $L(s) = \frac{1}{2}\Delta$, then $X(s) = W(s)$ (standard Wiener process or Brownian motion) and (5.63) can be rewritten as

$$\begin{aligned} u\left(t,W(t)\right) - u\left(\tau,W(\tau)\right) = &\int_\tau^t \left(\frac{\partial u}{\partial s}\left(s,W(s)\right) + \frac{1}{2}\Delta u\left(s,\cdot\right)\left(W(s)\right)\right) ds \\ &+ \int_\tau^t \nabla u\left(s,\cdot\right)\left(W(s)\right) dW(s) \end{aligned} \tag{5.64}$$

where $\int_\tau^t \nabla u\left(s,\cdot\right)\left(W(s)\right) dW(s)$ is to be interpreted as an Itô integral.

Remark 5.14. Suggestions for further research:

(a) Find "explicit solutions" to BSDEs with a linear drift part. This should be a type of Cameron-Martin formula or Girsanov transformation.
(b) Treat weak (and strong) solutions BDSDEs in a manner similar to what is presented here for BSDEs.

(c) Treat weak (strong) solutions to BSDEs generated by a function f which is not necessarily of linear growth but for example of quadratic growth in one or both of its entries $Y(t)$ and $Z_M(t)$.
(d) Can anything be done if f depends not only on s, x, $u(s,x)$, $\nabla_u(s,x)$, but also on $L(s)u(s,\cdot)(x)$?

In the following proposition it is assumed that the operator L generates a strong Markov process in the sense of the definitions 2.7 and 2.8.

Proposition 5.4. *Let the functions f, $g \in D(L)$ be such that their product fg also belongs to $D(L)$. Then $\Gamma_1(f,g)$ is well defined and for $(s,x) \in [0,T] \times E$ the following equality holds:*

$$\begin{aligned} &L(s)(fg)(s,\cdot)(x) - f(s,x)L(s)g(s,\cdot)(x) - L(s)f(s,\cdot)(x)g(s,x) \\ &= \Gamma_1(f,g)(s,x). \end{aligned} \tag{5.65}$$

Proof. Let the functions f and g be as in Proposition 5.4. For $h > 0$ we have:

$$\begin{aligned} &(f(X(s+h)) - f(X(s)))(g(X(s+h)) - g(X(s))) \\ &= f(X(s+h))g(X(s+h)) - f(X(s))g(X(s)) \qquad (5.66) \\ &\quad - f(X(s))(g(X(s+h)) - g(X(s))) - (f(X(s+h)) - f(X(s)))g(X(s)). \end{aligned}$$

Then we take expectations with respect to $\mathbb{E}_{s,x}$, divide by $h > 0$, and pass to the $\mathcal{T}_\beta$-limit as $h \downarrow 0$ to obtain equality (5.65) in Proposition 5.4. □

5.2 A probabilistic approach: Weak solutions

In this section and also in sections 5.3 we will study BSDE's on a single probability space. In Section 5.4 and Chapter 6 we will consider Markov families of probability spaces. In the present section we write $\mathbb{P}$ instead of $\mathbb{P}_{0,x}$, and similarly for the expectations $\mathbb{E}$ and $\mathbb{E}_{0,x}$. Here we work on the interval $[0,T]$. Since we are discussing the martingale problem and basically only the distributions of the process $t \mapsto X(t)$, $t \in [0,T]$, the solutions we obtain are of weak type. In case we consider strong solutions we apply a martingale representation theorem (in terms of Brownian Motion). In Section 5.4 we will also use this result for probability measures of the form $\mathbb{P}_{\tau,x}$ on the interval $[\tau,T]$. In this section we consider a pair of $\mathcal{F}_t = \mathcal{F}_t^0$-adapted processes $(Y,M) \in L^2(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^k) \times L^2(\Omega,\mathcal{F}_T,\mathbb{P} : \mathbb{R}^k)$ such that $Y(0) = M(0)$ and such that

$$Y(t) = Y(T) + \int_t^T f(s,X(s),Y(s),Z_M(s))\,ds + M(t) - M(T) \tag{5.67}$$

where M is a $\mathbb{P}$-martingale with respect to the filtration $\mathcal{F}_t = \sigma(X(s) : s \leqslant t)$. In [Van Casteren (2009)] we will employ the results of the present section with $\mathbb{P} = \mathbb{P}_{\tau,x}$, where $(\tau, x) \in [0,T] \times E$. For more details see §5.4 below.

Proposition 5.5. *Let the pair* (Y, M) *be as in (5.67), and suppose that* $Y(0) = M(0)$. *Then*

$$Y(t) = M(t) - \int_0^t f(s, X(s), Y(s), Z_M(s))\, ds, \quad \text{and} \tag{5.68}$$

$$Y(t) = \mathbb{E}\left[Y(T) + \int_t^T f(s, X(s), Y(s), Z_M(s))\, ds \;\Big|\; \mathcal{F}_t \right]; \tag{5.69}$$

$$M(t) = \mathbb{E}\left[Y(T) + \int_0^T f(s, X(s), Y(s), Z_M(s))\, ds \;\Big|\; \mathcal{F}_t \right]. \tag{5.70}$$

The equality in (5.68) shows that the process M is the martingale part of the semi-martingale Y.

Proof. The equality in (5.69) follows from (5.67) and from the fact that M is a martingale. Next we calculate

$$\begin{aligned}
&\mathbb{E}\left[Y(T) + \int_0^T f(s, X(s), Y(s), Z_M(s))\, ds \;\Big|\; \mathcal{F}_t \right] \\
&= \mathbb{E}\left[Y(T) + \int_t^T f(s, X(s), Y(s), Z_M(s))\, ds \;\Big|\; \mathcal{F}_t \right] \\
&\quad + \int_0^t f(s, X(s), Y(s), Z_M(s))\, ds \\
&= Y(t) + \int_0^t f(s, X(s), Y(s), Z_M(s))\, ds
\end{aligned}$$

(employ (5.67))

$$\begin{aligned}
&= Y(T) + \int_t^T f(s, X(s), Y(s), Z_M(s))\, ds + M(t) - M(T) \\
&\quad + \int_0^t f(s, X(s), Y(s), Z_M(s))\, ds \\
&= Y(T) + \int_0^T f(s, X(s), Y(s), Z_M(s))\, ds + M(t) - M(T) \\
&= M(T) + M(t) - M(T) = M(t).
\end{aligned} \tag{5.71}$$

The equality in (5.71) shows (5.70). Since

$$M(T) = Y(T) + \int_0^T f(s, X(s), Y(s), Z_M(s))\, ds$$

the equality in (5.68) follows. □

In the following theorem we write $z = Z_M(s) = Z_{M^1,\dots,M^k}(s)$ and y belongs to $\mathbb{R}^k$. The process M is a k-dimensional martingale in $\mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$. (see Definition 5.5) with the property that every function $t \mapsto \mathbb{E}_{\tau,x}\left[\left|M^j(t)\right|^2\right]$ is absolutely continuous. From Lemma 5.1 below it follows that it makes sense to write $\frac{d}{dt}\left\langle M^j, M^j\right\rangle(t)$. The expression $\frac{d}{dt}\langle M, M\rangle(t)$, $M = M_2 - M_1$, is shorthand for

$$\frac{d}{dt}\langle M, M\rangle(t) = \sum_{j=1}^k \frac{d}{dt}\left\langle M^j, M^j\right\rangle(t).$$

The notation $M \in \mathcal{M}^2_{\mathrm{AC}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ will be used: see Definition 5.6.

Theorem 5.2. *Fix $(\tau, x) \in [0, T] \times E$. Suppose that there exist finite constants C_1 and C_2 such that*

$$\left\langle y_2 - y_1, f\left(s, x', y_2, z\right) - f\left(s, x', y_1, z\right)\right\rangle \leqslant C_1 \left|y_2 - y_1\right|^2; \tag{5.72}$$

$$\left|f\left(s, x', y, Z_{M_2}(s)\right) - f\left(s, x', y, Z_{M_1}(s)\right)\right|^2 \leqslant C_2^2 \frac{d}{ds}\langle M_2 - M_1, M_2 - M_1\rangle(s) \tag{5.73}$$

for all $s \in [\tau, T]$, $x' \in E$, $y \in \mathbb{R}^k$, $z = Z_M(s) \in \mathcal{M}^{2,s}_{\mathrm{AC}}(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})$. Then there exists a unique pair of adapted processes $(Y, M) \in \mathcal{S}^2(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}) \times \mathcal{M}^2(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})$ such that $Y(\tau) = M(\tau)$ and such that the process M is the martingale part of the semi-martingale Y:

$$\begin{aligned} Y(t) &= M(t) - M(T) + Y(T) + \int_t^T f(s, X(s), Y(s), Z_M(s))\, ds \\ &= M(t) - \int_\tau^t f(s, X(s), Y(s), Z_M(s))\, ds, \quad \mathbb{P}_{\tau,x}\text{-almost surely,} \end{aligned} \tag{5.74}$$

for all $t \in [\tau, T]$

For the definition of the space $\mathcal{M}^{2,s}_{\mathrm{AC}}(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})$ see Definition 5.7 above. The symbols $Z_M(s)$ stand for the functionals:

$$Z_M(s)(N) = \sum_{j=1}^k \frac{d}{ds}\left\langle M^j, N^j\right\rangle(s).$$

Here M and N are k-dimensional martingales in $\mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ with the property that the functions $t \mapsto \mathbb{E}_{\tau,x}\left[|M(t)|^2\right]$ and $t \mapsto \mathbb{E}_{\tau,x}\left[|N(t)|^2\right]$ are absolutely continuous it the follows that the function $t \mapsto \langle M, N\rangle(t) = \sum_{j=1}^k \left\langle M^j, N^j\right\rangle(t)$ is also $\mathbb{P}_{\tau,x}$-almost surely absolutely continuous. Then the Borel measure determined by such a function are $\mathbb{P}_{\tau,x}$-surely continuous relative to the Lebesgue measure on $[\tau, T]$. As a consequence we see that

$$\langle M, N\rangle(t) - \langle M, N\rangle(\tau) = \int_\tau^t \frac{d}{d\rho}\langle M, N\rangle(\rho)\, d\rho, \quad \mathbb{P}_{\tau,x}\text{-almost surely.} \tag{5.75}$$

The equality in (5.75) also determines the domain of the function f which generates the BSDE in (5.74) in Theorem 5.2. It is defined on the space

$$\left\{\left(s, x', y, Z_M(s)\right) : s \in [\tau, T],\, x' \in E,\, y \in \mathbb{R}^k,\, Z_M(s) \in \mathcal{M}_{\mathrm{AC}}^{2,s}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)\right\},$$

and is continuous on this space. For the notation $\mathcal{M}_{\mathrm{AC}}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$ see Definition 5.6. Let M be a member of $\mathcal{M}_{\mathrm{AC}}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$. Then the process of functionals (see Definition 5.7 above)

$$t \mapsto \{N \mapsto \langle M, N\rangle(t) - \langle M(\tau), N\rangle(\tau)\}, \quad N \in \mathcal{M}_{\mathrm{AC}}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right),$$

can be written as an element of the Hilbert-integral

$$\int_\tau^t \mathcal{M}^{2,s}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)(\cdot)\, ds.$$

More precisely $\langle M, N\rangle(t) - \langle M, N\rangle(\tau) = \int_\tau^t \frac{d}{ds}\langle M, N\rangle(s)\, ds$, or briefly

$$\langle M, \cdot\rangle(t) - \langle M, \cdot\rangle(\tau) = \int_\tau^t \frac{d}{ds}\langle M, \cdot\rangle(s)\, ds = \int_\tau^t Z_M(s)\, ds,$$

where $Z_M(s) \in \mathcal{M}_{\mathrm{AC}}^{2,s}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$. The functionals $Z_M(s)$ can be considered as the reproducing kernel of the quadratic covariation process determined by the martingale $M \in \mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$. Lemma 5.1 below gives some formal arguments concerning the existence of these derivatives. For more details on direct Hilbert integrals and reproducing kernel techniques see e.g. [Thomas (1979, 1994)] and Berlinet and Thomas-Agnan [Berlinet and Thomas-Agnan (2004)].

5.2.1 *Some more explanation*

Suppose that the family of operators $L(s)$, $0 \leqslant s \leqslant T$, generates the strong Markov process

$$\left\{\left(\Omega, \mathcal{F}, \mathbb{P}_{s,x}\right), \left(X(t), 0 \leqslant t \leqslant T\right), (E, \mathcal{E})\right\}.$$

Consider the operators $L(s)$, $0 \leqslant s \leqslant T$, as an operator with domain and range in $C_b([0,T]\times E)$. Let $\mathcal{F}_t^s$ be the σ-field generated by $X(\rho)$, $s \leqslant \rho \leqslant t$. Let $u \in D(L)$. Then the process

$$M_{u,\tau} : t \mapsto u(t, X(t)) - u(\tau, X(\tau)) - \int_\tau^t \left(\frac{\partial}{\partial \rho} + L(\rho)\right) u(\rho, \cdot)(X(\rho))\, d\rho,$$

$t \in [\tau, T]$, is a $\mathbb{P}_{\tau,x}$-martingale. In addition, suppose that the squared gradient operators $\Gamma_1(s)$, $0 \leqslant s \leqslant T$, exist. Let v be another function in $D(L)$. To the function v there also corresponds a $\mathbb{P}_{\tau,x}$-martingale $t \mapsto M_{v,\tau}(t)$, $t \in [\tau, T]$. Then the covariation process $t \mapsto \langle M_{u,\tau}, M_{v,\tau}\rangle(t)$ is given by

$$\langle M_{u,\tau}, M_{v,\tau}\rangle(t) = \int_\tau^t \Gamma_1(u,v)(s, X(s))\, ds, \quad t \in [\tau, T].$$

In other words the covariation process $t \mapsto \langle M_{u,\tau}, M_{v,\tau}\rangle(t)$ is absolutely continuous with respect to the Lebesgue measure. So for such martingales it makes sense to write

$$\frac{d}{dt}\langle M_{u,\tau}, M_{v,\tau}\rangle(t) = \lim_{h\downarrow 0} \frac{\langle M_{u,\tau}, M_{v,\tau}\rangle(t+h) - \langle M_{u,\tau}, M_{v,\tau}\rangle(t)}{h}, \quad \mathbb{P}_{\tau,x}\text{-a.s..}$$

More generally we will consider martingales $M \in \mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ with the property that the function $t \mapsto \mathbb{E}_{\tau,x}\left[|M(t)|^2\right]$ is absolutely continuous with respect to the Lebesgue measure: for more details on the space $\mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ see Remark 5.5. If M is such a martingale, then the variation process $t \mapsto \langle M, M\rangle(t)$ is $\mathbb{P}_{s,x}$-almost surely absolutely continuous. The latter is explained in Lemma 5.1 below. It is assumed that the σ-field $\mathcal{F}_\tau^\tau$ contains the $\mathbb{P}_{\tau,x}$-negligible sets, and the filtration $(\mathcal{F}_t^\tau)_{t\in[\tau,T]}$ is continuous from the right.

Lemma 5.1. *Let λ be the Lebesgue measure on the interval $[\tau, T]$, and let M and N be martingales $\mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$, which by hypothesis are $\mathbb{P}_{\tau,x}$-almost surely continuous. Then the measure*

$$A \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T \mathbf{1}_A(\omega, s)\, d\langle M, N\rangle(s)\right], \quad A \in \mathcal{F}_T^\tau \otimes \mathcal{B}_{[\tau,T]}, \tag{5.76}$$

splits into two positive measures, one of which is absolutely continuous relative to the product measure $\mathbb{P}_{\tau,x}\times\lambda$, and another one which is singular relative to $\mathbb{P}_{\tau,x}\times\lambda$. Consequently, by the Lebesgue decomposition theorem there exists an adapted process $\rho \mapsto h_{M,N}(\rho)$ and a random measure $\nu_{M,N}$ such that

$$\langle M, M\rangle(t) - \langle M, M\rangle(\tau) = \int_\tau^t h_M(\rho)\, d\rho + \nu_{M,N}(\tau, t].$$

By Lebesgue's differentiation theorem it follows that

$$\langle M, N\rangle(t) - \langle M, N\rangle(\tau) = \int_\tau^t \frac{d}{d\rho}\langle M, N\rangle(\rho)\,d\rho = \int_\tau^t h_{M,N}(\rho)\,d\rho,$$

where the derivatives are in fact right derivatives.

We use the notation

$$Z_M(s)(N) = h_{M,N}(s) = \frac{d}{ds}\langle M, N\rangle(s) \quad \mathbb{P}_{\tau,x} \times \lambda\text{-almost everywhere.}$$

By definition functionals of the form $Z_M(s)$, $M \in \mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ belong to the space $\mathcal{M}^{2,s}_{\text{AC}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$. Often we use the shorthand notation: $\mathcal{M}^{2,s}_{\text{AC}}$.

Proof. By the Lebesgue decomposition theorem there exists a process $\rho \mapsto h_{M,N}(\rho)$, $\rho \in [\tau, T]$, and a measure $\nu_{M,N}$ which is singular relative to the measure $\mathbb{P}_{\tau,x} \times \lambda$ such that for $A \in \mathcal{F}_T^\tau \otimes \mathcal{B}_{[\tau,T]}$ we have

$$\begin{aligned}&\mathbb{E}_{\tau,x}\left[\int_\tau^T \mathbf{1}_A(\omega,\rho)\, d\langle M, M\rangle(\rho)\right]\\ &\quad = \mathbb{E}_{\tau,x}\left[\int_\tau^T \mathbf{1}_A(\omega,\rho)\, h_{M,N}(\rho)\,d\rho\right] + \nu_{M,N}(A).\end{aligned} \tag{5.77}$$

In (5.77) we take A of the form $A = C \times [\tau, t]$, $C \in \mathcal{F}_T^\tau$, $t \in [\tau, T]$. Then we see

$$\langle M, M\rangle(t) - \langle M, M\rangle(\tau) = \int_\tau^t h_{M,N}(\rho)\,d\rho + \nu_{M,N}((\tau, t]) \tag{5.78}$$

$\mathbb{P}_{\tau,x}$-almost surely. Since $t \in [\tau, T]$ is arbitrary, an application of Lebesgue's differentiation theorem yields the equality $h(\rho) = \dfrac{d}{d\rho}\langle M, M\rangle(\rho)$, $\mathbb{P}_{\tau,x} \times \lambda$-almost everywhere (derivative from the right).

This completes the proof of Lemma 5.1. □

From Lemma 5.1 it follows that for L^2-martingales it makes sense to write $\dfrac{d}{dt}\langle M, M\rangle(t)$, $\mathbb{P}_{\tau,x} \times \lambda$-almost everywhere (derivatives from the right).

In the situation that $t \mapsto M(t)$ is a k-dimensional martingale relative to a filtration determined by d-dimensional Brownian motion the martingale $t \mapsto M(t) = \left(M^1(t), \dots, M^k(t)\right)$ can be written in the form

$$M^j(t) = \mathbb{E}_{\tau,x}\left[M^j(\tau)\right] + \sum_{k=1}^d \int_\tau^t \sigma_{j,k}(\rho)\,dW_k(\rho), \quad 1 \leqslant j \leqslant k, \tag{5.79}$$

where we employed a martingale representation theorem: see e.g. [Protter (2005)] Theorem 43 in Chapter IV. Then the covariation process of the martingales M^{j_1} and M^{j_2} is given by $\left\langle M^{j_1}, M^{j_2}\right\rangle(t) = \sum_{k=1}^{d}\int_0^t \sigma_{j_1,k}(s)\sigma_{j_2,k}(s)\,ds$. It follows that the functional $Z_M(t)$ can be identified with the matrix process $\left(\sigma_{j,k}(t)\right)_{1\leqslant j\leqslant k,\,1\leqslant k\leqslant d}$. Moreover, the estimate in (5.73) is a classical Lipschitz condition:

$$\left|f\left(s,x,y,Z_{M_2}(s)\right)-f\left(s,x,y,Z_{M_1}(s)\right)\right|^2 \leqslant C_2^2\sum_{j=1}^{k}\sum_{k=1}^{d}\left|\sigma_{j,k}^2(s)-\sigma_{j,k}^1(s)\right|^2. \tag{5.80}$$

Here $M_1^j(t) = \mathbb{E}_{\tau,x}\left[M_1^j(\tau)\right] + \sum_{k=1}^{d}\int_\tau^t \sigma_{j,k}^1(\rho)\,dW_k(\rho)$, $1 \leqslant j \leqslant k$, and a similar expression for $M_2^j(t)$. In other words our setup encompasses the classical theory of Pardoux and others.

Next suppose that the Markov process

$$\left\{\left(\Omega,\mathcal{F},\mathbb{P}_{s,x}\right),\left(X(t),0\leqslant t\leqslant T\right),\left(\mathbb{R}^d,\mathcal{B}_{\mathbb{R}^d}\right)\right\}$$

is generated by a second-order differential operator of the form

$$L(s) = \frac{1}{2}\sum_{k,\ell=1}^{d} a_{k,\ell}(s)\frac{\partial^2}{\partial x_k\partial x_\ell} = \frac{1}{2}\sum_{j=1}^{n}\sum_{k,\ell=1}^{d}\sigma_{k,j}(s)\sigma_{\ell,j}(s)\frac{\partial^2}{\partial x_k\partial x_\ell}.$$

Then the corresponding squared gradient operators $\Gamma_1(s)$, $0 \leqslant s \leqslant T$, are given by

$$\Gamma_1(u,v)\left(s,X(s)\right) = \sum_{k,\ell=1}^{d} a_{k,\ell}\left(s,X(s)\right)\frac{\partial u\left(s,X(s)\right)}{\partial x_k}\frac{\partial v\left(s,X(s)\right)}{\partial x_\ell}.$$

Consider the operator L as an operator in $C_b\left([0,T]\times\mathbb{R}^d\right)$, and let u and v be a functions in $D(L)$. Let $M_{u,\tau}$ be the martingale given by

$$M_{u,\tau}(t) = u\left(t,X(t)\right) - u\left(\tau,X(\tau)\right) - \int_\tau^t\left(\frac{\partial}{\partial s}+L(s)\right)u\left(s,X(s)\right)\,ds,$$

and we use a similar notation for $M_{v,\tau}(t)$. Then the covariation process of the martingales $M_{u,\tau}$ and $M_{v,\tau}$ is given by

$$\begin{aligned}\left\langle M_{u,\tau},M_{v,\tau}\right\rangle(t) &= \int_\tau^t \Gamma_1\left(u,v\right)\left(s,X(s)\right)\,ds \qquad (5.81)\\ &= \sum_{k,\ell=1}^{d}\int_\tau^t a_{k,\ell}\left(s,X(s)\right)\frac{\partial u\left(s,X(s)\right)}{\partial x_k}\frac{\partial v\left(s,X(s)\right)}{\partial x_\ell}ds.\end{aligned}$$

From (5.81) we infer

$$Z_{M_{u,\tau}}(s)\left(M_{v,\tau}\right) = \sum_{k,\ell=1}^{d} a_{k,\ell}\left(s, X(s)\right) \frac{\partial u\left(s, X(s)\right)}{\partial x_k} \frac{\partial v\left(s, X(s)\right)}{\partial x_\ell}.$$

The author is convinced that the present setup of BSDEs is also very convenient for Brownian motion on a Riemannian manifold, where we have a Laplace-Beltrami operator, and a squared gradient operator. For more information on Brownian motion on manifolds see e.g. [Elworthy (1982)], and [Hsu (2002)]. The book by Hsu also contains results on logarithmic Sobolev inequalities, and on spectral gap theory. These items will also be discussed in Chapter 9.

The following proof contains just an outline of the proof of Theorem 5.2. Complete and rigorous arguments are found in the proof of Theorem 5.4: see Theorem 5.7 as well.

Proof. [Outline of a proof of Theorem 5.2.] The uniqueness follows from Corollary 5.2 to Theorem 5.3 below. In the existence part of the proof of Theorem 5.2 we will approximate the function f by Lipschitz continuous functions f_δ, $0 < \delta < (2C_1)^{-1}$, where each function f_δ has Lipschitz constant δ^{-1}, but at the same time inequality (5.73) remains valid for fixed second variable (in an appropriate sense). It follows that for the functions f_δ (5.73) remains valid and that (5.72) is replaced with

$$\left|f_\delta\left(s, x, y_2, z\right) - f_\delta\left(s, x, y_1, z\right)\right| \leqslant \frac{1}{\delta}\left|y_2 - y_1\right|. \tag{5.82}$$

In the uniqueness part of the proof it suffices to assume that (5.72) holds. In Theorem 5.5 we will see that the monotonicity condition (5.72) also suffices to prove the existence. For details the reader is referred to the propositions 5.6 and 5.7, Corollary 5.3, and to Proposition 5.8. In fact for $M \in \mathcal{M}^2$ fixed, and the function $y \mapsto f\left(s, x, y, Z_M(s)\right)$ satisfying (5.72) the function $y \mapsto y - \delta f\left(s, x, y, Z_M(s)\right)$, $Z_M(s) \in \mathcal{M}^{2,s}_{\mathrm{AC}}$ is surjective as a mapping from $\mathbb{R}^k$ to $\mathbb{R}^k$ and its inverse exists and is Lipschitz continuous with constant 2, for $\delta > 0$ small enough. The Lipschitz continuity is proved in Proposition 5.7. The surjectivity of this mapping is a consequence of Theorem 1 in [Crouzeix *et al.* (1983)]. As pointed out by Crouzeix *et al.* the result follows from a non-trivial homotopy argument. A relatively elementary proof of Theorem 1 in [Crouzeix *et al.* (1983)] can be found for a continuously differentiable function in Hairer and Wanner [Hairer and Wanner (1991)]: see Theorem 14.2 in Chapter IV. In fact the result also is a consequence of the Browder-Minty theorem applied to the mapping

$y \mapsto y - \delta f(s, x, y, Z_M(s))$ where $\delta > 0$ is such that $\delta C_1 < 1$; see Theorem 5.10 in Subsection 5.4.1. For a few more details see remarks 5.19 and Remark 5.20. Let $f_{s,x,M}$ be the mapping $y \mapsto f(s, x, y, Z_M(s))$, and put

$$f_\delta(s, x, y, Z_M(s)) = f\left(s, x, (I - \delta f_{s,x,M})^{-1} y, Z_M(s)\right). \tag{5.83}$$

Then the functions f_δ, $0 < \delta < (2C_1)^{-1}$, are Lipschitz continuous with constant δ^{-1}. Proposition 5.8 treats the transition from solutions of BSDE's with generator f_δ with fixed martingale $M \in \mathcal{M}^2$ to solutions of BSDE's driven by f with the same fixed martingale M. Proposition 5.6 contains the passage from solutions $(Y, N) \in \mathcal{S}^2 \times \mathcal{M}^2$ to BSDE's with generators of the form $(s, y) \mapsto f(s, y, Z_M(s))$ for any fixed martingale $M \in \mathcal{M}^2$ to solutions for BSDE's of the form (5.74) where the pair (Y, M) belongs to $\mathcal{S}^2 \times \mathcal{M}^2$. By hypothesis the process $s \mapsto f(s, x, Y(s), Z_M(s))$ satisfies (5.72) and (5.73). Essentially speaking a combination of these observations show the result in Theorem 5.2. □

Remark 5.15. In the literature functions with the monotonicity property are also called one-sided Lipschitz functions. In fact Theorem 5.2, with $f(t, x, \cdot, \cdot)$ Lipschitz continuous in both variables, will be superseded by Theorem 5.4 in the Lipschitz case and by Theorem 5.5 in case of monotonicity in the second variable and Lipschitz continuity in the third variable. The proof of Theorem 5.2 is part of the results in Section 5.3. Theorem 5.7 contains a corresponding result for a Markov family of probability measures. Its proof is omitted, it follows the same lines as the proof of Theorem 5.5.

5.3 Existence and uniqueness of solutions to BSDE's

The equation in (5.58) can be phrased in a semi-linear setting as follows. Find a function $u(t, x)$ which satisfies the following partial differential equation:

$$\begin{cases} \dfrac{\partial u}{\partial s}(s, x) + L(s)u(s, x) + f\left(s, x, u(s, x), \nabla_u^L(s, x)\right) = 0; \\ u(T, x) = \varphi(T, x), \quad x \in E. \end{cases} \tag{5.84}$$

Here $\nabla_{f_2}^L(s, x)$ is the linear functional $f_1 \mapsto \Gamma_1(f_1, f_2)(s, x)$ for smooth enough functions f_1 and f_2. For $s \in [0, T]$ fixed the symbol $\nabla_{f_2}^L$ stands for the linear mapping $f_1 \mapsto \Gamma_1(f_1, f_2)(s, \cdot)$. One way to treat this kind of equation is considering the following backward problem. Find a pair of adapted processes (Y, Z_Y), satisfying

$$Y(t) - Y(T) - \int_t^T f\left(s, X(s), Y(s), Z(s)\left(\cdot, Y\right)\right) ds = M(t) - M(T), \tag{5.85}$$

where $M(s)$, $t_0 < t \leqslant s \leqslant T$, is a forward local $\mathbb{P}_{t,x}$-martingale (for every $T > t > t_0$). The symbol Z_{Y_1}, $Y_1 \in \mathcal{S}^2\left([0,T], \mathbb{R}^k\right)$, stands for the functional

$$Z_{Y_1}\left(Y_2\right)(s) = Z(s)\left(Y_1(\cdot), Y_2(\cdot)\right) = \frac{d}{ds}\left\langle Y_1(\cdot), Y_2(\cdot)\right\rangle(s),\ Y_2 \in \mathcal{S}^2\left([0,T], \mathbb{R}^k\right). \tag{5.86}$$

If the pair (Y, Z_Y) satisfies (5.85), then $Z_Y = Z_M$. For a precise definition of the functional $Z_M(s)$ see Definition 5.7 above. Instead of trying to find the pair (Y, Z_Y) we will try to find a pair $(Y, M) \in \mathcal{S}^2\left([0,T], \mathbb{R}^k\right) \times \mathcal{M}^2\left([0,T], \mathbb{R}^k\right)$ such that

$$Y(t) = Y(T) + \int_t^T f\left(s, X(s), Y(s), Z_M(s)\right) ds + M(t) - M(T).$$

Next we define the spaces $\mathcal{S}^2\left([0,T], \mathbb{R}^k\right)$ and $\mathcal{M}^2\left([0,T], \mathbb{R}^k\right)$: compare with the definitions 5.4 and 5.5.

Definition 5.8. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{F}_t$, $t \in [0,T]$, be a filtration on $\mathcal{F}$. Let $t \mapsto Y(t)$ be an stochastic process with values in $\mathbb{R}^k$ which is adapted to the filtration $\mathcal{F}_t$ and which is $\mathbb{P}$-almost surely continuous. Then Y is said to belong to the space $\mathcal{S}^2\left([0,T], \mathbb{R}^k\right)$ provided that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t)|^2\right] < \infty.$$

Definition 5.9. The space of $\mathbb{P}$-almost surely continuous $\mathbb{R}^k$-valued martingales in $L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^k\right)$ is denoted by $\mathcal{M}^2\left([0,T], \mathbb{R}^k\right)$. So that a continuous martingale $t \mapsto M(t) - M(0)$ belongs to $\mathcal{M}^2\left([0,T], \mathbb{R}^k\right)$ if

$$\mathbb{E}\left[|M(T) - M(0)|^2\right] < \infty. \tag{5.87}$$

Since the process $t \mapsto |M(t)|^2 - |M(0)|^2 - \langle M, M\rangle(t) + \langle M, M\rangle(0)$ is a martingale difference we see that

$$\mathbb{E}\left[|M(T) - M(0)|^2\right] = \mathbb{E}\left[\langle M, M\rangle(T) - \langle M, M\rangle(0)\right], \tag{5.88}$$

and hence a martingale difference $t \mapsto M(t) - M(0)$ in $L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^k\right)$ belongs to $\mathcal{M}^2\left([0,T], \mathbb{R}^k\right)$ if and only if $\mathbb{E}\left[\langle M, M\rangle(T) - \langle M, M\rangle(0)\right]$ is finite. By the Burkholder-Davis-Gundy inequality this is the case if and only if

$$\mathbb{E}\left[\sup_{0<t<T} |M(t) - M(0)|^2\right] < \infty.$$

To be precise, let $M(s)$, $t \leqslant s \leqslant T$, be a continuous local L^2-martingale taking values in $\mathbb{R}^k$. Put $M^*(s) = \sup_{t\leqslant\tau\leqslant s} |M(\tau)|$. Fix $0 < p < \infty$. The Burkholder-Davis-Gundy inequality says that there exist universal finite and strictly positive constants c_p and C_p such that

$$c_p\mathbb{E}\left[(M^*(s))^{2p}\right] \leqslant \mathbb{E}\left[\langle M(\cdot), M(\cdot)\rangle^p (s)\right] \leqslant C_p\mathbb{E}\left[(M^*(s))^{2p}\right], \quad t \leqslant s \leqslant T. \tag{5.89}$$

If $p = 1$, then $c_p = \frac{1}{4}$, and if $p = \frac{1}{2}$, then $c_p = \frac{1}{8}\sqrt{2}$. For more details and a proof see e.g. [Ikeda and Watanabe (1998)]. A version for càdlàg martingales, and $p \geqslant 1$, can be found as Theorem 48 in [Protter (2005)]. For the original, and more general result with convex functions, see [Burkholder *et al.* (1972)].

As in Definition 5.6 there is a need for martingales with an absolutely continuous variation process. That is why we insert the following definitions. We also need a precise notion of the functionals $Z_M(s)$, where M is a continuous martingale in the space $\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$. Compare Definition 5.10 with Definition 5.7 above.

Definition 5.10. Let M be a ($\mathbb{P}$-almost surely continuous) martingale in $\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$. Then M is said to be absolutely continuous if the function $t \mapsto \mathbb{E}\left[|M(t)|^2\right]$ is absolutely continuous. The subspace of $\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ consisting of the absolutely continuous martingales is denoted by $\mathcal{M}^2_{\rm AC}\left([0,T],\mathbb{R}^k\right)$. Let $M \in \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$. As in the proof of Lemma 5.1 it follows that for every $N \in \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ the right derivative $s \mapsto \dfrac{d}{ds}\langle M,N\rangle(s) = \lim_{h\downarrow 0}\dfrac{\langle M,N\rangle(s+h) - \langle M,N\rangle(s)}{h}$ exists $\mathbb{P}\times\lambda$-almost everywhere. Here λ is the Lebesgue measure on $[0,T]$. So for $s \in [0,T)$ fixed it makes sense to introduce the spaces $\mathcal{M}^{2,s}_{\rm AC}\left([0,T],\mathbb{P};\mathbb{R}^k\right)$, $0 \leqslant s < T$. A functional $Z(s)$ belongs to the space $\mathcal{M}^{2,s}_{\rm AC}\left([0,T],\mathbb{P};\mathbb{R}^k\right)$ if there exists a martingale $M \in \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ such that for all $N \in \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ the limit

$$Z(s)(N) = \lim_{h\downarrow 0}\frac{\langle M,N\rangle(s+h) - \langle M,N\rangle(s)}{h}$$

exists $\mathbb{P}$-almost surely. In order to indicate that the functional $Z(s)$ originates from the martingale M the notation $Z(s) = Z_M(s)$ is used. As in formula (5.50) of Definition 5.7 the space $\mathcal{M}^{2,s}_{\rm AC}\left([0,T],\mathbb{P};\mathbb{R}^k\right)$ will be supplied with the inner-product:

$$\langle Z_M(s), Z_N(s)\rangle_{\mathcal{M}^{2,s}_{\rm AC}} = \mathbb{E}\left[\frac{d}{ds}\langle M,N\rangle(s)\right], \quad M,\ N \in \mathcal{M}^2_{\rm AC}\left([0,T],\mathbb{P};\mathbb{R}^k\right), \tag{5.90}$$

and the $\mathcal{M}^{2,s}_{\mathrm{AC}}$-norm of $Z_M(s)$ is defined by

$$\|Z_M(s)\|_{\mathcal{M}^{2,s}_{\mathrm{AC}}} = \left(\mathbb{E}\left[\frac{d}{ds}\langle M, M\rangle(s)\right]\right)^{1/2}.$$

In the notation we often suppress the dependence on $\mathbb{P}$: $\mathcal{M}^{2,s}_{\mathrm{AC}}\left([0,T],\mathbb{P};\mathbb{R}^k\right)$ is often replaced with $\mathcal{M}^{2,s}_{\mathrm{AC}}\left([0,T];\mathbb{R}^k\right)$ or even $\mathcal{M}^{2,s}_{\mathrm{AC}}$. Let the process M be a k-dimensional martingale in $\mathcal{M}^2\left([0,T],\mathbb{P};\mathbb{R}^k\right)$ (see Definition 5.9). From the proof of Lemma 5.1 it follows that it makes sense to write $\frac{d}{dt}\left\langle M^j, M^j\right\rangle(t)$. The expression $\frac{d}{dt}\langle M, M\rangle(t)$, is shorthand for

$$\frac{d}{dt}\langle M, M\rangle(t) = \sum_{j=1}^{k}\frac{d}{dt}\left\langle M^j, M^j\right\rangle(t).$$

The following theorem will be employed to prove continuity of solutions to BSDE's. It also implies that BSDE's as considered by us possess at most unique solutions. The variables (Y, M) and (Y', M') attain their values in $\mathbb{R}^k \times \mathbb{R}^k$ endowed with its Euclidean inner-product $\langle y', y\rangle = \sum_{j=1}^k y'_j y_j$, y', $y \in \mathbb{R}^k$. Processes of the form $s \mapsto f(s, Y(s), Z_M(s))$ are progressively measurable processes whenever the pair (Y, M) belongs to the space mentioned in (5.91) of the next theorem.

Theorem 5.3. *Let the pairs (Y, M) and (Y', M'), which belong to the space*

$$L^2\left([0,T]\times\Omega, \mathcal{F}^0_T, dt\times\mathbb{P}\right)\times\mathcal{M}^2\left(\Omega, \mathcal{F}^0_T, \mathbb{P}\right), \tag{5.91}$$

and are $\mathbb{P}$-almost surely continuous, be solutions to the following BSDE's:

$$Y(t) = Y(T) + \int_t^T f(s, Y(s), Z_M(s))\,ds + M(t) - M(T), \quad \textit{and} \tag{5.92}$$

$$Y'(t) = Y'(T) + \int_t^T f'\left(s, Y'(s), Z_{M'}(s)\right)ds + M'(t) - M'(T) \tag{5.93}$$

for $0 \leqslant t \leqslant T$. In particular this means that the processes (Y, M) and (Y', M') are progressively measurable and are square integrable. Suppose that the coefficient f' satisfies the following monotonicity and Lipschitz condition. There exist some positive and finite constants C'_1 and C'_2 such that the following inequalities hold for all $0 \leqslant t \leqslant T$:

$$\begin{aligned}&\left\langle Y'(t) - Y(t), f'\left(t, Y'(t), Z_{M'}(t)\right) - f'\left(t, Y(t), Z_{M'}(t)\right)\right\rangle\\ &\leqslant \left(C'_1\right)^2\left|Y'(t) - Y(t)\right|^2, \quad \textit{and}\\ &\left|f'\left(t, Y(t), Z_{M'}(t)\right) - f'\left(t, Y(t), Z_M(t)\right)\right|^2\end{aligned} \tag{5.94}$$

$$\leqslant \left(C_2'\right)^2 \frac{d}{dt} \left\langle M' - M, M' - M \right\rangle (t). \tag{5.95}$$

Then the pair $(Y' - Y, M' - M)$ *belongs to*

$$\mathcal{S}^2 \left(\Omega, \mathcal{F}_T^0, \mathbb{P}; \mathbb{R}^k\right) \times \mathcal{M}^2 \left(\Omega, \mathcal{F}_T^0, \mathbb{P}; \mathbb{R}^k\right),$$

and there exists a constant C' *which depends on* C_1', C_2' *and* T *such that*

$$\begin{aligned} &\mathbb{E}\left[\sup_{0<t<T} \left|Y'(t) - Y(t)\right|^2 + \left\langle M' - M, M' - M \right\rangle (T)\right] \\ &\leqslant C'\mathbb{E}\left[\left|Y'(T) - Y(T)\right|^2 \right. \\ &\qquad \left. + \int_0^T \left|f'\left(s, Y(s), Z_M(s)\right) - f\left(s, Y(s), Z_M(s)\right)\right|^2 ds\right]. \end{aligned} \tag{5.96}$$

The functionals $Z_M(t)$ *and* $Z_{M'}(t)$ *belong to the space*

$$\mathcal{M}_{\mathrm{AC}}^{2,t}\left([0,T], \mathbb{P}; \mathbb{R}^k\right) = \mathcal{M}_{\mathrm{AC}}^{2,t}\left([0,T]; \mathbb{R}^k\right), \quad 0 \leqslant t \leqslant T.$$

Remark 5.16. From the proof it follows that for C' we may choose $C' = 260e^{\gamma T}$, where $\gamma = 1 + 2\left(C_1'\right)^2 + 2\left(C_2'\right)^2$.

By taking $Y(T) = Y'(T)$ and $f\left(s, Y(s), Z_M(s)\right) = f'\left(s, Y(s), Z_M(s)\right)$ it also implies that BSDE's as considered by us possess at most unique solutions. A precise formulation reads as follows.

Corollary 5.2. *Suppose that the coefficient* f *satisfies the monotonicity condition (5.94) and the Lipschitz condition (5.95). Then there exists at most one* $\mathbb{P}$*-almost surely continuous pair* $(Y, M) \in L^2\left([0,T] \times \Omega, \mathcal{F}_T^0, dt \times \mathbb{P}\right) \times \mathcal{M}^2\left(\Omega, \mathcal{F}_T^0, \mathbb{P}\right)$ *which satisfies the backward stochastic differential equation in (5.92).*

Proof. [Proof of Theorem 5.3.] Put $\overline{Y} = Y' - Y$ and $\overline{M} = M' - M$. From Itô's formula it follows that

$$\begin{aligned} &\left|\overline{Y}(t)\right|^2 + \left\langle \overline{M}, \overline{M} \right\rangle (T) - \left\langle \overline{M}, \overline{M} \right\rangle (t) \\ &= \left|\overline{Y}(T)\right|^2 + 2\int_t^T \left\langle \overline{Y}(s), f'\left(s, Y'(s), Z_{M'}(s)\right) - f'\left(s, Y(s), Z_{M'}(s)\right)\right\rangle ds \\ &\quad + 2\int_t^T \left\langle \overline{Y}(s), f'\left(s, Y(s), Z_{M'}(s)\right) - f'\left(s, Y(s), Z_M(s)\right)\right\rangle ds \\ &\quad + 2\int_t^T \left\langle \overline{Y}(s), f'\left(s, Y(s), Z_M(s)\right) - f\left(s, Y(s), Z_M(s)\right)\right\rangle ds \end{aligned}$$

$$- 2\int_t^T \left\langle \overline{Y}(s), d\overline{M}(s)\right\rangle. \tag{5.97}$$

(Notice that in the left-hand side of (5.97) the brackets $\langle\cdot,\cdot\rangle$ denote the increment of the variation process of the $\mathbb{P}$-almost sure continuous martingale $M \in \mathcal{M}^2\left([0,T];\mathbb{R}^k\right)$, and that in the right-hand side the brackets denote an inner-product in $\mathbb{R}^k$.) Inserting the inequalities (5.94) and (5.95) into (5.97) shows:

$$\begin{aligned}
&\left|\overline{Y}(t)\right|^2 + \left\langle \overline{M},\overline{M}\right\rangle(T) - \left\langle \overline{M},\overline{M}\right\rangle(t) \\
&\leqslant \left|\overline{Y}(T)\right|^2 + 2\left(C_1'\right)^2\int_t^T \left|\overline{Y}(s)\right|^2 ds + 2C_2'\int_t^T \left|\overline{Y}(s)\right|\left(\frac{d}{ds}\left\langle \overline{M},\overline{M}\right\rangle(s)\right)^{1/2} ds \\
&\quad + 2\int_t^T \left|\overline{Y}(s)\right|\left|f'\left(s,Y(s),Z_M(s)\right) - f\left(s,Y(s),Z_M(s)\right)\right| ds \\
&\quad - 2\int_t^T \left\langle \overline{Y}(s), d\overline{M}(s)\right\rangle.
\end{aligned} \tag{5.98}$$

The elementary inequalities $2ab \leqslant 2C_2'a^2 + \dfrac{b^2}{2C_2'}$ and $2ab \leqslant a^2+b^2$, $0 \leqslant a$, $b \in \mathbb{R}$, apply to the effect that

$$\begin{aligned}
&\left|\overline{Y}(t)\right|^2 + \frac{1}{2}\left(\left\langle \overline{M},\overline{M}\right\rangle(T) - \left\langle \overline{M},\overline{M}\right\rangle(t)\right) \\
&\leqslant \left|\overline{Y}(T)\right|^2 + \left(1 + 2\left(C_1'\right)^2 + 2\left(C_2'\right)^2\right)\int_t^T \left|\overline{Y}(s)\right|^2 ds \\
&\quad + \int_t^T \left|f'\left(s,Y(s),Z_M(s)\right) - f\left(s,Y(s),Z_M(s)\right)\right|^2 ds \\
&\quad - 2\int_0^T \left\langle \overline{Y}(s), d\overline{M}(s)\right\rangle + 2\int_0^t \left\langle \overline{Y}(s), d\overline{M}(s)\right\rangle.
\end{aligned} \tag{5.99}$$

For a concise formulation of the relevant inequalities we introduce the following functions and the constant γ:

$$\begin{aligned}
A_{\overline{Y}}(t) &= \mathbb{E}\left[\left|\overline{Y}(t)\right|^2\right], \\
A_{\overline{M}}(t) &= \mathbb{E}\left[\left\langle \overline{M},\overline{M}\right\rangle(T) - \left\langle \overline{M},\overline{M}\right\rangle(t)\right], \\
C(s) &= \mathbb{E}\left[\left|f'\left(s,Y(s),Z_M(s)\right) - f\left(s,Y(s),Z_M(s)\right)\right|^2\right], \\
B(t) &= A_{\overline{Y}}(T) + \int_t^T C(s)ds = B(T) + \int_t^T C(s)ds, \quad \text{and} \\
\gamma &= 1 + 2\left(C_1'\right)^2 + 2\left(C_2'\right)^2.
\end{aligned} \tag{5.100}$$

Using the quantities in (5.100) and remembering the fact that the final term in (5.99) represents a martingale difference, the inequality in (5.99) implies:

$$A_{\overline{Y}}(t)+\frac{1}{2}A_{\overline{M}}(t)\leqslant B(t)+\gamma\int_t^T A_{\overline{Y}}(s)ds. \tag{5.101}$$

Using (5.101) and employing induction with respect to n yields:

$$\begin{aligned}&A_{\overline{Y}}(t)+\frac{1}{2}A_{\overline{M}}(t) \qquad (5.102)\\ &\leqslant B(t)+\int_t^T\sum_{k=0}^n\frac{\gamma^{k+1}(T-s)^k}{k!}B(s)ds+\int_t^T\frac{\gamma^{n+2}(T-s)^{n+1}}{(n+1)!}A_{\overline{Y}}(s)ds.\end{aligned}$$

Passing to the limit for $n\to\infty$ in (5.102) results in:

$$A_{\overline{Y}}(t)+\frac{1}{2}A_{\overline{M}}(t)\leqslant B(t)+\gamma\int_t^T e^{\gamma(T-s)}B(s)ds. \tag{5.103}$$

Since $B(t)=A_{\overline{Y}}(T)+\int_t^T C(s)ds$ from (5.103) we infer:

$$A_{\overline{Y}}(t)+\frac{1}{2}A_{\overline{M}}(t)\leqslant e^{\gamma(T-t)}\left(A_{\overline{Y}}(T)+\int_t^T C(s)ds\right). \tag{5.104}$$

By first taking the supremum over $0<t<T$ and then taking expectations in (5.99) gives:

$$\begin{aligned}&\mathbb{E}\left[\sup_{0<t<T}\left|\overline{Y}(t)\right|^2\right]\\ &\leqslant \mathbb{E}\left[\left|\overline{Y}(T)\right|^2\right]+\left(1+2\left(C_1'\right)^2+2\left(C_2'\right)^2\right)\int_0^T\mathbb{E}\left[\left|\overline{Y}(s)\right|^2\right]ds\\ &\quad+\int_0^T\mathbb{E}\left[\left|f'\left(s,Y(s),Z_M(s)\right)-f\left(s,Y(s),Z_M(s)\right)\right|^2\right]ds\\ &\quad+2\mathbb{E}\left[\sup_{0<t<T}\int_0^t\left\langle\overline{Y}(s),d\overline{M}(s)\right\rangle\right]. \qquad (5.105)\end{aligned}$$

The quadratic variation process of the martingale $t\mapsto\int_0^t\left\langle\overline{Y}(s),d\overline{M}(s)\right\rangle$ is given by the increasing process $t\mapsto\sum_{j_1=1}^k\sum_{j_2=1}^k\int_0^t Y_{j_1}(s)Y_{j_2}(s)\,d\left\langle M_{j_1},M_{j_2}\right\rangle(s)$ which is dominated by the process $t\mapsto\int_0^t\left|\overline{Y}(s)\right|^2 d\left\langle\overline{M},\overline{M}\right\rangle(s)$. The inequality $\sum_{j_1=1}^k\sum_{j_2=1}^k\int_0^t Y_{j_1}(s)Y_{j_2}(s)\,d\left\langle M_{j_1},M_{j_2}\right\rangle(s)\leqslant\int_0^t|Y(s)|^2\,d\langle M,M\rangle(s)$ follows from inequalities of the form $(1\leqslant j_1,\,j_2\leqslant k)$

$$2\int_0^t Y_{j_1}(s)Y_{j_2}(s)d\left\langle M_{j_1},M_{j_2}\right\rangle(s)$$

$$\leqslant \int_0^t |Y_{j_1}(s)|^2 \, d\langle M_{j_2}, M_{j_2}\rangle(s) + \int_0^t |Y_{j_2}(s)|^2 \, d\langle M_{j_1}, M_{j_1}\rangle(s). \tag{5.106}$$

Here $Y = (Y_1, \dots, Y_k)$, $M = (M_1, \dots, M_k)$, $|Y|^2 = \sum_{j=1}^k |Y_j|^2$, and by definition $\langle M, M\rangle(t) = \sum_{j=1}^k \langle M_j, M_j\rangle(t)$. From the Burkholder-Davis-Gundy inequality (5.89) we know that

$$\mathbb{E}\left[\sup_{0<t<T} \int_0^t \left\langle \overline{Y}(s), d\overline{M}(s)\right\rangle\right] \leqslant 4\sqrt{2}\mathbb{E}\left[\left(\int_0^T \left|\overline{Y}(s)\right|^2 d\left\langle \overline{M}, \overline{M}\right\rangle(s)\right)^{1/2}\right]. \tag{5.107}$$

(For more details on the Burkholder-Davis-Gundy inequality, see e.g. [Ikeda and Watanabe (1998)].) Again we use an elementary inequality $4\sqrt{2}ab \leqslant \frac{1}{4}a^2 + 32b^2$ and plug it into (5.107) to obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T} \int_0^t \left\langle \overline{Y}(s), d\overline{M}(s)\right\rangle\right] \\
&\leqslant 4\sqrt{2}\mathbb{E}\left[\sup_{0<t<T} \left|\overline{Y}(t)\right| \left(\int_0^T d\left\langle \overline{M}, \overline{M}\right\rangle(s)\right)^{1/2}\right] \\
&\leqslant \frac{1}{4}\mathbb{E}\left[\sup_{0<t<T} \left|\overline{Y}(t)\right|^2\right] + 32\mathbb{E}\left[\left\langle \overline{M}, \overline{M}\right\rangle(T)\right].
\end{aligned} \tag{5.108}$$

From (5.104) we also infer

$$\begin{aligned}
\gamma \int_0^T A_{\overline{Y}}(s) ds &\leqslant \gamma \int_0^T e^{\gamma(T-s)} \left(A_{\overline{Y}}(T) + \int_s^T C(\rho) d\rho\right) ds \\
&= \left(e^{\gamma T} - 1\right) A_{\overline{Y}}(T) + \int_0^T \left(e^{\gamma T} - e^{\gamma \rho}\right) C(\rho) d\rho.
\end{aligned} \tag{5.109}$$

Inserting the inequalities (5.108) and (5.109) into (5.105) yields:

$$\begin{aligned}
\mathbb{E}\left[\sup_{0<t<T} \left|\overline{Y}(t)\right|^2\right] &\leqslant e^{\gamma T}\mathbb{E}\left[\left|\overline{Y}(T)\right|^2\right] + e^{\gamma T} \int_0^T C(s) ds + \frac{1}{2}\mathbb{E}\left[\sup_{0<t<T} \left|\overline{Y}(t)\right|^2\right] \\
&\quad + 64\mathbb{E}\left[\left\langle \overline{M}, \overline{M}\right\rangle(T)\right].
\end{aligned} \tag{5.110}$$

From (5.104) we also get

$$\begin{aligned}
&\mathbb{E}\left[\left\langle \overline{M}, \overline{M}\right\rangle(T)\right] = A_{\overline{M}}(0) \qquad (5.111) \\
&\leqslant 2e^{\gamma T}\left(A_{\overline{Y}}(T) + \int_0^T C(s) ds\right) = 2e^{\gamma T}\left(\mathbb{E}\left[\left|\overline{Y}(T)\right|^2\right] + \int_0^T C(s) ds\right).
\end{aligned}$$

A combination of (5.111) and (5.110) results in

$$\mathbb{E}\left[\sup_{0<t<T}\left|\overline{Y}(t)\right|^2\right] \leqslant 258e^{\gamma T}\left(\mathbb{E}\left[\left|\overline{Y}(T)\right|^2\right]+\int_0^T C(s)ds\right). \tag{5.112}$$

Adding the right- and left-hand sides of (5.110) and (5.111) proves Theorem 5.3 with the constant C' given by $C' = 260e^{\gamma T}$, where $\gamma = 1 + 2\left(C_1'\right)^2 + 2\left(C_2'\right)^2$.

This completes the proof of Theorem 5.3. □

In the definitions 5.8 and 5.9 the spaces $\mathcal{S}^2\left([0,T],\mathbb{R}^k\right)$ and $\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ are defined.

In Theorem 5.5 we will replace the Lipschitz condition (5.113) in Theorem 5.4 for the function $Y(s) \mapsto f\left(s, Y(s), Z_M(s)\right)$ with the (weaker) monotonicity condition (5.137). Here we write y for the variable $Y(s)$ and z for $Z_M(s)$. It is noticed that we consider a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with a filtration $(\mathcal{F}_t)_{t\in[0,T]} = \left(\mathcal{F}_t^0\right)_{t\in[0,T]}$ where $\mathcal{F}_T = \mathcal{F}$. In Theorem 5.4 for every fixed $s \in [0,T]$ the function $f(s,\cdot,\cdot)$ is defined on $\mathbb{R}^k \times \mathcal{M}^{2,s}_{\mathrm{AC}}\left([0,T],\mathbb{P};\mathbb{R}^k\right)$: see Definition 5.10. Instead of $\mathcal{M}^{2,s}_{\mathrm{AC}}\left([0,T],\mathbb{P};\mathbb{R}^k\right)$ we write $\mathcal{M}^{2,s}_{\mathrm{AC}}$.

Theorem 5.4. *Let $f(s) : \mathbb{R}^k \times \mathcal{M}^{2,s}_{\mathrm{AC}} \to \mathbb{R}^k$, $0 \leqslant s \leqslant T$, be a Lipschitz continuous in the sense that there exists finite constants C_1 and C_2 such that for any two pairs of processes (Y,M) and $(U,N) \in \mathcal{S}^2\left([0,T],\mathbb{R}^k\right) \times \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ the following inequalities hold for all $0 \leqslant s \leqslant T$:*

$$\left|f\left(s,Y(s),Z_M(s)\right) - f\left(s,U(s),Z_M(s)\right)\right| \leqslant C_1\left|Y(s)-U(s)\right|, \text{ and} \tag{5.113}$$

$$\left|f(s,Y(s),Z_M(s)) - f(s,Y(s),Z_N(s))\right| \leqslant C_2\left(\frac{d}{ds}\left\langle M-N,M-N\right\rangle(s)\right)^{1/2}. \tag{5.114}$$

Suppose that $\mathbb{E}\left[\int_0^T \left|f(s,0,0)\right|^2 ds\right] < \infty$. Then there exists a unique pair $(Y,M) \in \mathcal{S}^2\left([0,T],\mathbb{R}^k\right) \times \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ such that

$$Y(t) = \xi + \int_t^T f\left(s,Y(s),Z_M(s)\right)ds + M(t) - M(T), \tag{5.115}$$

where $Y(T) = \xi \in L^2\left(\Omega,\mathcal{F}_T,\mathbb{R}^k\right)$ is given and $Y(0) = M(0)$.

For brevity we write

$$\begin{aligned}\mathcal{S}^2 \times \mathcal{M}^2 &= \mathcal{S}^2\left([0,T],\mathbb{R}^k\right) \times \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)\\ &= \mathcal{S}^2\left(\Omega,\mathcal{F}_T^0,\mathbb{P};\mathbb{R}^k\right) \times \mathcal{M}^2\left(\Omega,\mathcal{F}_T^0,\mathbb{P};\mathbb{R}^k\right).\end{aligned}$$

In fact we employ this theorem with the function f replaced by f_δ, $0 < \delta < (2C_1)^{-1}$, defined by

$$f_\delta\left(s, y, Z_M(s)\right) = f\left(s, \left(I - \delta f_{s,M}\right)^{-1} y, Z_M(s)\right). \tag{5.116}$$

Here $f_{s,M}(y) = f\left(s, y, Z_M(s)\right)$. If the function f is monotone (or one-sided Lipschitz) in the second variable with constant C_1, and Lipschitz in the third variable with constant C_2, then the function f_δ is Lipschitz in y with Lipschitz constant δ^{-1}.

Proof. The proof of the uniqueness part follows from Corollary 5.2.

In order to prove existence we proceed as follows. By induction we define a sequence (Y_n, M_n) in the space $\mathcal{S}^2 \times \mathcal{M}^2$ as follows. Put

$$\widetilde{Y}_{n+1}(t) = \mathbb{E}\left[\xi + \int_t^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\Big|\; \mathcal{F}_t\right], \text{ and} \tag{5.117}$$

$$\widetilde{M}_{n+1}(t) = \mathbb{E}\left[\xi + \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\Big|\; \mathcal{F}_t\right]. \tag{5.118}$$

The processes $t \mapsto \widetilde{Y}_{n+1}(t)$ and $t \mapsto \widetilde{M}_{n+1}(t)$ need not be continuous. It is easy to see that the jumps of the processes $t \mapsto \widetilde{Y}_{n+1}(t)$ and $t \mapsto \widetilde{M}_{n+1}(t)$ coincide. Moreover, if we subtract the jumps from $M_{n+1}(t)$ we still have a martingale, i.e. the process

$$t \mapsto M_{n+1}(t) := \widetilde{M}_{n+1}(t) - \sum_{s \leqslant t}\left(\widetilde{M}_{n+1}(s) - \widetilde{M}_{n+1}(s-)\right) \tag{5.119}$$

is still a martingale which belongs to $\mathcal{M}^2$. In particular it is continuous. In the same manner we introduce the process $Y_{n+1}(t)$:

$$\begin{aligned} t \mapsto Y_{n+1}(t) &:= \widetilde{Y}_{n+1}(t) - \sum_{s \leqslant t}\left(\widetilde{M}_{n+1}(s) - \widetilde{M}_{n+1}(s-)\right) \\ &= \widetilde{Y}_{n+1}(t) - \sum_{s \leqslant t}\left(\widetilde{Y}_{n+1}(s) - \widetilde{Y}_{n+1}(s-)\right). \end{aligned} \tag{5.120}$$

By construction the processes M_{n+1} and Y_{n+1} are continuous. Since by assumption the variable $\xi + \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds$ belongs to $L^2\left(\Omega, \mathcal{F}_T, \mathbb{P}\right)$ it follows that the martingale M_{n+1} belongs to $L^2\left(\Omega, \mathcal{F}_T, \mathbb{P}\right)$. Moreover, using the fact that the process $s \mapsto f\left(s, Y_n(s), M_n(s)\right)$ is adapted we have:

$$\begin{aligned} &\xi + \int_t^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds + \widetilde{M}_{n+1}(t) - \widetilde{M}_{n+1}(T) \\ &= \xi + \int_t^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds + \mathbb{E}\left[\xi + \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\Big|\; \mathcal{F}_t\right] \end{aligned}$$

$$
\begin{aligned}
&- \mathbb{E}\left[\xi + \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\middle|\; \mathcal{F}_T\right] \\
&= \xi + \int_t^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds + \mathbb{E}\left[\xi + \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\middle|\; \mathcal{F}_t\right] \\
&- \xi - \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \\
&= \mathbb{E}\left[\xi + \int_0^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\middle|\; \mathcal{F}_t\right] - \int_0^t f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \\
&= \mathbb{E}\left[\xi + \int_t^T f\left(s, Y_n(s), Z_{M_n}(s)\right) ds \;\middle|\; \mathcal{F}_t\right] = \widetilde{Y}_{n+1}(t). \qquad (5.121)
\end{aligned}
$$

Since $\xi = Y_{n+1}(T)$ (5.121) implies that the jump parts of the process Y_{n+1} occur at the same time instances as those of M_{n+1}. As a consequence these jump parts cancel each other. So without loss of generality we assume that the processes Y_{n+1} and M_{n+1} are $\mathbb{P}$-almost surely continuous, and that (5.121) is satisfied with $Y_{n+1}(t)$ and $M_{n+1}(t)$ instead of $\widetilde{Y}_{n+1}(t)$ and $\widetilde{M}_{n+1}(t)$ respectively.

Suppose that the pair (Y_n, M_n) belongs $\mathcal{S}^2 \times \mathcal{M}^2$. We first prove that the pair (Y_{n+1}, M_{n+1}) is a member of $\mathcal{S}^2 \times \mathcal{M}^2$. As explained above we may and do assume that the pair of precesses $t \mapsto (Y_{n+1}(t), M_{n+1}(t))$ is continuous $\mathbb{P}$-almost surely. Therefore we fix $\alpha = 1 + C_1^2 + C_2^2 \in \mathbb{R}$ where C_1 and C_2 are as in (5.113) and (5.114) respectively. From Itô's formula we get:

$$
\begin{aligned}
&e^{2\alpha t} \left|Y_{n+1}(t)\right|^2 + 2\alpha \int_t^T e^{2\alpha s} \left|Y_{n+1}(s)\right|^2 ds + \int_t^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s) \\
&= e^{2\alpha T} \left|Y_{n+1}(T)\right|^2 \\
&\quad + 2\int_t^T e^{2\alpha s} \left\langle Y_{n+1}(s), f\left(s, Y_n(s), Z_{M_n}(s)\right) - f\left(s, Y_n(s), 0\right)\right\rangle ds \\
&\quad + 2\int_t^T e^{2\alpha s} \left\langle Y_{n+1}(s), f\left(s, Y_n(s), 0\right) - f\left(s, 0, 0\right)\right\rangle ds \\
&\quad + 2\int_t^T e^{2\alpha s} \left\langle Y_{n+1}(s), f\left(s, 0, 0\right)\right\rangle ds - 2\int_t^T e^{2\alpha s} \left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle.
\end{aligned}
\qquad (5.122)
$$

(Notice that the bracket in the left-hand side of (5.122) denotes the variation process of the k-dimensional martingale M, and that the brackets in the right-hand side denote inner-products in $\mathbb{R}^k$.) We employ (5.113) and (5.114) to obtain from (5.122):

$$\begin{aligned} & e^{2\alpha t}\left|Y_{n+1}(t)\right|^2 + 2\alpha\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \int_t^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s) \\ & \leqslant e^{2\alpha T}\left|Y_{n+1}(T)\right|^2 + 2C_2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|\left(\frac{d}{ds}\left\langle M_n, M_n\right\rangle(s)\right)^{1/2} ds \\ & \quad + 2C_1\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|\left|Y_n(s)\right| ds \\ & \quad + 2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|\left|f(s,0,0)\right| ds - 2\int_t^T e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle. \end{aligned} \tag{5.123}$$

The elementary inequalities $2ab \leqslant 2C_j a^2 + \dfrac{b^2}{2C_j}$, $a, b \in \mathbb{R}$, $j = 0, 1, 2$, with $C_0 = 1$, in combination with (5.123) yield

$$\begin{aligned} & e^{2\alpha t}\left|Y_{n+1}(t)\right|^2 + 2\alpha\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \int_t^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s) \\ & \leqslant e^{2\alpha T}\left|Y_{n+1}(T)\right|^2 + 2C_2^2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \frac{1}{2}\int_t^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s) \\ & \quad + 2C_1^2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \frac{1}{2}\int_t^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds \\ & \quad + \int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \int_t^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds \\ & \quad - 2\int_t^T e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle, \end{aligned} \tag{5.124}$$

and hence by the choice of α from (5.124) we infer:

$$\begin{aligned} & e^{2\alpha t}\left|Y_{n+1}(t)\right|^2 + \int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \int_t^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s) \\ & \quad + 2\int_0^T e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle \\ & \leqslant e^{2\alpha T}\left|Y_{n+1}(T)\right|^2 + \frac{1}{2}\int_t^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s) + \frac{1}{2}\int_t^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds \\ & \quad + \int_t^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds + 2\int_0^t e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle. \end{aligned} \tag{5.125}$$

The following steps can be justified by observing that the process Y_{n+1} belongs to the space $L^2\left(\Omega, \mathcal{F}_T^0, \mathbb{P}\right)$, and that $\sup_{0\leqslant t\leqslant T}\left|Y_{n+1}(t)\right| < \infty$ $\mathbb{P}$-almost surely. By stopping the process $Y_{n+1}(t)$ at the stopping time τ_N being the first time $t \leqslant T$ that $|Y_{n+1}(t)|$ exceeds N. In inequality (5.125) we then replace t by $t \wedge \tau_N$, T by τ_N, and proceed as below with the stopped processes instead of the processes itself. Then we use the monotone convergence theorem to obtain inequality (5.128). By the same approximation argument we may assume that $\mathbb{E}\left[\int_t^T e^{2\alpha s} \langle Y_{n+1}(s), dM_{n+1}(s)\rangle\right] = 0$. Hence (5.125) implies that

$$\begin{aligned}
&\mathbb{E}\left[e^{2\alpha t}\left|Y_{n+1}(t)\right|^2 + \int_t^T e^{2\alpha s}\left|Y_{n+1}(s)\right|^2 ds + \int_t^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s)\right] \\
&\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)\right|^2\right] + \frac{1}{2}\mathbb{E}\left[\int_t^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_t^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] + \mathbb{E}\left[\int_t^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] < \infty. \quad (5.126)
\end{aligned}$$

Invoking the Burkholder-Davis-Gundy inequality with $p = \frac{1}{2}$ (see inequality (5.89)) and applying the inequality (see inequality (5.106))

$$\begin{aligned}
&\left\langle \int_0^{\cdot} e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle, \int_0^{\cdot} e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle \right\rangle(t) \\
&\leqslant \int_0^t e^{4\alpha s}\left|Y_{n+1}(s)\right|^2 d\left\langle M_{n+1}, M_{n+1}\right\rangle(s)
\end{aligned}$$

to (5.125) yields:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T} e^{2\alpha t}\left|Y_{n+1}(t)\right|^2\right] \\
&\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)\right|^2\right] + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] \\
&\quad + \mathbb{E}\left[\int_0^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] - 2\mathbb{E}\left[\int_0^T e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle\right] \\
&\quad + 8\sqrt{2}\mathbb{E}\left[\left(\int_0^T e^{4\alpha s}\left|Y_{n+1}(s)\right|^2 d\left\langle M_{n+1}, M_{n+1}\right\rangle(s)\right)^{1/2}\right]
\end{aligned}$$

(without loss of generality assume that $\mathbb{E}\left[\int_0^T e^{2\alpha s}\left\langle Y_{n+1}(s), dM_{n+1}(s)\right\rangle\right] = 0$; this can be achieved by localization)

$$\begin{aligned}
&\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)\right|^2\right] + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] + \mathbb{E}\left[\int_0^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] \\
&\quad + 8\sqrt{2}\mathbb{E}\left[\sup_{0<t<T} e^{\alpha t}\left|Y_{n+1}(t)\right|\left(\int_0^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s)\right)^{1/2}\right]
\end{aligned}$$

($8\sqrt{2}ab \leqslant \dfrac{a^2}{2} + 64b^2$, a, $b \in \mathbb{R}$)

$$\begin{aligned}
&\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)\right|^2\right] + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] + \mathbb{E}\left[\int_0^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\sup_{0<t<T} e^{2\alpha t}\left|Y_{n+1}(t)\right|^2\right] + 64\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_{n+1}, M_{n+1}\right\rangle(s)\right]
\end{aligned}$$

(apply (5.126))

$$\begin{aligned}
&\leqslant 65 e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)\right|^2\right] + \frac{65}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] \\
&\quad + \frac{65}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] \\
&\quad + 65\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] + \frac{1}{2}\mathbb{E}\left[\sup_{0<t<T} e^{2\alpha t}\left|Y_{n+1}(t)\right|^2\right]. \qquad (5.127)
\end{aligned}$$

From (5.127) it follows that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T} e^{2\alpha t}\left|Y_{n+1}(t)\right|^2\right] \\
&\quad \leqslant 130 e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)\right|^2\right] + 130\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] \qquad (5.128)
\end{aligned}$$

$$+ 65\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] + 65\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] < \infty.$$

From (5.126) and (5.128) it follows that the pair (Y_{n+1}, M_{n+1}) belongs to $\mathcal{S}^2 \times \mathcal{M}^2$.

Another application of Itô's formula shows:

$$\begin{aligned}
&e^{2\alpha t}\left|Y_{n+1}(t) - Y_n(t)\right|^2 + 2\alpha\int_t^T e^{2\alpha s}\left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad + \int_t^T e^{2\alpha s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle(s) \\
&= e^{2\alpha T}\left|Y_{n+1}(T) - Y_n(T)\right|^2 \\
&\quad + 2\int_t^T e^{2\alpha s}\left\langle \triangle Y_n(s), f\left(s, Y_n(s), Z_{M_n}(s)\right) - f\left(s, Y_n(s), Z_{M_{n-1}}(s)\right)\right\rangle ds \\
&\quad + 2\int_t^T e^{2\alpha s}\left\langle \triangle Y_n(s), f(s, Y_n(s), Z_{M_{n-1}}(s)) - f(s, Y_{n-1}(s), Z_{M_{n-1}}(s))\right\rangle ds \\
&\quad - 2\int_t^T e^{2\alpha s}\left\langle Y_{n+1}(s) - Y_n(s), dM_{n+1}(s) - dM_n(s)\right\rangle, \qquad (5.129)
\end{aligned}$$

where for brevity we wrote $\triangle Y_n(s) = Y_{n+1}(s) - Y_n(s)$. From (5.113), (5.114), and (5.129) we infer

$$\begin{aligned}
&e^{2\alpha t}\left|Y_{n+1}(t) - Y_n(t)\right|^2 + 2\alpha\int_t^T e^{2\alpha s}\left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad + \int_t^T e^{2\alpha s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle(s) \\
&\leqslant e^{2\alpha T}\left|Y_{n+1}(T) - Y_n(T)\right|^2 \\
&\quad + 2C_2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s) - Y_n(s)\right|\left(\frac{d}{ds}\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle(s)\right)^{1/2} ds \\
&\quad + 2C_1\int_t^T e^{2\alpha s}\left|Y_{n+1}(s) - Y_n(s)\right|\left|Y_n(s) - Y_{n-1}(s)\right| ds \\
&\quad - 2\int_t^T e^{2\alpha s}\left\langle Y_{n+1}(s) - Y_n(s), dM_{n+1}(s) - dM_n(s)\right\rangle \\
&\leqslant e^{2\alpha T}\left|Y_{n+1}(T) - Y_n(T)\right|^2 + 2C_2^2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad + \frac{1}{2}\int_t^T e^{2\alpha s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle(s) \\
&\quad + 2C_1^2\int_t^T e^{2\alpha s}\left|Y_{n+1}(s) - Y_n(s)\right|^2 ds + \frac{1}{2}\int_t^T e^{2\alpha s}\left|Y_n(s) - Y_{n-1}(s)\right|^2 ds
\end{aligned}$$

$$- 2\int_t^T e^{2\alpha s} \left\langle Y_{n+1}(s) - Y_n(s), dM_{n+1}(s) - dM_n(s)\right\rangle . \tag{5.130}$$

Since $\alpha = 1 + C_1^2 + C_2^2$ the inequality in (5.130) implies:

$$\begin{aligned}
&e^{2\alpha t} \left|Y_{n+1}(t) - Y_n(t)\right|^2 + 2\int_t^T e^{2\alpha s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad + \int_t^T e^{2\alpha s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle(s) \\
&\leqslant e^{2\alpha T} \left|Y_{n+1}(T) - Y_n(T)\right|^2 \\
&\quad + \frac{1}{2}\int_t^T e^{2\alpha s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle(s) \\
&\quad + \frac{1}{2}\int_t^T e^{2\alpha s} \left|Y_n(s) - Y_{n-1}(s)\right|^2 ds \\
&\quad - 2\int_0^T e^{2\alpha s} \left\langle Y_{n+1}(s) - Y_n(s), dM_{n+1}(s) - dM_n(s)\right\rangle \\
&\quad + 2\int_0^t e^{2\alpha s} \left\langle Y_{n+1}(s) - Y_n(s), dM_{n+1}(s) - dM_n(s)\right\rangle .
\end{aligned} \tag{5.131}$$

Upon taking expectations in (5.131) we see

$$\begin{aligned}
&e^{2\alpha t}\mathbb{E}\left[\left|Y_{n+1}(t) - Y_n(t)\right|^2\right] + 2\mathbb{E}\left[\int_t^T e^{2\alpha s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds\right] \\
&\quad + \mathbb{E}\left[\int_t^T e^{2\alpha s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle(s)\right] \\
&\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T) - Y_n(T)\right|^2\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_t^T e^{2\alpha s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle(s)\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_t^T e^{2\alpha s} \left|Y_n(s) - Y_{n-1}(s)\right|^2 ds\right].
\end{aligned} \tag{5.132}$$

In particular it follows that

$$\begin{aligned}
&2\mathbb{E}\left[\int_t^T e^{2\alpha s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds\right] \\
&\quad + \mathbb{E}\left[\int_t^T e^{2\alpha s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle(s)\right]
\end{aligned}$$

$$\leqslant \frac{1}{2}\mathbb{E}\left[\int_t^T e^{2\alpha s}\left|Y_n(s)-Y_{n-1}(s)\right|^2 ds\right]$$
$$+\frac{1}{2}\mathbb{E}\left[\int_t^T e^{2\alpha s} d\left\langle M_n-M_{n-1}, M_n-M_{n-1}\right\rangle(s)\right],$$

provided that $Y_{n+1}(T)=Y_n(T)$. As a consequence we see that the sequence (Y_n, M_n) converges with respect to the norm $\|\cdot\|_\alpha$ defined by

$$\left\|\begin{pmatrix} Y \\ M \end{pmatrix}\right\|_\alpha^2=\mathbb{E}\left[\int_0^T e^{2\alpha s}|Y(s)|^2 ds+\int_0^T e^{2\alpha s} d\langle M, M\rangle(s)\right].$$

Employing a similar reasoning as the one we used to obtain (5.127) and (5.128) from (5.131) we also obtain:

$$\begin{aligned}
&\sup_{0\leqslant t\leqslant T} e^{2\alpha t}\left|Y_{n+1}(t)-Y_n(t)\right|^2 \\
&\leqslant e^{2\alpha T}\left|Y_{n+1}(T)-Y_n(T)\right|^2 \\
&\quad+\frac{1}{2}\int_0^T e^{2\alpha s} d\left\langle M_n-M_{n-1}, M_n-M_{n-1}\right\rangle(s) \\
&\quad+\frac{1}{2}\int_0^T e^{2\alpha s}\left|Y_n(s)-Y_{n-1}(s)\right|^2 ds \\
&\quad-2\int_0^T e^{2\alpha s}\left\langle Y_{n+1}(s)-Y_n(s), dM_{n+1}(s)-dM_n(s)\right\rangle \\
&\quad+2\sup_{0\leqslant t\leqslant T}\int_0^t e^{2\alpha s}\left\langle Y_{n+1}(s)-Y_n(s), dM_{n+1}(s)-dM_n(s)\right\rangle.
\end{aligned} \tag{5.133}$$

By taking expectations in (5.133), and invoking the Burkholder-Davis-Gundy inequality (5.89) for $p=\frac{1}{2}$ we obtain:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} e^{2\alpha t}\left|Y_{n+1}(t)-Y_n(t)\right|^2\right] \\
&\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right] \\
&\quad+\frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n-M_{n-1}, M_n-M_{n-1}\right\rangle(s)\right] \\
&\quad+\frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)-Y_{n-1}(s)\right|^2 ds\right] \\
&\quad+2\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\int_0^t e^{2\alpha s}\left\langle Y_{n+1}(s)-Y_n(s), dM_{n+1}(s)-dM_n(s)\right\rangle\right]
\end{aligned}$$

$$\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right]$$
$$+\frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_n-M_{n-1},M_n-M_{n-1}\right\rangle(s)\right]$$
$$+\frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_n(s)-Y_{n-1}(s)\right|^2 ds\right]$$
$$+8\sqrt{2}\mathbb{E}\left[\left(\int_0^T e^{4\alpha s}\left|Y_{n+1}(s)-Y_n(s)\right|^2 d\left\langle M_{n+1}-M_n,M_{n+1}-M_n\right\rangle(s)\right)^{1/2}\right]$$

(insert the definition of $\|\cdot\|_\alpha$)

$$\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right]+\frac{1}{2}\left\|\begin{pmatrix} Y_n-Y_{n-1}\\ M_n-M_{n-1}\end{pmatrix}\right\|_\alpha^2$$
$$+8\sqrt{2}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{\alpha s}\left|Y_{n+1}(s)-Y_n(s)\right|\right.$$
$$\left.\times\left(\int_0^T e^{2\alpha s} d\left\langle M_{n+1}-M_n,M_{n+1}-M_n\right\rangle(s)\right)^{1/2}\right]$$

($8\sqrt{2}ab\leqslant\frac{1}{2}a^2+64b^2$, a, $b\in\mathbb{R}$)

$$\leqslant e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right]+\frac{1}{2}\left\|\begin{pmatrix} Y_n-Y_{n-1}\\ M_n-M_{n-1}\end{pmatrix}\right\|_\alpha^2$$
$$+\frac{1}{2}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{2\alpha s}\left|Y_{n+1}(s)-Y_n(s)\right|^2\right]$$
$$+64\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_{n+1}-M_n,M_{n+1}-M_n\right\rangle(s)\right]. \tag{5.134}$$

Employing inequality (5.132) (with $t=0$) together with (5.134), and the definition of the norm $\|\cdot\|_\alpha$ yields the inequality

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} e^{2\alpha t}\left|Y_{n+1}(t)-Y_n(t)\right|^2\right]+129\mathbb{E}\left[\int_0^T e^{2\alpha s}\left|Y_{n+1}(s)-Y_n(s)\right|^2 ds\right]$$
$$+\frac{1}{2}\mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_{n+1}-M_n,M_{n+1}-M_n\right\rangle(s)\right]$$

$$\leqslant \frac{131}{2}e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right]+\frac{131}{4}\left\|\begin{pmatrix} Y_n - Y_{n-1} \\ M_n - M_{n-1}\end{pmatrix}\right\|_\alpha^2 + \frac{1}{2}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{2\alpha s}\left|Y_{n+1}(s)-Y_n(s)\right|^2\right]. \tag{5.135}$$

(In order to justify the transition from (5.133) to (5.135) like in passing from inequality (5.125) to (5.128) a stopping time argument might be required. This time an appropriate stopping time τ_N would be the first time $t \leqslant T$ the process $|Y_{n+1}(t) - Y_n(t)|$ exceeds N. The time T should then be replaced with τ_N.) Consequently, from (5.135) we get

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} e^{2\alpha t}\left|Y_{n+1}(t)-Y_n(t)\right|^2\right] + \mathbb{E}\left[\int_0^T e^{2\alpha s} d\left\langle M_{n+1}-M_n, M_{n+1}-M_n\right\rangle(s)\right] \leqslant 131 e^{2\alpha T}\mathbb{E}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right]+\frac{131}{2}\left\|\begin{pmatrix} Y_n - Y_{n-1} \\ M_n - M_{n-1}\end{pmatrix}\right\|_\alpha^2. \tag{5.136}$$

Since by definition $Y_n(T) = \mathbb{E}\left[\xi \mid \mathcal{F}_T^T\right]$ for all $n \in \mathbb{N}$, this sequence also converges with respect to the norm $\|\cdot\|_{\mathcal{S}^2\times\mathcal{M}^2}$ defined by

$$\left\|\begin{pmatrix} Y \\ M\end{pmatrix}\right\|_{\mathcal{S}^2\times\mathcal{M}^2}^2 = \mathbb{E}\left[\sup_{0<s<T}\left|Y(s)\right|^2\right] + \mathbb{E}\left[\left\langle M, M\right\rangle(T) - \left\langle M, M\right\rangle(0)\right],$$

because

$$Y_{n+1}(0) = M_{n+1}(0) = \mathbb{E}\left[\xi + \int_0^T f_n\left(s, Y_n(s), Z_{M_n}(s)\right) ds \mid \mathcal{F}_0^0\right], \quad n \in \mathbb{N}.$$

This concludes the proof of Theorem 5.4. □

In the following Theorem 5.5 we replace the Lipschitz condition (5.113) in Theorem 5.4 for the function $Y(s) \mapsto f\left(s, Y(s), Z_M(s)\right)$ with the (weaker) monotonicity condition (5.137). Here we write y for the variable $Y(s)$ and z for $Z_M(s)$. As in Theorem 5.4 for every $s \in [0,T]$ the function $f(s) = f(s,\cdot,\cdot)$ is defined on $\mathbb{R}^k \times \mathcal{M}_{\mathrm{AC}}^{2,s}$ and by hypothesis it is continuous.

Theorem 5.5. *Let the function $f(s) : \mathbb{R}^k \times \mathcal{M}_{\mathrm{AC}}^{2,s} \to \mathbb{R}^k$ be monotone in the variable y and Lipschitz in z. More precisely, suppose that there exist finite constants C_1 and C_2 such that for any two pairs of processes (Y, M) and $(U, N) \in \mathcal{S}^2\left([0,T], \mathbb{R}^k\right) \times \mathcal{M}^2\left([0,T], \mathbb{R}^k\right)$ the following inequalities hold for all $0 \leqslant s \leqslant T$:*

$$\left\langle Y(s) - U(s), f\left(s, Y(s), Z_M(s)\right) - f\left(s, U(s), Z_M(s)\right)\right\rangle \leqslant C_1 \left|Y(s) - U(s)\right|^2, \tag{5.137}$$

$$|f(s,Y(s),Z_M(s)) - f(s,Y(s),Z_N(s))| \leqslant C_2\left(\frac{d}{ds}\langle M-N, M-N\rangle(s)\right)^{1/2}, \tag{5.138}$$

and

$$|f(s,Y(s),0)| \leqslant \overline{f}(s) + K|Y(s)|. \tag{5.139}$$

If $\mathbb{E}\left[\int_0^T \left|\overline{f}(s)\right|^2 ds\right] < \infty$, *then there exists a unique pair*

$$(Y,M) \in \mathcal{S}^2\left([0,T],\mathbb{R}^k\right) \times \mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$$

such that

$$Y(t) = \xi + \int_t^T f(s,Y(s),Z_M(s))\,ds + M(t) - M(T), \tag{5.140}$$

where $Y(T) = \xi \in L^2\left(\Omega,\mathcal{F}_T,\mathbb{R}^k\right)$ *is given and where* $Y(0) = M(0)$.

In order to prove Theorem 5.5 we need the following proposition, the proof of which uses the monotonicity condition (5.137) in an explicit manner.

Proposition 5.6. *Suppose that for every* $\xi \in L^2\left(\Omega,\mathcal{F}_T^0,\mathbb{P}\right)$ *and* $M \in \mathcal{M}^2$ *there exists a pair* $(Y,N) \in \mathcal{S}^2 \times \mathcal{M}^2$ *such that*

$$Y(t) = \xi + \int_t^T f(s,Y(s),Z_M(s))\,ds + N(t) - N(T). \tag{5.141}$$

Then for every $\xi \in L^2\left(\Omega,\mathcal{F}_T^0,\mathbb{P}\right)$ *there exists a unique pair* $(Y,M) \in \mathcal{S}^2 \times \mathcal{M}^2$ *which satisfies (5.140).*

The following proposition can be viewed as a consequence of Theorem 12.4 in [Hairer and Wanner (1991)]. The result is due to Burrage and Butcher [Burrage and Butcher (1979)] and Crouzeix [Crouzeix (1979)]. The constants obtained by these authors are somewhat different from ours.

Proposition 5.7. *Fix a martingale* $M \in \mathcal{M}^2$, *and choose* $\delta > 0$ *in such a way that* $\delta C_1 < 1$. *Here* C_1 *is the constant which occurs in inequality (5.137). Choose, for given* $y \in \mathbb{R}^k$, *the random variable* $\widetilde{Y}(t) \in \mathbb{R}^k$ *in such a way that* $y = \widetilde{Y}(t) - \delta f\left(t,\widetilde{Y}(t),Z_M(t)\right)$. *Then the mapping* $y \mapsto f\left(t,\widetilde{Y}(t),Z_M(t)\right)$ *is Lipschitz continuous with a Lipschitz constant which is equal to* $\frac{1}{\delta}\max\left(1,\frac{\delta C_1}{1-\delta C_1}\right)$. *Moreover, the mapping* $y \mapsto I - \delta f(t,y,Z_M(t))$ *is surjective and has a Lipschitz continuous inverse with Lipschitz constant* $\frac{1}{1-\delta C_1}$.

Proof. [Proof of Proposition 5.7.] Let the pair $(y_1, y_2) \in \mathbb{R}^k \times \mathbb{R}^k$ and the pair of $\mathbb{R}^k \times \mathbb{R}^k$-valued random variables $\left(\widetilde{Y}_1(t), \widetilde{Y}_2(t)\right)$ be such that the following equalities are satisfied:

$$y_1 = \widetilde{Y}_1(t) - \delta f\left(t, \widetilde{Y}_1(t), Z_M(t)\right) \quad \text{and} \quad y_2 = \widetilde{Y}_2(t) - \delta f\left(t, \widetilde{Y}_2(t), Z_M(t)\right). \tag{5.142}$$

We have to show that there exists a constant $C(\delta)$ such that

$$\left|f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right| \leqslant C(\delta)\left|y_2 - y_1\right|. \tag{5.143}$$

In order to achieve this we will exploit the inequality:

$$\begin{aligned}&\left\langle \widetilde{Y}_2(t) - \widetilde{Y}_1(t), f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right\rangle \\ &\leqslant C_1\left|\widetilde{Y}_2(t) - \widetilde{Y}_1(t)\right|^2.\end{aligned} \tag{5.144}$$

Inserting the equalities in (5.142) into (5.144) results in

$$\begin{aligned}&\left\langle y_2 - y_1, f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right\rangle \\ &\quad + \delta\left|f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right|^2 \\ &\leqslant C_1\left|y_2 - y_1\right|^2 + 2\delta C_1\left\langle y_2 - y_1, f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right\rangle \\ &\quad + C_1\delta^2\left|f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right|^2.\end{aligned} \tag{5.145}$$

Notice that (5.145) is equivalent to:

$$\begin{aligned}&\delta\left|f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right|^2 \\ &\leqslant C_1\left|y_2 - y_1\right|^2 \\ &\quad + 2\left(\delta C_1 - \frac{1}{2}\right)\left\langle y_2 - y_1, f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right\rangle \\ &\quad + C_1\delta^2\left|f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right|^2.\end{aligned} \tag{5.146}$$

Put $\alpha = \dfrac{1 - |1 - 2\delta C_1|}{2\delta C_1}$. Notice that, since $1 - \delta C_1 > 0$, the constant α is positive as well, $\alpha = 1$ provided $2\delta C_1 < 1$. Since $\delta C_1 < 1$ and

$$\begin{aligned}&2\left|\left\langle y_2 - y_1, f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right\rangle\right| \\ &\leqslant \frac{1}{\alpha\delta}\left|y_2 - y_1\right|^2 + \alpha\delta\left|f\left(t, \widetilde{Y}_2(t), Z_M(t)\right) - f\left(t, \widetilde{Y}_1(t), Z_M(t)\right)\right|^2,\end{aligned} \tag{5.147}$$

the inequality in (5.146) implies

$$\delta\left|f\left(t,\widetilde{Y}_2(t),Z_M(t)\right)-f\left(t,\widetilde{Y}_1(t),Z_M(t)\right)\right|\leqslant\max\left(1,\frac{\delta C_1}{1-\delta C_1}\right)|y_2-y_1|. \tag{5.148}$$

The Lipschitz constant is given by $C(\delta)=\frac{1}{\delta}\max\left(1,\frac{\delta C_1}{1-\delta C_1}\right)$: compare (5.148) and (5.143). The surjectivity of the mapping $y\mapsto y-\delta f\left(t,y,Z_M(t)\right)$ is a consequence of Theorem 1 in Croezeix et al [Crouzeix *et al.* (1983)]. Denote the mapping $y\mapsto t\left(t,y,Z_M(t)\right)$ by $f_{t,M}$. Then for $0<2\delta C_1<1$ the mapping $I-\delta f_{t,M}$ is invertible. Since

$$\left(I-\delta f_{t,M}\right)^{-1}=I+\delta f\left(t,\left(I-\delta f_{t,M}\right)^{-1},Z_M(t)\right),$$

and since by (5.148) the mapping $y\mapsto f\left(t,\left(I-\delta f_{t,M}\right)^{-1}y,Z_M(t)\right)$ is Lipschitz continuous with Lipschitz constant $\frac{1}{\delta}\max\left(1,\frac{\delta C_1}{1-\delta C_1}\right)$ we see that the mapping $y\mapsto\left(I-\delta f_{t,M}\right)^{-1}y$ is Lipschitz continuous with constant $\max\left(2,\frac{1}{1-\delta C_1}\right)$. A somewhat better constant is obtained by again using (5.144), and replacing

$$f\left(t,\widetilde{Y}_2(t),Z_M(t)\right)-f\left(t,\widetilde{Y}_1(t),Z_M(t)\right)$$

with $\delta^{-1}\left(\widetilde{y}_2-\widetilde{y}_1-y_2+y_1\right)$. Then we see:

$$\left|\widetilde{y}_2-\widetilde{y}_1\right|^2-\left\langle\widetilde{y}_2-\widetilde{y}_1,y_2-y_1\right\rangle\leqslant\delta C_1\left|\widetilde{y}_2-\widetilde{y}_1\right|^2, \tag{5.149}$$

and hence

$$\left(1-\delta C_1\right)\left|\widetilde{y}_2-\widetilde{y}_1\right|^2\leqslant\left\langle\widetilde{y}_2-\widetilde{y}_1,y_2-y_1\right\rangle\leqslant\left|\widetilde{y}_2-\widetilde{y}_1\right|\,\left|y_2-y_1\right|. \tag{5.150}$$

Altogether this proves Proposition 5.7. □

In Corollary 5.3 the process $t\mapsto M(t)$, $t\in[0,T]$, is a martingale in the space of $\mathbb{R}^k$-valued martingales in $L^2\left(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^k\right)$ which is denoted by $\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$: see Definition 5.9.

Corollary 5.3. *For $\delta>0$ such that $2\delta C_1<1$ there exist processes Y_δ and $\widetilde{Y}_\delta\in\mathcal{S}^2$ and a martingale $M_\delta\in\mathcal{M}^2$ such that the following equalities are satisfied:*

$$\begin{aligned}Y_\delta(t)&=\widetilde{Y}_\delta(t)-\delta f\left(t,\widetilde{Y}_\delta(t),Z_M(t)\right)\\&=Y_\delta(T)+\int_t^T f\left(s,\widetilde{Y}_\delta(s),Z_M(s)\right)ds+M_\delta(t)-M_\delta(T)\\&=Y_\delta(T)+\int_t^T f_\delta\left(s,Y_\delta(s),Z_M(s)\right)ds+M_\delta(t)-M_\delta(T).\end{aligned} \tag{5.151}$$

Proof. From Theorem 1 (page 87) in [Crouzeix *et al.* (1983)] it follows that the mapping $y \mapsto y - \delta f(t, y, Z_M(t))$ is a surjective map from $\mathbb{R}^k$ onto itself, provided $0 < \delta C_1 < 1$. If y_2 and y_1 in $\mathbb{R}^k$ are such that $y_2 - \delta f(t, y_2, Z_M(t)) = y_1 - \delta f(t, y_1, Z_M(t))$. Then

$$|y_2 - y_1|^2 = \langle y_2 - y_1, \delta f(t, y_2, Z_M(t)) - \delta f(t, y_1, Z_M(t))\rangle \leqslant \delta C_1 |y_2 - y_1|^2,$$

and hence $y_2 = y_1$. It follows that the continuous mapping $y \mapsto y - \delta f(t, y, Z_M(t))$ has a continuous inverse. Denote this inverse by $(I - \delta f_{t,M})^{-1}$. Moreover, for $0 < 2\delta C_1 < 1$, the mapping $y \mapsto f\left(t, (I - \delta f_{t,M})^{-1}, Z_m(t)\right)$ is Lipschitz continuous with Lipschitz constant δ^{-1}, which follows from Proposition 5.7. The remaining assertions in Corollary 5.3 are consequences of Theorem 5.4 where the Lipschitz condition in (5.113) was used with δ^{-1} instead of C_1.

This establishes the proof of Corollary 5.3. □

Remark 5.17. For more information on the surjectivity of the mapping $y \mapsto y - \delta f(s, y, z)$ the reader is referred to Remark 5.19 in Subsection 5.4.1.

Proof. [Proof of Proposition 5.6.] The proof of the uniqueness part follows from Corollary 5.2.

Fix $\xi \in L^2\left(\Omega, \mathcal{F}_T^0, \mathbb{P}\right)$, and let the martingale $M_{n-1} \in \mathcal{M}^2$ be given. Then by hypothesis there exists a pair $(Y_n, M_n) \in \mathcal{S}^2 \times \mathcal{M}^2$ which satisfies:

$$Y_n(t) = \xi + \int_t^T f\left(s, Y_n(s), Z_{M_{n-1}}(s)\right) ds + M_n(t) - M_n(T). \tag{5.152}$$

Another use of this hypothesis yields the existence of a pair $(Y_{n+1}, M_{n+1}) \in \mathcal{S}^2 \times \mathcal{M}^2$ which again satisfies (5.152) with $n+1$ instead of n. We will prove that the sequence (Y_n, M_n) is a Cauchy sequence in the space $\mathcal{S}^2 \times \mathcal{M}^2$. Put $\gamma = 1 + 2C_1 + 2C_2^2$. We apply Itô's formula to obtain

$$\begin{aligned}
& e^{\gamma T} |Y_{n+1}(T) - Y_n(T)|^2 - e^{\gamma t} |Y_{n+1}(t) - Y_n(t)|^2 \\
& = \gamma \int_t^T e^{\gamma s} |Y_{n+1}(s) - Y_n(s)|^2 ds \\
& \quad + 2 \int_t^T e^{\gamma s} \langle Y_{n+1}(s) - Y_n(s), d(Y_{n+1}(s) - Y_n(s))\rangle \\
& \quad + \int_t^T e^{\gamma s} d\langle M_{n+1} - M_n, M_{n+1} - M_n\rangle(s) \\
& = \gamma \int_t^T e^{\gamma s} |Y_{n+1}(s) - Y_n(s)|^2 ds
\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\int_t^T e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), d\left(M_{n+1}(s) - M_n(s)\right)\right\rangle \\
&- 2\int_t^T e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), f(s, Y_{n+1}(s), Z_{M_n}(s)) - f(s, Y_n(s), Z_{M_n}(s))\right\rangle ds \\
&+ 2\int_t^T e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), f(s, Y_n(s), Z_{M_n}(s)) - f(s, Y_n(s), Z_{M_{n-1}}(s))\right\rangle ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s)
\end{aligned}
$$

(employ (5.137) and (5.138))

$$
\begin{aligned}
&\geqslant \gamma \int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad + 2\int_t^T e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), d\left(M_{n+1}(s) - M_n(s)\right)\right\rangle \\
&\quad - 2C_1 \int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad - 2C_2 \int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right| \left(\frac{d}{ds}\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle (s)\right)^{1/2} ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s)
\end{aligned}
$$

(employ the elementary inequality $2ab \leqslant 2a^2 + \frac{1}{2}b^2$)

$$
\begin{aligned}
&\geqslant \left(\gamma - 2C_1 - 2C_2^2\right) \int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\quad - \frac{1}{2}\int_t^T e^{\gamma s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle (s) \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s) \\
&\quad + 2\int_t^T e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), d\left(M_{n+1}(s) - M_n(s)\right)\right\rangle . \qquad (5.153)
\end{aligned}
$$

From (5.153) we infer the inequality

$$
\begin{aligned}
&\left(\gamma - 2C_1 - 2C_2^2\right) \int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds \\
&\ + \int_t^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s)
\end{aligned}
$$

$$+ e^{\gamma t} \left|Y_{n+1}(t) - Y_n(t)\right|^2 + 2\int_t^T e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), d(M_{n+1}(s) - M_n(s))\right\rangle$$
$$\leqslant e^{\gamma T} \left|Y_{n+1}(T) - Y_n(T)\right|^2 + \frac{1}{2}\int_t^T e^{\gamma s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle (s). \tag{5.154}$$

By taking expectations in (5.154) we get, since $\gamma = 1 + 2C_1 + 2C_2^2$,

$$\mathbb{E}\left[\int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds\right]$$
$$+ \mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s)\right] + e^{\gamma t}\mathbb{E}\left[\left|Y_{n+1}(t) - Y_n(t)\right|^2\right]$$
$$\leqslant e^{\gamma T}\mathbb{E}\left[\left|Y_{n+1}(T) - Y_n(T)\right|^2\right]$$
$$+ \frac{1}{2}\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle (s)\right]. \tag{5.155}$$

Iterating (5.155) yields:

$$\mathbb{E}\left[\int_t^T e^{\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 ds\right]$$
$$+ \mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s)\right] + e^{\gamma t}\mathbb{E}\left[\left|Y_{n+1}(t) - Y_n(t)\right|^2\right]$$
$$\leqslant \sum_{k=1}^{n} \frac{1}{2^{n-k}} e^{\gamma T}\mathbb{E}\left[\left|Y_{k+1}(T) - Y_k(T)\right|^2\right]$$
$$+ \frac{1}{2^n}\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M_1 - M_0, M_1 - M_0\right\rangle (s)\right]$$
$$= \frac{1}{2^n}\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M_1 - M_0, M_1 - M_0\right\rangle (s)\right] \tag{5.156}$$

where in the last line we used the equalities $Y_k(T) = \xi$, $k \in \mathbb{N}$. From the Burkholder-Davis-Gundy inequality with $p = \frac{1}{2}$ (see inequality (5.89)) together with (5.156) it follows that

$$\mathbb{E}\left[\max_{0\leqslant t\leqslant T}\int_0^t e^{\gamma s} \left\langle Y_{n+1}(s) - Y_n(s), d\left(M_{n+1} - M_n\right)(s)\right\rangle\right]$$
$$\leqslant 4\sqrt{2}\mathbb{E}\left[\left(\int_0^T e^{2\gamma s} \left|Y_{n+1}(s) - Y_n(s)\right|^2 d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle (s)\right)^{1/2}\right]$$

$$\leqslant 4\sqrt{2}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{\frac{1}{2}\gamma s}\left|Y_{n+1}(s)-Y_n(s)\right| \right.$$
$$\left.\times\left(\int_0^T e^{\gamma s} d\left\langle M_{n+1}-M_n, M_{n+1}-M_n\right\rangle(s)\right)^{1/2}\right]$$

(use the elementary inequality $4\sqrt{2}ab \leqslant \frac{1}{4}a^2 + 32b^2$)

$$\leqslant \frac{1}{4}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{\gamma s}\left|Y_{n+1}(s)-Y_n(s)\right|^2\right]$$
$$+ 32\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_{n+1}-M_n, M_{n+1}-M_n\right\rangle(s)\right]$$
$$\leqslant \frac{1}{4}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{\gamma s}\left|Y_{n+1}(s)-Y_n(s)\right|^2\right]$$
$$+ \frac{1}{2^{n-5}}\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_1-M_0, M_1-M_0\right\rangle(s)\right]. \tag{5.157}$$

(In the first step of (5.157) we employed inequality (5.106) once more.) From (5.154) and (5.157) we obtain

$$\sup_{0\leqslant t\leqslant T} e^{\gamma t}\left|Y_{n+1}(t)-Y_n(t)\right|^2$$
$$+ 2\int_0^T e^{\gamma s}\left\langle Y_{n+1}(s)-Y_n(s), d\left(M_{n+1}(s)-M_n(s)\right)\right\rangle$$
$$\leqslant e^{\gamma T}\left|Y_{n+1}(T)-Y_n(T)\right|^2 + \frac{1}{2}\int_0^T e^{\gamma s} d\left\langle M_n-M_{n-1}, M_n-M_{n-1}\right\rangle(s)$$
$$+ 2\sup_{0\leqslant t\leqslant T}\int_0^t e^{\gamma s}\left\langle Y_{n+1}(s)-Y_n(s), d\left(M_{n+1}(s)-M_n(s)\right)\right\rangle. \tag{5.158}$$

From (5.156) (for $n-1$ instead of n), (5.157), and the fact that $Y_{n+1}(T) = Y_n(T) = \xi$ from (5.156) we infer the inequalities:

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} e^{\gamma t}\left|Y_{n+1}(t)-Y_n(t)\right|^2\right]$$
$$\leqslant \frac{1}{2}\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_n-M_{n-1}, M_n-M_{n-1}\right\rangle(s)\right]$$
$$+ 2\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\int_0^t e^{\gamma s}\left\langle Y_{n+1}(s)-Y_n(s), d\left(M_{n+1}(s)-M_n(s)\right)\right\rangle\right]$$

$$
\begin{aligned}
&\leqslant \frac{1}{2}\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_n - M_{n-1}, M_n - M_{n-1}\right\rangle(s)\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{\gamma s}\left|Y_{n+1}(s) - Y_n(s)\right|^2\right] \\
&\quad + 64\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_{n+1} - M_n, M_{n+1} - M_n\right\rangle(s)\right] \\
&\leqslant \frac{1}{2}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} e^{\gamma s}\left|Y_{n+1}(s) - Y_n(s)\right|^2\right] \\
&\quad + \frac{65}{2^n}\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_1 - M_0, M_1 - M_0\right\rangle(s)\right]. \qquad (5.159)
\end{aligned}
$$

From (5.159) we infer the inequality

$$
\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} e^{\gamma t}\left|Y_{n+1}(t) - Y_n(t)\right|^2\right] \leqslant \frac{65}{2^n}\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M_1 - M_0, M_1 - M_0\right\rangle(s)\right]. \tag{5.160}
$$

(In order to justify the passage from (5.154) to (5.160) like in passing from inequality (5.125) to (5.128) a stopping time argument might be required.) From (5.156) and (5.160) it follows that the sequence (Y_n, M_n) converges in the space $\mathcal{S}^2 \times \mathcal{M}^2$, and that its limit (Y, M) satisfies (5.140) in Theorem 5.5. This completes the proof of Proposition 5.6. □

Proposition 5.8. *Let the notation and hypotheses be as in Theorem 5.5. Let for $\delta > 0$ with $2\delta C_1 < 1$ the processes Y_δ, $\widetilde{Y}_\delta \in \mathcal{S}^2$ and the martingale $M_\delta \in \mathcal{M}^2$ be such that the equalities of (5.151) in Corollary 5.3 are satisfied. Then the family*

$$
\left\{(Y_\delta, M_\delta) : 0 < \delta < \frac{1}{2C_1}\right\}
$$

converges in the space $\mathcal{S}^2 \times \mathcal{M}^2$ if δ decreases to 0, provided that the terminal value $\xi = Y_\delta(T)$ is given.

Let (Y, M) be the limit in the space $\mathcal{S}^2 \times \mathcal{M}^2$. In fact from the proof of Proposition 5.8 it follows that

$$
\left\| \begin{pmatrix} Y_\delta - Y \\ M_\delta - M \end{pmatrix} \right\|_{\mathcal{S}^2 \times \mathcal{M}^2} = \mathcal{O}(\delta) \tag{5.161}
$$

as $\delta \downarrow 0$, provided that $\left\|Y_{\delta_2}(T) - Y_{\delta_1}(T)\right\|_{L^2\left(\Omega, \mathcal{F}_T^0, \mathbb{P}\right)} = \mathcal{O}\left(\left|\delta_2 - \delta_1\right|\right)$.

Proof. [Proof of Proposition 5.8.] Let C_1 be the constant which occurs in inequality (5.137) in Theorem 5.5, and fix $0 < \delta_2 < \delta_1 < (2C_1)^{-1}$. Our estimates give quantitative bounds in case we restrict the parameters δ, δ_1 and δ_2 to the interval $\left(0, (4C_1+4)^{-1}\right)$. An appropriate choice for the constant γ in the present proof turns out to be $\gamma = 6 + 4C_1$ (see e.g. the inequalities (5.163), (5.175), (5.176), and (5.177) below). An appropriate choice for the positive number a, which may be a function of the parameters δ_1 and δ_2, in (5.174), (5.175) and subsequent inequalities below is given by $a = (\delta_1+\delta_2)^{-1}$. For convenience we introduce the following notation: $\triangle Y(s) = Y_{\delta_2}(s) - Y_{\delta_1}(s)$, $\triangle M(s) = M_{\delta_2}(s) - M_{\delta_1}(s)$, $\triangle \widetilde{Y}(s) = \widetilde{Y}_{\delta_2}(s) - \widetilde{Y}_{\delta_1}(s)$, and $\triangle \widetilde{f}(s) = \widetilde{f}_{\delta_2}(s) - \widetilde{f}_{\delta_1}(s)$ where $\widetilde{f}_\delta(s) = f\left(s, \widetilde{Y}_\delta(s), Z_M(s)\right)$. From the equalities in (5.151) we infer

$$Y_\delta(t) = \widetilde{Y}_\delta(t) - \delta \widetilde{f}_\delta(t) = Y_\delta(T) + \int_t^T \widetilde{f}_\delta(s)ds + M_\delta(t) - M_\delta(T). \tag{5.162}$$

First we prove that the family $\left\{(Y_\delta, M_\delta) :\ 0 < \delta < (4C_1+4)^{-1}\right\}$ is bounded in the space $\mathcal{S}^2 \times \mathcal{M}^2$. Therefore we fix $\gamma > 0$ and apply Itô's formula to the process $t \mapsto e^{\gamma t}\left|Y_\delta(t)\right|^2$ to obtain:

$$\begin{aligned}
&e^{\gamma T}\left|Y_\delta(T)\right|^2 - e^{\gamma t}\left|Y_\delta(t)\right|^2 \\
&= \gamma \int_t^T e^{\gamma s}\left|Y_\delta(s)\right|^2 ds + 2\int_t^T e^{\gamma s}\left\langle Y_\delta(s), dY_\delta(s)\right\rangle + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) \\
&= \gamma \int_t^T e^{\gamma s}\left|\widetilde{Y}_\delta(s) - \delta \widetilde{f}_\delta(s)\right|^2 ds - 2\int_t^T e^{\gamma s}\left\langle \widetilde{Y}_\delta(s), \widetilde{f}_\delta(s)\right\rangle ds \\
&\quad - 2\int_t^T e^{\gamma s}\left\langle Y_\delta(s) - \widetilde{Y}_\delta(s), \widetilde{f}_\delta(s)\right\rangle ds + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) \\
&\quad + 2\int_t^T e^{\gamma s}\left\langle Y_\delta(s), dM_\delta(s)\right\rangle \\
&= \gamma \int_t^T e^{\gamma s}\left|\widetilde{Y}_\delta(s)\right|^2 ds + \gamma \int_t^T e^{\gamma s}\left|\delta \widetilde{f}_\delta(s)\right|^2 ds \\
&\quad - 2(1+\gamma\delta)\int_t^T e^{\gamma s}\left\langle \widetilde{Y}_\delta(s), \widetilde{f}_\delta(s) - f(s,0,Z_M(s))\right\rangle ds \\
&\quad + 2\int_t^T e^{\gamma s}\left\langle \delta \widetilde{f}_\delta(s), \widetilde{f}_\delta(s)\right\rangle ds - 2(1+\gamma\delta)\int_t^T e^{\gamma s}\left\langle \widetilde{Y}_\delta(s), f(s,0,Z_M(s))\right\rangle ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) + 2\int_t^T e^{\gamma s}\left\langle Y_\delta(s), dM_\delta(s)\right\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \gamma \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right|^2 ds + \left(\gamma\delta^2 + 2\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{f}_\delta(s)\right|^2 ds \\
&\quad - 2\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left\langle \widetilde{Y}_\delta(s), \widetilde{f}_\delta(s) - f\left(s,0,Z_M(s)\right)\right\rangle ds \\
&\quad - 2\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left\langle \widetilde{Y}_\delta(s), f\left(s,0,Z_M(s)\right) - f\left(s,0,0\right)\right\rangle ds \\
&\quad - 2\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left\langle \widetilde{Y}_\delta(s), f\left(s,0,0\right)\right\rangle ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) + 2\int_t^T e^{\gamma s} \left\langle Y_\delta(s), dM_\delta(s)\right\rangle
\end{aligned}
$$

(employ the inequalities (5.137), (5.138), and (5.139) of Theorem 5.5)

$$
\begin{aligned}
&\geqslant \gamma \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right|^2 ds + \left(\gamma\delta^2 + 2\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{f}_\delta(s)\right|^2 ds \\
&\quad - 2C_1\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right|^2 ds \\
&\quad - 2C_2\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right| \left(\frac{d}{ds}\left\langle M, M\right\rangle(s)\right)^{1/2} ds \\
&\quad - 2\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right| \left|f\left(s,0,0\right)\right| ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) + 2\int_t^T e^{\gamma s} \left\langle Y_\delta(s), dM_\delta(s)\right\rangle \\
&\geqslant \left(\gamma - 2\left(C_1+1\right)\left(1+\gamma\delta\right)\right) \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right|^2 ds + \left(\gamma\delta^2 + 2\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{f}_\delta(s)\right|^2 ds \\
&\quad - C_2^2\left(1+\gamma\delta\right) \int_t^T e^{\gamma s} d\left\langle M, M\right\rangle(s) - \left(1+\gamma\delta\right) \int_t^T e^{\gamma s} \left|f\left(s,0,0\right)\right|^2 ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) + 2\int_t^T e^{\gamma s} \left\langle Y_\delta(s), dM_\delta(s)\right\rangle .
\end{aligned}
\tag{5.163}
$$

From (5.163) we infer the inequality:

$$
\begin{aligned}
&\left(\gamma - 2\left(C_1+1\right)\left(1+\gamma\delta\right)\right) \int_t^T e^{\gamma s} \left|\widetilde{Y}_\delta(s)\right|^2 ds + \left(\gamma\delta^2 + 2\delta\right) \int_t^T e^{\gamma s} \left|\widetilde{f}_\delta(s)\right|^2 ds \\
&\quad + \int_t^T e^{\gamma s} d\left\langle M_\delta, M_\delta\right\rangle(s) + 2\int_t^T e^{\gamma s} \left\langle Y_\delta(s), dM_\delta(s)\right\rangle + e^{\gamma t}\left|Y_\delta(t)\right|^2 \\
&\leqslant e^{\gamma T}\left|Y_\delta(T)\right|^2 + \left(1+\gamma\delta\right)\left(C_2^2 \int_t^T e^{\gamma s} d\left\langle M, M\right\rangle(s) + \int_t^T e^{\gamma s}\left|f\left(s,0,0\right)\right|^2 ds\right).
\end{aligned}
\tag{5.164}
$$

From (5.164) we deduce

$$\begin{aligned}
&(\gamma - 2(C_1+1)(1+\gamma\delta))\,\mathbb{E}\left[\int_t^T e^{\gamma s}\left|\widetilde{Y}_\delta(s)\right|^2 ds\right] \\
&\quad + \left(\gamma\delta^2 + 2\delta\right)\mathbb{E}\left[\int_t^T e^{\gamma s}\left|\widetilde{f}_\delta(s)\right|^2 ds\right] + e^{\gamma t}\mathbb{E}\left[\left|Y_\delta(t)\right|^2\right] \\
&\leqslant e^{\gamma T}\mathbb{E}\left[\left|Y_\delta(T)\right|^2\right] \qquad (5.165)\\
&\quad + (1+\gamma\delta)\left(C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\langle M, M\rangle(s)\right] + \mathbb{E}\left[\int_t^T e^{\gamma s}\left|f(s,0,0)\right|^2 ds\right]\right).
\end{aligned}$$

In particular from (5.165) we see

$$\begin{aligned}
&\mathbb{E}\left[\int_t^T e^{\gamma s}\left|\widetilde{Y}_\delta(s)\right|^2 ds\right] \\
&\leqslant \frac{1}{\gamma - 2(C_1+1)(1+\gamma\delta)} e^{\gamma T}\mathbb{E}\left[\left|Y_\delta(T)\right|^2\right] \\
&\quad + \frac{1+\gamma\delta}{\gamma - 2(C_1+1)(1+\gamma\delta)} C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\langle M, M\rangle(s)\right] \\
&\quad + \frac{1+\gamma\delta}{\gamma - 2(C_1+1)(1+\gamma\delta)}\mathbb{E}\left[\int_t^T e^{\gamma s}\left|\overline{f}(s)\right|^2 ds\right]. \qquad (5.166)
\end{aligned}$$

In addition, from (5.164) we obtain the following inequalities

$$\begin{aligned}
&2\int_0^T e^{\gamma s}\langle Y_\delta(s), dM_\delta(s)\rangle + 2\sup_{0<t<T} e^{\gamma t}\left|Y_\delta(t)\right| \\
&\leqslant e^{\gamma T}\left|Y_\delta(T)\right|^2 + 2\sup_{0<t<T}\int_0^t e^{\gamma s}\langle Y_\delta(s), dM_\delta(s)\rangle \\
&\quad + (1+\gamma\delta)\left(C_2^2\int_t^T e^{\gamma s} d\langle M, M\rangle(s) + \int_t^T e^{\gamma s}\left|f(s,0,0)\right|^2 ds\right), \qquad (5.167)
\end{aligned}$$

and hence by using the Burkholder-Davis-Gundy inequality (5.89) for $p = \frac{1}{2}$ in combination with inequality (5.106) we get:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t}\left|Y_\delta(t)\right|\right] \\
&\leqslant e^{\gamma T}\mathbb{E}\left[\left|Y_\delta(T)\right|^2\right] + \mathbb{E}\left[\sup_{0<t<T}\int_0^t e^{\gamma s}\langle Y_\delta(s), dM_\delta(s)\rangle\right] \\
&\quad + (1+\gamma\delta)\left(C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\langle M, M\rangle(s)\right] + \mathbb{E}\left[\int_t^T e^{\gamma s}\left|f(s,0,0)\right|^2 ds\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&\leqslant e^{\gamma T}\mathbb{E}\left[\left|Y_\delta(T)\right|^2\right] + 8\sqrt{2}\mathbb{E}\left[\left(\int_0^T e^{2\gamma s}\left|Y_\delta(s)\right|^2 d\left\langle M_\delta, M_\delta\right\rangle(s)\right)^{1/2}\right] \\
&\quad + (1+\gamma\delta)\left(C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M, M\right\rangle(s)\right] + \mathbb{E}\left[\int_t^T e^{\gamma s}\left|f(s,0,0)\right|^2 ds\right]\right) \\
&\leqslant e^{\gamma T}\mathbb{E}\left[\left|Y_\delta(T)\right|^2\right] + \frac{1}{2}\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t}\left|Y_\delta(t)\right|\right] \\
&\quad + 64\mathbb{E}\left[\int_0^T e^{\gamma s}\left|Y_\delta(s)\right|^2 d\left\langle M_\delta, M_\delta\right\rangle(s)\right] \\
&\quad + (1+\gamma\delta)\left(C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M, M\right\rangle(s)\right] + \mathbb{E}\left[\int_t^T e^{\gamma s}\left|f(s,0,0)\right|^2 ds\right]\right).
\end{aligned}
\tag{5.168}
$$

(In the second step in (5.168) inequality (5.106) has been used again.) From (5.165) and (5.168) we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t}\left|Y_\delta(t)\right|\right] \\
&\leqslant 130 e^{\gamma T}\mathbb{E}\left[\left|Y_\delta(T)\right|^2\right] \\
&+ 130(1+\gamma\delta)\left(C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M, M\right\rangle(s)\right] + \mathbb{E}\left[\int_t^T e^{\gamma s}\left|f(s,0,0)\right|^2 ds\right]\right).
\end{aligned}
\tag{5.169}
$$

(In order to justify the passage from (5.167) to (5.169) like in passing from inequality (5.125) to (5.128) a stopping time argument might be required. (An appropriate stopping time τ_N would be the first time $t \leqslant T$ the process $|Y_\delta(t)|$ exceeds N. The time T should then be replaced with τ_N.) Next we notice that

$$
\left|\widetilde{f}_\delta(s)\right|^2 \leqslant 2\left|\overline{f}(s)\right|^2 + 2K^2\left|\widetilde{Y}_\delta(s)\right|^2 + 2C_2^2\frac{d}{ds}\left\langle M, M\right\rangle(s), \tag{5.170}
$$

and hence

$$
\begin{aligned}
&2\left\langle \delta_2\widetilde{f}_{\delta_2}(s) - \delta_1\widetilde{f}_{\delta_1}(s), \triangle\widetilde{f}(s)\right\rangle \geqslant -2\left|\delta_2-\delta_1\right|\left(\left|\widetilde{f}_{\delta_2}(s)\right|^2 - \left|\widetilde{f}_{\delta_1}(s)\right|^2\right) \\
&\geqslant -4\left|\delta_2-\delta_1\right|\left(\left|\overline{f}(s)\right|^2 + K^2\left|\widetilde{Y}_{\delta_2}(s)\right|^2 + K^2\left|\widetilde{Y}_{\delta_1}(s)\right|^2 + C_2^2\frac{d}{ds}\left\langle M, M\right\rangle(s)\right).
\end{aligned}
\tag{5.171}
$$

In a similar manner we also get

$$\left|\delta_2\widetilde{f}_{\delta_2}(s)-\delta_1\widetilde{f}_{\delta_1}(s)\right|^2 \leqslant 4\left(\delta_2^2+\delta_1^2\right)\left(\left|\overline{f}(s)\right|^2+K^2\left|\widetilde{Y}_{\delta_2}(s)\right|^2+K^2\left|\widetilde{Y}_{\delta_1}(s)\right|^2+C_2^2\frac{d}{ds}\left\langle M,M\right\rangle(s)\right). \tag{5.172}$$

Fix $\gamma>0$, and apply Itô's lemma to the process $t\mapsto e^{\gamma t}\left|\triangle Y(t)\right|^2$ to obtain

$$\begin{aligned}
&e^{\gamma T}\left|\triangle Y(T)\right|^2-e^{\gamma t}\left|\triangle Y(t)\right|^2\\
&=\gamma\int_t^T e^{\gamma s}\left|\triangle Y(s)\right|^2ds+2\int_t^T e^{\gamma s}\left\langle \triangle Y(s),d\triangle Y(s)\right\rangle\\
&\quad+\int_t^T e^{\gamma s}d\left\langle \triangle M,\triangle M\right\rangle(s)\\
&=\gamma\int_t^T e^{\gamma s}\left|\triangle\widetilde{Y}(s)-\delta_2\widetilde{f}_{\delta_2}(s)+\delta_1\widetilde{f}_{\delta_1}(s)\right|^2ds-2\int_t^T e^{\gamma s}\left\langle\triangle\widetilde{Y}(s),\triangle\widetilde{f}(s)\right\rangle ds\\
&\quad-2\int_t^T e^{\gamma s}\left\langle\triangle Y(s)-\triangle\widetilde{Y}(s),\triangle\widetilde{f}(s)\right\rangle ds+\int_t^T e^{\gamma s}d\left\langle \triangle M,\triangle M\right\rangle(s)\\
&\quad+2\int_t^T e^{\gamma s}\left\langle \triangle Y(s),d\triangle M(s)\right\rangle\\
&=\gamma\int_t^T e^{\gamma s}\left|\triangle\widetilde{Y}(s)\right|^2ds+\gamma\int_t^T e^{\gamma s}\left|\delta_2\widetilde{f}_{\delta_2}(s)-\delta_1\widetilde{f}_{\delta_1}(s)\right|^2ds\\
&\quad-2\int_t^T e^{\gamma s}\left\langle\triangle\widetilde{Y}(s),\triangle\widetilde{f}(s)\right\rangle ds\\
&\quad-2\gamma\int_t^T e^{\gamma s}\left\langle\delta_2\widetilde{f}_{\delta_2}(s)-\delta_1\widetilde{f}_{\delta_1}(s),\triangle\widetilde{Y}(s)\right\rangle ds\\
&\quad+2\int_t^T e^{\gamma s}\left\langle\delta_2\widetilde{f}_{\delta_2}(s)-\delta_1\widetilde{f}_{\delta_1}(s),\triangle\widetilde{f}(s)\right\rangle ds\\
&\quad+\int_t^T e^{\gamma s}d\left\langle \triangle M,\triangle M\right\rangle(s)+2\int_t^T e^{\gamma s}\left\langle \triangle Y(s),d\triangle M(s)\right\rangle.
\end{aligned} \tag{5.173}$$

Employing the inequalities (5.137), (5.171), (5.172) and an elementary one like

$$2\left|\left\langle y_1,y_2\right\rangle\right|\leqslant(a+1)\left|y_1\right|^2+(a+1)^{-1}\left|y_2\right|^2,\quad y_1,\ y_2\in\mathbb{R}^k,\quad a>0, \tag{5.174}$$

together with (5.173) we obtain

$$\begin{aligned}
&e^{\gamma T}\left|\triangle Y(T)\right|^2-e^{\gamma t}\left|\triangle Y(t)\right|^2\\
&\geqslant\left(\gamma-2C_1-\frac{\gamma}{a+1}\right)\int_t^T e^{\gamma s}\left|\triangle\widetilde{Y}(s)\right|^2ds
\end{aligned}$$

$$
\begin{aligned}
& - a\gamma \int_t^T e^{\gamma s} \left| \delta_2 \widetilde{f}_{\delta_2}(s) - \delta_1 \widetilde{f}_{\delta_1}(s) \right|^2 ds \\
& - 8\gamma \left| \delta_2 - \delta_1 \right| \left(\int_t^T e^{\gamma s} \left| \overline{f}(s) \right|^2 ds + C_2^2 \int_t^T e^{\gamma s} d \left\langle M, M \right\rangle (s) \right) \\
& - 8\gamma K^2 \left| \delta_2 - \delta_1 \right| \left(\int_t^T e^{\gamma s} \left(\left| \widetilde{Y}_{\delta_1}(s) \right|^2 + \left| \widetilde{Y}_{\delta_2}(s) \right|^2 \right) ds \right) \\
& + \int_t^T e^{\gamma s} d \left\langle \triangle M, \triangle M \right\rangle (s) + 2 \int_t^T e^{\gamma s} \left\langle \triangle Y(s), d\triangle M(s) \right\rangle \\
\geqslant & \left(\gamma - 2C_1 - \frac{\gamma}{a+1} \right) \int_t^T e^{\gamma s} \left| \triangle \widetilde{Y}(s) \right|^2 ds \\
& - 4\gamma \left(2 \left| \delta_2 - \delta_1 \right| + a \left(\delta_1^2 + \delta_2^2 \right) \right) \\
& \quad \times \left(\int_t^T e^{\gamma s} \left| \overline{f}(s) \right|^2 ds + C_2^2 \int_t^T e^{\gamma s} d \left\langle M, M \right\rangle (s) \right) \\
& - 4\gamma K^2 \left(2 \left| \delta_2 - \delta_1 \right| + a \left(\delta_1^2 + \delta_2^2 \right) \right) \left(\int_t^T e^{\gamma s} \left(\left| \widetilde{Y}_{\delta_1}(s) \right|^2 + \left| \widetilde{Y}_{\delta_2}(s) \right|^2 \right) ds \right) \\
& + \int_t^T e^{\gamma s} d \left\langle \triangle M, \triangle M \right\rangle (s) + 2 \int_t^T e^{\gamma s} \left\langle \triangle Y(s), d\triangle M(s) \right\rangle .
\end{aligned} \tag{5.175}
$$

From (5.175) we obtain

$$
\begin{aligned}
& \left(\frac{\gamma a}{a+1} - 2C_1 \right) \int_t^T e^{\gamma s} \left| \triangle \widetilde{Y}(s) \right|^2 ds + e^{\gamma t} \left| \triangle Y(t) \right|^2 \\
& \quad + \int_t^T e^{\gamma s} d \left\langle \triangle M, \triangle M \right\rangle (s) + 2 \int_t^T e^{\gamma s} \left\langle \triangle Y(s), d\triangle M(s) \right\rangle \\
& \leqslant e^{\gamma T} \left| \triangle Y(T) \right|^2 \\
& \quad + 4\gamma \left(2 \left| \delta_2 - \delta_1 \right| + a \left(\delta_1^2 + \delta_2^2 \right) \right) \\
& \quad \quad \times \left(\int_t^T e^{\gamma s} \left| \overline{f}(s) \right|^2 ds + C_2^2 \int_t^T e^{\gamma s} d \left\langle M, M \right\rangle (s) \right) \\
& \quad + 4\gamma K^2 \left(2 \left| \delta_2 - \delta_1 \right| + a \left(\delta_1^2 + \delta_2^2 \right) \right) \left(\int_t^T e^{\gamma s} \left(\left| \widetilde{Y}_{\delta_1}(s) \right|^2 + \left| \widetilde{Y}_{\delta_2}(s) \right|^2 \right) ds \right) .
\end{aligned} \tag{5.176}
$$

From (5.166) and (5.176) we infer

$$
\left(\frac{\gamma a}{a+1} - 2C_1 \right) \mathbb{E} \left[\int_t^T e^{\gamma s} \left| \triangle \widetilde{Y}(s) \right|^2 ds \right] + e^{\gamma t} \mathbb{E} \left[\left| \triangle Y(t) \right|^2 \right]
$$

$$
\begin{aligned}
&+ \mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle \triangle M, \triangle M\right\rangle(s)\right] \\
&\leqslant e^{\gamma T}\mathbb{E}\left[\left|\triangle Y(T)\right|^2\right] + \gamma_1\left(\delta_1,\delta_2\right)e^{\gamma T}\mathbb{E}\left[\left|Y_{\delta_1}(T)\right|^2\right] \\
&+ \gamma_1\left(\delta_2,\delta_1\right)e^{\gamma T}\mathbb{E}\left[\left|Y_{\delta_2}(T)\right|^2\right] \qquad (5.177)\\
&+ \gamma_2\left(\delta_1,\delta_2\right)\left(\mathbb{E}\left[\int_t^T e^{\gamma s}\left|\overline{f}(s)\right|^2 ds\right] + C_2^2\mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M, M\right\rangle(s)\right]\right)
\end{aligned}
$$

where

$$
\begin{aligned}
\gamma_1\left(\delta_1,\delta_2\right) &= 4\gamma K^2\left(2\left|\delta_2-\delta_1\right| + a\left(\delta_1^2+\delta_2^2\right)\right)\frac{1}{\gamma - 2\left(C_1+1\right)\left(1+\gamma\delta_1\right)}; \\
\gamma_2\left(\delta_1,\delta_2\right) &= 4\gamma\left(2\left|\delta_2-\delta_1\right| + a\left(\delta_1^2+\delta_2^2\right)\right) \qquad (5.178)\\
&\qquad \times\left(1 + \frac{K^2\left(1+\gamma\delta_1\right)}{\gamma - 2\left(C_1+1\right)\left(1+\gamma\delta_1\right)} + \frac{K^2\left(1+\gamma\delta_2\right)}{\gamma - 2\left(C_1+1\right)\left(1+\gamma\delta_2\right)}\right).
\end{aligned}
$$

From (5.176) we also get:

$$
\begin{aligned}
&\sup_{0<t<T}\left(e^{\gamma t}\left|\triangle Y(t)\right|^2\right) + 2\int_0^T e^{\gamma s}\left\langle \triangle Y(s), d\triangle M(s)\right\rangle \\
&\leqslant e^{\gamma T}\left|\triangle Y(T)\right|^2 \\
&\quad + 4\gamma\left(2\left|\delta_2-\delta_1\right| + a\left(\delta_1^2+\delta_2^2\right)\right) \\
&\qquad \times\left(\int_0^T e^{\gamma s}\left|\overline{f}(s)\right|^2 ds + C_2^2\int_0^T e^{\gamma s} d\left\langle M, M\right\rangle(s)\right) \\
&\quad + 4\gamma K^2\left(2\left|\delta_2-\delta_1\right| + a\left(\delta_1^2+\delta_2^2\right)\right)\left(\int_0^T e^{\gamma s}\left(\left|\widetilde{Y}_{\delta_1}(s)\right|^2 + \left|\widetilde{Y}_{\delta_2}(s)\right|^2\right) ds\right) \\
&\quad + 2\sup_{0<t<T}\int_0^t e^{\gamma s}\left\langle \triangle Y(s), d\triangle M(s)\right\rangle. \qquad (5.179)
\end{aligned}
$$

In what follows a stopping time argument might be required. This time an appropriate stopping time τ_N would be the first time $t \leqslant T$ the process $\left|\triangle Y_n\right| = \left|Y_{n+1}(t) - Y_n(t)\right|$ exceeds N. The time T should then be replaced with τ_N. From (5.179), (5.166), the inequality of Burkholder-Davis-Gundy (5.89) for $p = \frac{1}{2}$ and (5.177) with $t = 0$ we obtain:

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T}\left(e^{\gamma t}\left|\triangle Y(t)\right|^2\right)\right] \\
&\leqslant e^{\gamma T}\mathbb{E}\left[\left|\triangle Y(T)\right|^2\right] \\
&\quad + 4\gamma\left(2\left|\delta_2-\delta_1\right| + a\left(\delta_1^2+\delta_2^2\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&\times \left(\mathbb{E}\left[\int_0^T e^{\gamma s} \left| \overline{f}(s) \right|^2 ds \right] + C_2^2 \mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M, M \right\rangle (s) \right] \right) \\
&+ 4\gamma K^2 \left(2\left| \delta_2 - \delta_1 \right| + a\left(\delta_1^2 + \delta_2^2 \right) \right) \\
&\times \left(\mathbb{E}\left[\int_0^T e^{\gamma s} \left(\left| \widetilde{Y}_{\delta_1}(s) \right|^2 + \left| \widetilde{Y}_{\delta_2}(s) \right|^2 \right) ds \right] \right) \\
&+ 2\mathbb{E}\left[\sup_{0<t<T} \int_0^t e^{\gamma s} \left\langle \triangle Y(s), d\triangle M(s) \right\rangle \right] \\
\leqslant\ & e^{\gamma T} \mathbb{E}\left[\left| \triangle Y(T) \right|^2 \right] + \gamma_1\left(\delta_1, \delta_2 \right) e^{\gamma T} \mathbb{E}\left[\left| Y_{\delta_1}(T) \right|^2 \right] \\
&+ \gamma_1\left(\delta_2, \delta_1 \right) e^{\gamma T} \mathbb{E}\left[\left| Y_{\delta_2}(T) \right|^2 \right] \\
&+ \gamma_2\left(\delta_1, \delta_2 \right) \left(C_2^2 \mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M, M \right\rangle (s) \right] + \mathbb{E}\left[\int_t^T e^{\gamma s} \left| \overline{f}\left(s \right) ds \right|^2 \right] \right) \\
&+ 8\sqrt{2}\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t} \left| \triangle Y(t) \right| \left(\int_0^T e^{\gamma s} d\left\langle \triangle M, \triangle M \right\rangle (s) \right)^{1/2} \right] \\
\leqslant\ & e^{\gamma T} \mathbb{E}\left[\left| \triangle Y(T) \right|^2 \right] + \gamma_1\left(\delta_1, \delta_2 \right) e^{\gamma T} \mathbb{E}\left[\left| Y_{\delta_1}(T) \right|^2 \right] \\
&+ \gamma_1\left(\delta_2, \delta_1 \right) e^{\gamma T} \mathbb{E}\left[\left| Y_{\delta_2}(T) \right|^2 \right] \\
&+ \gamma_2\left(\delta_1, \delta_2 \right) \left(C_2^2 \mathbb{E}\left[\int_t^T e^{\gamma s} d\left\langle M, M \right\rangle (s) \right] + \mathbb{E}\left[\int_t^T e^{\gamma s} \left| \overline{f}\left(s \right) ds \right|^2 \right] \right) \\
&+ \frac{1}{2}\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t} \left| \triangle Y(t) \right|^2 \right] + 64\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle \triangle M, \triangle M \right\rangle (s) \right]. \qquad (5.180)
\end{aligned}
$$

Consequently, from (5.177) and (5.180) we deduce, like in the proof of inequality (5.169),

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t} \left| \triangle Y(t) \right|^2 \right] \\
&\leqslant 130\, e^{\gamma T} \mathbb{E}\left[\left| \triangle Y(T) \right|^2 \right] \\
&\quad + 130 \left(\gamma_1\left(\delta_1, \delta_2 \right) e^{\gamma T} \mathbb{E}\left[\left| Y_{\delta_1}(T) \right|^2 \right] + \gamma_1\left(\delta_2, \delta_1 \right) e^{\gamma T} \mathbb{E}\left[\left| Y_{\delta_2}(T) \right|^2 \right] \right) \\
&\quad + 130\gamma_2\left(\delta_1, \delta_2 \right) \left(C_2^2 \mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M, M \right\rangle (s) \right] + \mathbb{E}\left[\int_0^T e^{\gamma s} \left| \overline{f}\left(s \right) ds \right|^2 \right] \right). \qquad (5.181)
\end{aligned}
$$

(Again it is noticed that the passage from (5.179) to (5.181) is justified by a stopping time argument. The same argument was used several times. The first time we used it in passing from inequality (5.125) to (5.128).) Another appeal to (5.177) and (5.181) shows:

$$\begin{aligned}
&\left(\frac{\gamma a}{a+1}-2C_1\right)\mathbb{E}\left[\int_t^T e^{\gamma s}\left|\triangle\widetilde{Y}(s)\right|^2 ds\right] \\
&\quad+\mathbb{E}\left[\sup_{0<t<T} e^{\gamma t}\left|\triangle Y(t)\right|^2\right]+\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle\triangle M,\triangle M\right\rangle(s)\right] \\
&\leqslant 131\, e^{\gamma T}\mathbb{E}\left[\left|\triangle Y(T)\right|^2\right] \\
&\quad+131\left(\gamma_1\left(\delta_1,\delta_2\right)e^{\gamma T}\mathbb{E}\left[\left|Y_{\delta_1}(T)\right|^2\right]+\gamma_1\left(\delta_2,\delta_1\right)e^{\gamma T}\mathbb{E}\left[\left|Y_{\delta_2}(T)\right|^2\right]\right) \\
&\quad+131\gamma_2\left(\delta_1,\delta_2\right)\left(C_2^2\mathbb{E}\left[\int_0^T e^{\gamma s} d\left\langle M,M\right\rangle(s)\right]+\mathbb{E}\left[\int_0^T e^{\gamma s}\left|\overline{f}\left(s\right)ds\right|^2\right]\right).
\end{aligned} \tag{5.182}$$

The result in Proposition 5.8 now follows from (5.182) and the continuity of the functions $y\mapsto f\left(s,y,Z_M(s)\right)$, $y\in\mathbb{R}^k$. The fact that the convergence of the family (Y_δ,M_δ), $0<\delta\leqslant\left(4C_1+4\right)^{-1}$ is of order δ, as $\delta\downarrow 0$, follows by the choice of our parameters: $\gamma=4C_1+4$ and $a=\left(\delta_1+\delta_2\right)^{-1}$. □

Proof. [Proof of Theorem 5.5.] The proof of the uniqueness part follows from Corollary 5.2. The existence is a consequence of Theorem 5.4, Proposition 5.8 and Corollary 5.3. □

The following result shows that in the monotonicity condition we may always assume that the constant C_1 can be chosen as we like provided we replace the equation in (5.115) by (5.183) and adapt its solution.

Theorem 5.6. *Let the pair (Y,M) belong to $\mathcal{S}^2\left([0,T],\mathbb{R}^k\right)\times\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$. Fix $\lambda\in\mathbb{R}$, and put*

$$\left(Y_\lambda(t),M_\lambda(t)\right)=\left(e^{\lambda t}Y(t),Y(0)+\int_0^t e^{\lambda s}dM(s)\right).$$

Then the pair (Y_λ,M_λ) belongs to $\mathcal{S}^2\times\mathcal{M}^2$. Moreover, the following assertions are equivalent:

(i) The pair $(Y,M)\in\mathcal{S}^2\times\mathcal{M}^2$ satisfies $Y(0)=M(0)$ and

$$Y(t)=Y(T)+\int_t^T f\left(s,Y(s),Z_M(s)\right)ds+M(t)-M(T).$$

(ii) The pair (Y_λ, M_λ) *satisfies* $Y_\lambda(0) = M_\lambda(0)$ *and*

$$Y_\lambda(t) = Y_\lambda(T) + \int_t^T e^{\lambda s} f\left(s, e^{-\lambda s} Y_\lambda(s), e^{-\lambda s} Z_{M_\lambda}(s)\right) ds$$
$$- \lambda \int_t^T Y_\lambda(s) ds + M_\lambda(t) - M_\lambda(T). \tag{5.183}$$

Remark 5.18. Put $f_\lambda(s,y,z) = e^{\lambda s} f\left(s, e^{-\lambda s} y, e^{-\lambda s} z\right) - \lambda y$. If the function $y \mapsto f(s,y,z)$ has monotonicity constant C_1, then the function $y \mapsto f_\lambda(s,y,z)$ has monotonicity constant $C_1 - \lambda$. It follows that by reformulating the problem one always may assume that the monotonicity constant is 0.

Proof. [Proof of Theorem 5.6.] First notice the equality $e^{-\lambda s} Z_{M_\lambda}(s) = Z_M(s)$: see Remark 5.5. The equivalence of (i) and (ii) follows by considering the equalities in (i) and (ii) in differential form. □

5.4 Backward stochastic differential equations and Markov processes

In this section the coefficient $f(s) = f(s, \cdot, \cdot, \cdot)$, $s \in [0,T]$, of our BSDE is a mapping from $E \times \mathbb{R}^k \times \mathcal{M}_{\mathrm{AC}}^{2,s}$ to $\mathbb{R}^k$. For the definition of the space $\mathcal{M}_{\mathrm{AC}}^{2,s} = \mathcal{M}_{\mathrm{AC}}^{2,s}(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x})$ see Definition 5.7.

Theorem 5.7 below is the analogue of Theorem 5.5 with a Markov family of measures $\{\mathbb{P}_{\tau,x} : (\tau,x) \in [0,T] \times E\}$ instead of a single measure. Put

$$f_n(s) = f(s, X(s), Y_n(s), Z_{M_n}(s)),$$

and suppose that the processes $Y_n(s)$ and $Z_{M_n}(s)$ only depend of the state-time variable $(s, X(s))$. Put $Y(\tau,t) g(x) = \mathbb{E}_{\tau,x}[g(X(t))]$, $g \in C_b(E)$, and suppose that for every $g \in C_b(E)$ the function $(\tau,x,t) \mapsto Y(\tau,t) f(x)$ is continuous on the set $\{(\tau,x,t) \in [0,T] \times E \times [0,T] : 0 \leqslant \tau \leqslant T\}$. Then it can be proved that the Markov process

$$\{(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}), (X(t) : T \geqslant t \geqslant 0), (E, \mathcal{E})\} \tag{5.184}$$

has left limits and is right-continuous: see e.g. Theorem 2.9. Theorem 2.22 in [Gulisashvili and van Casteren (2006)] contains a similar result in case the state space E is locally compact and second countable. Suppose that the $\mathbb{P}_{\tau,x}$-martingale $t \mapsto N(t) - N(\tau)$, $t \in [\tau,T]$, belongs to the space $\mathcal{M}^2([\tau,T], \mathbb{P}_{\tau,x}, \mathbb{R}^k)$ (see Definition 5.5). It follows that the quantity $Z_M(s)(N)$ is measurable with respect to $\sigma\left(\mathcal{F}_{s+}^s, N(s+)\right)$: see equalities

(5.188), (5.189) and (5.190) below. The following iteration formulas play an important role:

$$Y_{n+1}(t) = \mathbb{E}_{t,X(t)}\left[\xi\right] + \int_t^T \mathbb{E}_{t,X(t)}\left[f_n(s)\right] ds,$$

$$M_{n+1}(t) = \mathbb{E}_{t,X(t)}\left[\xi\right] + \int_0^t f_n(s)ds + \int_t^T \mathbb{E}_{t,X(t)}\left[f_n(s)\right] ds.$$

Then the processes Y_{n+1} and M_{n+1} are related as follows:

$$Y_{n+1}(T) + \int_t^T f_n(s)ds + M_{n+1}(t) - M_{n+1}(T) = Y_{n+1}(t).$$

Moreover, by the Markov property, the process

$$\begin{aligned} t &\mapsto M_{n+1}(t) - M_{n+1}(\tau) \\ &= \mathbb{E}_{\tau,X(\tau)}\left[\xi \mid \mathcal{F}_t^\tau\right] - \mathbb{E}_{\tau,X(\tau)}\left[\xi\right] + \mathbb{E}_{\tau,X(\tau)}\left[\int_\tau^T f_n(s)ds \mid \mathcal{F}_t^\tau\right] \\ &\quad - \mathbb{E}_{\tau,X(\tau)}\left[\int_\tau^T f_n(s)ds\right] \\ &= \mathbb{E}_{\tau,X(\tau)}\left[\xi + \int_\tau^T f_n(s)ds \mid \mathcal{F}_t^\tau\right] - \mathbb{E}_{\tau,X(\tau)}\left[\xi + \int_\tau^T f_n(s)ds\right] \end{aligned}$$

is a $\mathbb{P}_{\tau,x}$-martingale on the interval $[\tau, T]$ for every $(\tau, x) \in [0, T] \times E$.

In Theorem 5.7 below we replace the Lipschitz condition (5.113) in Theorem 5.4 for the function $Y(s) \mapsto f(s, Y(s), Z_M(s))$ with the (weaker) monotonicity condition (5.193) for the function $Y(s) \mapsto f(s, X(s), Y(s), Z_M(s))$. Sometimes we write y for the variable $Y(s)$ and z for $Z_M(s)$. Notice that the functional $Z_{M_n}(t)$ only depends on $\mathcal{F}_{t+}^t := \bigcap_{h:T \geqslant t+h>t} \sigma(X(t+h))$ and that this σ-field belongs to the $\mathbb{P}_{t,x}$-completion of $\sigma(X(t))$ for every $x \in E$. This is the case, because by assumption the process $s \mapsto X(s)$ is right-continuous at $s = t$: see Proposition 5.3. In order to show this we have to prove equalities of the following type:

$$\mathbb{E}_{s,x}\left[Y \mid \mathcal{F}_{t+}^s\right] = \mathbb{E}_{t,X(t)}\left[Y\right], \quad \mathbb{P}_{s,x}\text{-almost surely}, \tag{5.185}$$

for all bounded random variables which are $\mathcal{F}_T^t$-measurable. By the monotone class theorem and density arguments the proof of (5.185) reduces to showing these equalities for $Y = \prod_{j=1}^n f_j(t_j, X(t_j))$, where $t = t_1 < t_2 < \cdots < t_n \leqslant T$, and the functions $x \mapsto f_j(t_j, x)$, $1 \leqslant j \leqslant n$, belong to the space $C_b(E)$. So we consider

$$\mathbb{E}_{s,x}\left[\prod_{j=1}^n f_j(t_j, X(t_j)) \mid \mathcal{F}_{t+}^s\right]$$

$$
\begin{aligned}
&= f_1\left(t, X(t)\right) \mathbb{E}_{t,X(t)}\left[\prod_{j=2}^{n} f_j\left(t_j, X\left(t_j\right)\right) \,\Big|\, \mathcal{F}_{t+}^s\right] \\
&= f_1\left(t, X(t)\right) \lim_{h\downarrow 0, 0<h<t_2-t} \mathbb{E}_{s,x}\left[\mathbb{E}_{s,x}\left[\prod_{j=2}^{n} f_j\left(t_j, X\left(t_j\right)\right) \,\Big|\, \mathcal{F}_{t+h}^s\right] \,\Big|\, \mathcal{F}_{t+}^s\right] \\
&= f_1\left(t, X(t)\right) \lim_{h\downarrow 0, 0<h<t_2-t} \mathbb{E}_{s,x}\left[\mathbb{E}_{t+h,X(t+h)}\left[\prod_{j=2}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right] \,\Big|\, \mathcal{F}_{t+}^s\right]
\end{aligned}
$$

(the function $\rho \mapsto \mathbb{E}_{\rho,X(\rho)}\left[\prod_{j=2}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right]$ is right-continuous)

$$
\begin{aligned}
&= f_1\left(t, X(t)\right) \mathbb{E}_{s,x}\left[\mathbb{E}_{t,X(t)}\left[\prod_{j=2}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right] \,\Big|\, \mathcal{F}_{t+}^s\right] \\
&= f_1\left(t, X(t)\right) \mathbb{E}_{t,X(t)}\left[\prod_{j=2}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right] \\
&= \mathbb{E}_{t,X(t)}\left[\prod_{j=1}^{n} f_j\left(t_j, X\left(t_j\right)\right)\right], \quad \mathbb{P}_{s,x}\text{-almost surely.}
\end{aligned}
\tag{5.186}
$$

Next suppose that the bounded random variable Y is measurable with respect to $\mathcal{F}_{t+}^t$. From (5.185) with $s = t$ it follows that $Y = \mathbb{E}_{t,X(t)}[Y]$, $\mathbb{P}_{t,x}$-almost surely. Hence such a variable Y only depends on the space-time variable $(t, X(t))$. Since $X(t) = x$ $\mathbb{P}_{t,x}$-almost surely it follows that the variable $\mathbb{E}_{t,x}\left[Y \,\middle|\, \mathcal{F}_{t+}^t\right]$ is $\mathbb{P}_{t,x}$-almost equal to the deterministic constant $\mathbb{E}_{t,x}[Y]$. A similar argument shows the following result. Let $0 \leqslant s < t \leqslant T$, and let Y be a bounded $\mathcal{F}_T^s$-measurable random variable. Then the following equality holds $\mathbb{P}_{s,x}$-almost surely:

$$
\mathbb{E}_{s,x}\left[Y \,\middle|\, \mathcal{F}_{t+}^s\right] = \mathbb{E}_{s,x}\left[Y \,\middle|\, \mathcal{F}_t^s\right]. \tag{5.187}
$$

In particular it follows that an $\mathcal{F}_{t+}^s$-measurable bounded random variable coincides with the $\mathcal{F}_t^s$-measurable variable $\mathbb{E}_{s,x}\left[Y \,\middle|\, \mathcal{F}_t^s\right]$ $\mathbb{P}_{s,x}$-almost surely for all $x \in E$. Hence (5.187) implies that the σ-field $\mathcal{F}_{t+}^s$ is contained in the $\mathbb{P}_{s,x}$-completion of the σ-field $\mathcal{F}_t^s$.

In addition, notice that the functional $Z_M(s)$ is defined by

$$
Z_M(s)(N) = \lim_{t\downarrow s} \frac{\langle M, N\rangle(t) - \langle M, N\rangle(s)}{t-s} \tag{5.188}
$$

where

$$
\langle M, N\rangle(t) - \langle M, N\rangle(s)
$$

$$= \lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left(M\left(t_{j+1,n}\right) - M\left(t_{j,n}\right)\right)\left(N\left(t_{j+1,n}\right) - N\left(t_{j,n}\right)\right). \tag{5.189}$$

For this the reader is referred to the remarks 5.5, 5.6, 5.9, and to formula (5.86). The symbol $t_{j,n}$ represents the real number $t_{j,n} = s + j2^{-n}(t-s)$. The limit in (5.189) exists $\mathbb{P}_{\tau,x}$-almost surely for all $\tau \in [0,s]$. As a consequence the process $Z_M(s)$ is $\mathcal{F}^\tau_{s+}$-measurable for all $\tau \in [0,s]$. It follows that the process $N \mapsto Z_M(s)(N)$ is $\mathbb{P}_{\tau,x}$-almost surely equal to the functional $N \mapsto \mathbb{E}_{\tau,x}\left[Z_M(s)(N) \mid \sigma\left(\mathcal{F}^\tau_s, N(s)\right)\right]$ provided that $Z_M(s)(N)$ is $\sigma\left(\mathcal{F}^\tau_{s+}, N(s+)\right)$-measurable. If the martingale M is of the form $M(s) = u\left(s, X(s)\right) + \int_0^s f(\rho)d\rho$, then the functional $Z_M(s)(N)$ is automatically $\sigma\left(\mathcal{F}^s_{s+}, N(s+)\right)$-measurable. It follows that, for every $\tau \in [0,s]$, the following equality holds $\mathbb{P}_{\tau,x}$-almost surely:

$$\mathbb{E}_{\tau,x}\left[Z_M(s)(N) \mid \sigma\left(\mathcal{F}^\tau_{s+}, N(s+)\right)\right] = \mathbb{E}_{\tau,x}\left[Z_M(s)(N) \mid \sigma\left(\mathcal{F}^\tau_{s}, N(s+)\right)\right]. \tag{5.190}$$

Moreover, in the next Theorem 5.7 the filtered probability measure

$$\left(\Omega, \mathcal{F}, \left(\mathcal{F}^0_t\right)_{t\in[0,T]}, \mathbb{P}\right)$$

is replaced with a Markov family of measures

$$\left(\Omega, \mathcal{F}^\tau_T, \left(\mathcal{F}^\tau_t\right)_{\tau\leqslant t\leqslant T}, \mathbb{P}_{\tau,x}\right), \quad (\tau,x) \in [0,T]\times E.$$

Its proof follows the lines of the proof of Theorem 5.5: it will not be repeated here. Relevant equalities which play a dominant role are the following ones: (5.128), (5.136), (5.169), and (5.182). In these inequalities the measure $\mathbb{P}_{\tau,x}$ replaces $\mathbb{P}$ and the coefficient $f\left(s, Y(s), Z_M(s)\right)$ is replaced with $f\left(s, X(s), Y(s), Z_M(s)\right)$. Then (5.191), which is the same as (5.128), is satisfied and with $\alpha = 1 + C_1^2 + C_2^2$ the following inequalities play a dominant role for the sequence (Y_n, M_n):

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\sup_{\tau<t<T} e^{2\alpha t}\left|Y_{n+1}(t)\right|^2\right] \\
&\leqslant 130 e^{2\alpha T}\mathbb{E}_{\tau,x}\left[\left|Y_{n+1}(T)\right|^2\right] + 130\mathbb{E}_{\tau,x}\left[\int_\tau^T e^{2\alpha s}\left|f(s,0,0)\right|^2 ds\right] \\
&\quad + 65\mathbb{E}_{\tau,x}\left[\int_\tau^T e^{2\alpha s} d\left\langle M_n, M_n\right\rangle(s)\right] + 65\mathbb{E}_{\tau,x}\left[\int_\tau^T e^{2\alpha s}\left|Y_n(s)\right|^2 ds\right] < \infty,
\end{aligned} \tag{5.191}$$

and

$$\mathbb{E}_{\tau,x}\left[\sup_{\tau\leqslant t\leqslant T} e^{2\alpha t}\left|Y_{n+1}(t) - Y_n(t)\right|^2\right]$$

$$+\mathbb{E}_{\tau,x}\left[\int_{\tau}^{T} e^{2\alpha s} d\left\langle M_{n+1}-M_n, M_{n+1}-M_n\right\rangle(s)\right]$$
$$\leqslant 131 e^{2\alpha T}\mathbb{E}_{\tau,x}\left[\left|Y_{n+1}(T)-Y_n(T)\right|^2\right]+\frac{131}{2}\left\|\begin{pmatrix} Y_n-Y_{n-1} \\ M_n-M_{n-1}\end{pmatrix}\right\|_{\tau,x,\alpha}^2. \tag{5.192}$$

Compare these inequalities with (5.128) and (5.192). The inequality in (5.192) plays only a direct role in case we are dealing with a Lipschitz continuous generator f. In case the generator f is only monotone (or one-sided Lipschitz) in the variable y, then we need the propositions 5.6, 5.7, 5.8, and Corollary 5.3.

The norm $\left\|\begin{pmatrix} Y \\ M\end{pmatrix}\right\|_{\tau,x,\alpha}$ is defined by:

$$\left\|\begin{pmatrix} Y \\ M\end{pmatrix}\right\|_{\tau,x,\alpha}^2=\mathbb{E}_{\tau,x}\left[\int_{\tau}^{T} e^{2\alpha s}\left|Y(s)\right|^2 ds+\int_{\tau}^{T} e^{2\alpha s} d\left\langle M, M\right\rangle(s)\right].$$

A proof of these inequalities can be found in [Van Casteren (2008)] and in the proof of Theorem 5.4 in the present Chapter 5. The following theorem contains the most important results of the present section 5.4.

Theorem 5.7. *Let for every $s\in[0,T]$ the function $f(s)=f(s,\cdot,\cdot,\cdot)$ be a function from $E\times\mathbb{R}^k\times\mathcal{M}_{\mathrm{AC}}^{2,s}$ to $\mathbb{R}^k$ which is monotone in the variable y and Lipschitz in z. More precisely, suppose that there exist finite constants C_1 and C_2 such that for any two pairs of processes (Y,M) and $(U,N)\in\mathcal{S}^2\left([0,T],\mathbb{R}^k\right)\times\mathcal{M}^2\left([0,T],\mathbb{R}^k\right)$ the following inequalities hold for all $0\leqslant s\leqslant T$:*

$$\left\langle Y(s)-U(s), f\left(s, X(s), Y(s), Z_M(s)\right)-f\left(s, X(s), U(s), Z_M(s)\right)\right\rangle$$
$$\leqslant C_1\left|Y(s)-U(s)\right|^2, \tag{5.193}$$
$$\left|f\left(s, X(s), Y(s), Z_M(s)\right)-f\left(s, X(s), Y(s), Z_N(s)\right)\right|$$
$$\leqslant C_2\left(\frac{d}{ds}\left\langle M-N, M-N\right\rangle(s)\right)^{1/2}, \tag{5.194}$$

and

$$\left|f\left(s, X(s), Y(s), 0\right)\right|\leqslant\overline{f}\left(s, X(s)\right)+K\left|Y(s)\right|. \tag{5.195}$$

Fix $(\tau,x)\in[0,T]\times E$ and let $Y(T)=\xi\in L^2\left(\Omega,\mathcal{F}_T^{\tau},\mathbb{P}_{\tau,x};\mathbb{R}^k\right)$ be given. In addition, suppose $\mathbb{E}_{\tau,x}\left[\int_{\tau}^{T}\left|\overline{f}\left(s,X(s)\right)\right|^2 ds\right]<\infty$. Then there exists a unique pair

$$(Y,M)\in\mathcal{S}^2\left([\tau,T],\mathbb{P}_{\tau,x},\mathbb{R}^k\right)\times\mathcal{M}^2\left([\tau,T],\mathbb{P}_{\tau,x},\mathbb{R}^k\right)$$

with $Y(\tau) = M(\tau)$ *such that*

$$Y(t) = \xi + \int_t^T f\left(s, X(s), Y(s), Z_M(s)\right) ds + M(t) - M(T). \tag{5.196}$$

Next let $\xi = \mathbb{E}_{T,X(T)}\left[\xi\right] \in \bigcap_{(\tau,x)\in[0,T]\times E} L^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$ *be given. Suppose that the functions* $(\tau,x) \mapsto \mathbb{E}_{\tau,x}\left[\left|\xi\right|^2\right]$ *and* $(\tau,x) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T \left|\overline{f}\left(s, X(s)\right)\right|^2 ds\right]$ *are locally bounded. Then there exists a unique pair*

$$(Y, M) \in \mathcal{S}^2_{\text{loc,unif}}\left([\tau,T], \mathbb{R}^k\right) \times \mathcal{M}^2_{\text{loc,unif}}\left([\tau,T], \mathbb{R}^k\right)$$

with $Y(0) = M(0)$ *such that equation (5.196) is satisfied.*

Again let $\xi = \mathbb{E}_{T,X(T)}\left[\xi\right] \in \bigcap_{(\tau,x)\in[0,T]\times E} L^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right)$ *be given. Suppose that the functions*

$$(\tau,x) \mapsto \mathbb{E}_{\tau,x}\left[\left|\xi\right|^2\right] \quad \textit{and} \quad (\tau,x) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T \left|\overline{f}\left(s, X(s)\right)\right|^2 ds\right]$$

are uniformly bounded. Then there exists a unique pair

$$(Y, M) \in \mathcal{S}^2_{\text{unif}}\left([\tau,T], \mathbb{R}^k\right) \times \mathcal{M}^2_{\text{unif}}\left([\tau,T], \mathbb{R}^k\right)$$

with $Y(0) = M(0)$ *such that equation (5.196) is satisfied.*

The notations

$$\begin{aligned} \mathcal{S}^2\left([\tau,T], \mathbb{P}_{\tau,x}, \mathbb{R}^k\right) &= \mathcal{S}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right) \quad \text{and} \\ \mathcal{M}^2\left([\tau,T], \mathbb{P}_{\tau,x}, \mathbb{R}^k\right) &= \mathcal{M}^2\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right) \end{aligned}$$

are explained in the definitions 5.4 and 5.5 respectively. The same is true for the notions

$$\begin{aligned} \mathcal{S}^2_{\text{loc,unif}}\left([0,T], \mathbb{R}^k\right) &= \mathcal{S}^2_{\text{loc,unif}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right), \\ \mathcal{M}^2_{\text{loc,unif}}\left([0,T], \mathbb{R}^k\right) &= \mathcal{M}^2_{\text{loc,unif}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right), \\ \mathcal{S}^2_{\text{unif}}\left([0,T], \mathbb{R}^k\right) &= \mathcal{S}^2_{\text{unif}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right), \quad \text{and} \\ \mathcal{M}^2_{\text{unif}}\left([0,T], \mathbb{R}^k\right) &= \mathcal{M}^2_{\text{unif}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right). \end{aligned}$$

In addition, the space $\mathcal{M}^{2,s}_{\text{AC}} = \mathcal{M}^{2,s}_{\text{AC}}\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}; \mathbb{R}^k\right)$ is explained in Definition 5.7 (see Lemma 5.1 as well). The probability measure $\mathbb{P}_{\tau,x}$ is defined on the σ-field $\mathcal{F}_T^\tau$. Since the existence properties of the solutions to backward stochastic equations are based on explicit inequalities, the proofs carry over to Markov families of measures. Ultimately these inequalities imply that boundedness and continuity properties of the function $(\tau,x) \mapsto \mathbb{E}_{\tau,x}\left[Y(t)\right]$,

$0 \leqslant \tau \leqslant t \leqslant T$, depend on the continuity of the function $x \mapsto \mathbb{E}_{T,x}[\xi]$, where ξ is a terminal value function which is supposed to be $\sigma(X(T))$-measurable. In addition, in order to be sure that the function $(\tau, x) \mapsto \mathbb{E}_{\tau,x}[Y(t)]$ is continuous, functions of the form $(\tau, x) \mapsto \mathbb{E}_{\tau,x}[f(t, u(t, X(t)), Z_M(t))]$ have to be continuous, whenever the following mappings

$$(\tau, x) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T |u(s, X(s))|^2 \, ds\right] \text{ and } (\tau, x) \mapsto \mathbb{E}_{\tau,x}[\langle M, M\rangle(T) - \langle M, M\rangle]$$

represent finite and continuous functions.

In the next example we see how the classical Feynman-Kac formula is related to backward stochastic differential equations.

Example 5.2. Suppose that the coefficient f has the special form:

$$f(t, x, r, z) = c(t, x)r + h(t, x)$$

and that the process $s \mapsto X^{x,t}(s)$ is a solution to a stochastic differential equation:

$$\begin{cases} X^{t,x}(s) - X^{t,x}(t) = \int_t^s b\left(\tau, X^{t,x}(\tau)\right) d\tau + \int_t^s \sigma\left(\tau, X^{t,x}(\tau)\right) dW(\tau), \\ \qquad\qquad t \leqslant s \leqslant T; \\ X^{t,x}(s) = x, \quad 0 \leqslant s \leqslant t. \end{cases}$$

In that case, the BSDE is linear,

$$\begin{aligned} Y^{t,x}(s) = g(X^{t,x}(T)) &+ \int_s^T [c(r, X^{t,x}(r)) Y^{t,x}(s) + h(r, X^{t,x}(r))] \, dr \\ &- \int_s^T Z^{t,x}(r) \, dW(r), \end{aligned}$$

and hence it has an explicit solution. From an extension of the classical "variation of constants formula" (see the argument in the proof of the comparison theorem 1.6 in [Pardoux (1998a)]) or by direct verification we get:

$$\begin{aligned} Y^{t,x}(s) = g&\left(X^{t,x}(T)\right) e^{\int_s^T c\left(r, X^{t,x}(r)\right) dr} \\ &+ \int_s^T h\left(r, X^{t,x}(r)\right) e^{\int_s^r c\left(\alpha, X^{t,x}(\alpha)\right) d\alpha} \, dr \\ &- \int_s^T e^{\int_s^r c\left(\alpha, X^{t,x}(\alpha)\right) d\alpha} Z^{t,x}(r) \, dW(r). \end{aligned}$$

Now we have $Y^{t,x}(t) = \mathbb{E}\left[Y^{t,x}(t)\right]$, so that

$$\begin{aligned} &Y^{t,x}(t) \\ &= \mathbb{E}\left[g(X^{t,x}(T))e^{\int_t^T c\left(s,X^{t,x}(s)\right)\,ds} + \int_t^T h\left(s,X^{t,x}(s)\right)e^{\int_t^s c\left(r,X^{t,x}(r)\right)\,dr}ds\right], \end{aligned}$$

which is the well-known Feynman-Kac formula. Clearly, solutions to stochastic backward stochastic differential equations can be used to represent solutions to classical differential equations of parabolic type, and as such they can be considered as a nonlinear extension of the Feynman-Kac formula.

Example 5.3. In this example the family of operators $L(s)$, $0 \leqslant s \leqslant T$, generates a Markov process in the sense of Definition 5.3: see (5.11). For a "smooth" function v we introduce the martingales:

$$M_{v,t}(s) = v\left(s, X(s)\right) - v\left(t, X(t)\right) - \int_t^s \left(\frac{\partial}{\partial\rho} + L(\rho)\right) v\left(\rho, X(\rho)\right) d\rho. \tag{5.197}$$

Its quadratic variation part $\left\langle M_{v,t}\right\rangle(s) := \left\langle M_{v,t}, M_{v,t}\right\rangle(s)$ is given by

$$\left\langle M_{v,t}\right\rangle(s) = \int_t^s \Gamma_1\left(v,v\right)\left(\rho, X(\rho)\right) d\rho.$$

In this example we will mainly be concerned with the Hamilton-Jacobi-Bellman equation as exhibited in (5.198). We have the following result for generators of diffusions: it refines Theorem 2.4 in [Zambrini (1998a)]. Observe that $\mathbb{P}_{t,x}^{M_{v,t}}$ stands for a Girsanov transformation of the measure $\mathbb{P}_{t,x}$.

Theorem 5.8. *Suppose that the operator $L = L(s)$ does not depend on $s \in [0,T]$. Let $\chi : (\tau, T] \times E \to [0,\infty]$ be a function such that for all $\tau < t \leqslant T$ and for sufficiently many functions v*

$$\mathbb{E}_{t,x}^{M_{v,t}}\left[\left|\log\chi\left(T, X(T)\right)\right|\right] < \infty.$$

Let S_L be a (classical) solution to the following Riccati type equation. For $\tau < s \leqslant T$ and $x \in E$ the following identity is true:

$$\begin{cases} \dfrac{\partial S_L}{\partial s}(s,x) - \dfrac{1}{2}\Gamma_1\left(S_L, S_L\right)(s,x) + L(s)S_L(s,x) + V(s,x) = 0; \\ S_L(T,x) = -\log\chi(T,x), \quad x \in E. \end{cases} \tag{5.198}$$

Then for any nice real valued $v(s,x)$ the following inequality is valid:

$$S_L(t,x) \leqslant \mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T \left(\frac{1}{2}\Gamma_1\left(v,v\right) + V\right)(\tau, X(\tau))d\tau\right]$$

$$- \mathbb{E}_{t,x}^{M_{v,t}} \left[\log \chi \left(T, X(T) \right) \right],$$

and equality is attained for the "Lagrangian action" $v = S_L$:

$$S_L(t,x) = -\log \mathbb{E}_{t,x} \left[\exp \left(- \int_t^T V\left(\sigma, X(\sigma)\right) d\sigma \right) \chi\left(T, X(T)\right) \right]. \quad (5.199)$$

The probability $\mathbb{P}_{t,x}^{M_{v,t}}$ is determined by following equality (5.200). For all finite n-tuples $t_1, \dots, t_n$ in $(t,T]$ and all bounded Borel functions $f_j : [t,T] \times E \to \mathbb{R}$, $1 \leqslant j \leqslant n$, we have:

$$\mathbb{E}_{t,x}^{M_{v,t}} \left[\prod_{j=1}^n f_j\left(t_j, X\left(t_j\right)\right) \right] \quad (5.200)$$

$$= \mathbb{E}_{t,x} \left[\exp \left(-\frac{1}{2} \int_t^T \Gamma_1(v,v)\left(\tau, X(\tau)\right) d\tau - M_{v,t}(T) \right) \prod_{j=1}^n f_j\left(t_j, X\left(t_j\right)\right) \right].$$

Proof. This result is proved in Chapter 7: see Theorem 7.1. There is only a notational difference: here we write $L(s)$ instead of $-K_0(s)$ in Theorem 7.1. □

It is just mentioned that Theorem 5.8 is fully proved with $L(s) = L$ time-independent in [Van Casteren (2003)]. In Theorem 5.8 the operator family $\{L(s) : s \in [0,T]\}$ should be the generator of a diffusion process in the sense as in Definition 5.1. In addition, it should generate a Feller evolution in the sense of Theorem 2.11. Moreover, the squared gradient operator should exist in $\mathcal{T}_\beta$-sense, i.e. in the sense of (5.2).

5.4.1 *Remarks on the Runge-Kutta method and on monotone operators*

We conclude this chapter with an explanation of the relation which exists between surjectivity of the mapping $y \mapsto y - \delta f(t,y,z)$, $y \in \mathbb{R}^k$, and the Runge-Kutta method. Here t is a time variable, $\delta > 0$ is a (small) constant, and z is a functional which plays no role here. In the text which follows the z-dependence is suppressed, and h plays the role of δ.

Remark 5.19. The surjectivity of the mapping $y \mapsto y - \delta f(s,y,Z_M(s))$ from $\mathbb{R}^k$ onto itself follows from Theorem 1 in [Crouzeix *et al.* (1983)]. The authors use a homotopy argument to prove this theorem for $C_1 = 0$. Upon replacing $f(t,y,Z_M(t))$ with $f(t,y,Z_M(t)) - C_1 y$, where C_1 is as in (5.72) the result follows in our version, and the conditions in [Crouzeix *et al.*

(1983)] are satisfied. An elementary proof of Theorem 1 in [Crouzeix *et al.* (1983)] can be found for a continuously differentiable function in [Hairer and Wanner (1991)]: see Theorem 14.2 in Chapter IV. The author is grateful to Karel in't Hout (University of Antwerp) for pointing out closely related Runge-Kutta type results and these references. In [Hairer and Wanner (1991)] and also in the newer version [Hairer and Wanner (1996)] (Theorem 14.2) the authors study the existence of a Runge-Kutta solution $(g_j)_{1\leqslant j\leqslant s}$, $g_j \in \mathbb{R}^k$, which is implicitly defined by an equation of the form

$$\begin{aligned} g_i &= y_0 + h\sum_{j=1}^{s} a_{ij} f\left(t_0 + c_j h, g_j\right), \; i = 1, \dots, s, \\ y_1 &= y_0 + h\sum_{j=1}^{s} b_j f\left(t_0 + c_j h, g_j\right). \end{aligned} \tag{5.201}$$

Here c_j and b_j are given constants which depend on the precise numerical method under discussion, and the same is true for the constants a_{ij}, $1 \leqslant i, j \leqslant s$. The equations in (5.201) are motivated by a numerical treatment of ordinary differential equations of the form $\dfrac{\partial y(t)}{\partial t} = f\left(t, y(t)\right)$, $y(0) = y_0$. The function f satisfies a one-sided Lipschitz condition of the form

$$\left\langle f\left(t, y_2\right) - f\left(t, y_1\right), y_2 - y_1\right\rangle \leqslant C\left|y_2 - y_1\right|^2 \tag{5.202}$$

for all t in an open interval of $\mathbb{R}$ and for all y_1, $y_2 \in \mathbb{R}^k$. Here the symbol y with or without subscript is a vector in $\mathbb{R}^k$. The Runge-Kutta matrix $A = \left(a_{ij}\right)_{i,j=1}^{s}$ is supposed to be an invertible $s \times s$ matrix. The vector y_1 is the new initial condition. Put

$$\alpha_0\left(A^{-1}\right) = \sup_D \inf_{u\in\mathbb{R}^k,\, u\neq 0} \frac{\left\langle u, DA^{-1}u\right\rangle}{\langle u, Du\rangle}$$

where the supremum is taken over all diagonal matrices D with strictly positive entries. If A is the identity matrix, then $\alpha_0\left(A^{-1}\right) = 1$. In terms of the matrix A and the mapping

$$F : \left(g_1, \dots, g_s\right) \mapsto \left(f\left(t_0 + c_1 h, g_1\right), \dots, f\left(t_0 + c_s h, g_s\right)\right)$$

the solvability of the Runge-Kutta equation (5.201) for all initial values $y_0 \in \mathbb{R}^k$ is equivalent to the surjectivity of the mapping $g \mapsto g - hAF(g)$, $g \in \mathbb{R}^k$.

Theorem 5.9. *Let $h > 0$ be such that $hC < \alpha_0\left(A^{-1}\right)$, where C is as in (5.202). Then the equation in (5.201) has a solution.*

Under the extra assumption of continuously differentiability of the function f the authors of [Hairer and Wanner (1991)] base their proof on a study of the homotopy properties of the mapping:

$$g_i = y_0 + h\sum_{j=1}^{s} a_{ij} f\left(t_0 + c_j h, g_j\right) + (\tau - 1)h\sum_{j=1}^{s} a_{ij} f\left(t_0 + c_j h, y_0\right). \quad (5.203)$$

The equation in (5.203) has a solution $g_j = y_0$ for $\tau = 0$, and for $\tau = 1$ it reduces to the equation in (5.201).

Remark 5.20. Finally we notice that in [Ramm (2007)] the author treats one-sided Lipschitz and monotone operators in the context of Hilbert and Banach spaces. He uses the so-called Dynamical System Method. We also mention that for Hilbert spaces the problem of surjectivity of the operator $I - \delta F$ is closely related to the fact that the operator $-F$ is m-accretive in the sense that F is one-sided monotone with monotonicity constant 0, and that $I - \delta F$ is surjective for some, and by Minty's theorem, for all $\delta > 0$: for details see [Showalter (1997)]. For a closely related result, called the Browder-Minty theorem, see Theorem 9.45 in [Renardy and Rogers (2004)] or Theorem 2.2 in [Showalter (1997)].

Theorem 5.10. *The Browder-Minty theorem states that a bounded, demi-continuous, coercive and monotone function T from a real, reflexive Banach space X into its continuous dual space X^* is automatically surjective. That is, for each continuous linear functional $g \in X^*$, there exists a solution $u \in X$ of the equation $T(u) = g$.*

The (non-linear) operator T is called coercive if the equality $\lim_{|y|\to\infty} \frac{\langle y, Ty\rangle}{|y|} = \infty$ holds, and monotonicity means that $\langle y_2 - y_1, Ty_2 - Ty_1\rangle \geqslant 0$ for all y_1, $y_2 \in X$. The operator T is said to be demi-continuous if $u_n \to u$ in X implies $\langle x, Tu_n - Tu\rangle \to 0$ for all $x \in X$. It is bounded if it sends bounded sets to bounded sets. The operator T is said to be hemi-continuous if the function $t \mapsto \langle x, T(x + tu)\rangle \to 0$ is continuous for all x, $u \in X$. The result in Theorem 5.10 was proved independently by Minty [Minty (1963)] and Browder [Browder (1963)]: see [Browder (1967)] as well. By a result due to Browder and Rockafellar for monotone operators hemi-continuity and demi-continuity are equivalent: see [Rockafellar (1997)] and [Rockafellar (1969)]. It is noticed that in order to pass from the finite-dimensional to the infinite-dimensional setting authors use a Galerkin method. In view of Theorem 5.10 it is quite well possible that the results in this chapter can also be formulated and proved in the Hilbert space context, i.e. when the variables $Y(t)$ and $M(t)$ take their values in a Hilbert or (reflexive) Banach space instead of $\mathbb{R}^k$: see [Browder *et al.* (1970)]. Theorem 2.25 in [Phelps (1993)] states that the multi-valued sub-differential of a continuous convex function which is everywhere defined in a Banach space is a maximal monotone operator. For set valued monotone operators the reader is referred to [Tarafdar and Chowdhury (2008)].

To conclude this chapter we insert a sample result in the Hilbert space setting: for details see Van [Van Casteren (2010)]. In what follows the symbol $\mathbb{H}$ stands for a Hilbert space, and the family $\{C(t,s) : 0 \leqslant t \leqslant s \leqslant T\}$ stands for a strongly continuous family of linear operators on $\mathbb{H}$ which is a forward evolution family, i.e. $C(t_1,s)\,C(s,t_2) = C(t_1,t_2)$, $0 \leqslant t_1 \leqslant s \leqslant t_2 \leqslant T$, and $C(t,t) = I$. In addition we write $A(t)h = \lim_{s\downarrow 0} \dfrac{C(t,t+s)h-h}{s}$, $h \in D(A(t))$. Let E be a Polish space, and

$$\left\{\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x}\right),\left(X(t),\tau\leqslant t\leqslant T\right),\left(E,\mathcal{E}\right)\right\} \tag{5.204}$$

an E-valued strong time-dependent Markov process with continuous paths, $\mathbb{H}$ a Hilbert space, and $u : [0,T]\times E \to \mathbb{H}$ a function with the property that the limit

$$\lim_{\rho\downarrow t}\frac{\mathbb{E}_{t,x}\left[u\left(\rho,X(\rho)\right)\right]-u(t,x)}{\rho-t} = \left(\frac{\partial}{\partial t}+L(t)\right)u\left(t,x\right) \tag{5.205}$$

exists for all $(t,x) \in [0,T]\times E$. By hypothesis it is assumed that this convergence takes place in the Hilbert space $\mathbb{H}$ and is uniformly on compact subsets of $[0,T]\times E$. Observe that the operators $L(s)$, $0 \leqslant s \leqslant T$, are defined on a subspace of the space of continuous $\mathbb{H}$-valued functions. The equality in (5.206) below should be compared with equality (2.77) in Definition 2.8.

Theorem 5.11. *Let $\mathbb{H}$ be a real Hilbert space. Let $u : [0,T]\times E \to \mathbb{H}$ be a continuous function with the property that for every $(t,x) \in [0,T]\times E$ the function $s \mapsto \mathbb{E}_{t,x}\left[u\left(s,X(s)\right)\right]$ is differentiable and that for the derivatives from the right*

$$\frac{d}{ds}\mathbb{E}_{t,x}\left[u\left(s,X(s)\right)\right] = \mathbb{E}_{t,x}\left[L(s)u\left(s,X(s)\right)+\frac{\partial}{\partial s}u\left(s,X(s)\right)\right], \quad t\leqslant s<T. \tag{5.206}$$

Then the following assertions are equivalent:

(a) The function u satisfies the following differential equation:

$$L(t)u\left(t,x\right)+A(t)u(t,x)+\frac{\partial}{\partial t}u\left(t,x\right)+f\left(t,x,u\left(t,x\right),\nabla_u^L\left(t,x\right)\right)=0.$$

(b) The function u satisfies the following type of Feynman-Kac integral equation:

$$u(t,x) = \mathbb{E}_{t,x}\left[C(t,T)u(T,X(T))\right] + \mathbb{E}_{t,x}\left[\int_t^T C(t,\tau) f\left(\tau, X(\tau), u(\tau, X(\tau)), \nabla_u^L(\tau, X(\tau))\right) d\tau\right].$$

(c) For every $t \in [0,T]$ there exists a $\mathbb{P}_{t,x}$-martingale $M_t(s)$ on the interval $[t,T]$ such that for $t \leqslant s \leqslant T$

$$\begin{aligned} u(t,X(t)) &= C(t,s)u(s,X(s)) \\ &\quad + \int_t^s C(t,\tau) f\left(\tau, X(\tau), u(\tau, X(\tau)), \nabla_u^L(\tau, X(\tau))\right) d\tau \\ &\quad - \int_t^s C(t,\tau)\, dM_t(\tau). \end{aligned}$$

The result in Theorem 5.11 should be compared with Theorem 5.1. In [Van Casteren (2010)] conditions are given in order that an equation of the form

$$Y(t) = C(t,T)\xi + \int_t^T C(t,s) f(s, X(s), Y(s), Z_M(s))\, ds - \int_t^T C(t,s)\, dM(s).$$

admits solutions in an appropriate stochastic phase space $\mathcal{S}^2 \times \mathcal{M}^2$: cf. equality (5.196) in Theorem 5.7. Again, as in the remaining part of this chapter the pair of $\mathbb{H}$-valued processes $(Y(t), M(t))$ is adapted to the underlying (strong) Markov process (5.204). Again one-sided Lipschitz conditions play a role (in the variable Y) and a two-sided Lipschitz condition is required in the variable Z_M. Instead of the identity operator I as in Theorem 5.7 we now have a propagator $\{C(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$. It generator $t \mapsto A(t)$ is supposed to be bounded from above: there exists a constant C_0 such that $\langle y, A(s)y\rangle_{\mathbb{H}} \leqslant C_0 \langle y, y\rangle_{\mathbb{H}}$, $0 \leqslant s \leqslant T$, $y \in D(A(s))$.

This page is intentionally left blank

Chapter 6

Viscosity solutions, backward stochastic differential equations and Markov processes

In this chapter we explain the notion of stochastic backward differential equations and its relationship with classical (backward) parabolic differential equations of second order. The chapter contains a combination of stochastic processes like Markov processes and martingale theory and semi-linear partial differential equations of parabolic type. Emphasis is put on the fact that the solutions to BSDE's obtained by stochastic methods to BSDE's are often viscosity solutions.

The notations

$$\mathcal{S}^2\left([\tau,T],\mathbb{P}_{\tau,x},\mathbb{R}^k\right) = \mathcal{S}^2\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right) \quad \text{and}$$
$$\mathcal{M}^2\left([\tau,T],\mathbb{P}_{\tau,x},\mathbb{R}^k\right) = \mathcal{M}^2\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right)$$

were explained in the definitions 5.4 and 5.5 respectively. The same is true for the notions

$$\begin{aligned}
\mathcal{S}^2_{\text{loc,unif}}\left([0,T],\mathbb{R}^k\right) &= \mathcal{S}^2_{\text{loc,unif}}\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right),\\
\mathcal{M}^2_{\text{loc,unif}}\left([0,T],\mathbb{R}^k\right) &= \mathcal{M}^2_{\text{loc,unif}}\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right),\\
\mathcal{S}^2_{\text{unif}}\left([0,T],\mathbb{R}^k\right) &= \mathcal{S}^2_{\text{unif}}\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right), \quad \text{and}\\
\mathcal{M}^2_{\text{unif}}\left([0,T],\mathbb{R}^k\right) &= \mathcal{M}^2_{\text{unif}}\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right).
\end{aligned}$$

The space $\mathcal{M}^{2,s}_{\text{AC}} = \mathcal{M}^{2,s}_{\text{AC}}\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x};\mathbb{R}^k\right)$ is explained in Definition 5.7 (see Lemma 5.1 as well). The probability measure $\mathbb{P}_{\tau,x}$ is defined on the σ-field $\mathcal{F}_T^\tau$. Since the existence properties of the solutions to backward stochastic equations are based on explicit inequalities, the proofs carry over to Markov families of measures. Ultimately these inequalities imply that boundedness and continuity properties of the function $(\tau,x)\mapsto\mathbb{E}_{\tau,x}\left[Y(t)\right]$, $0\leqslant\tau\leqslant t\leqslant T$, depend on the continuity of the function $x\mapsto\mathbb{E}_{T,x}\left[\xi\right]$, where ξ is a terminal value function which is supposed to be $\sigma\left(X(T)\right)$-measurable. In addition, in order to be sure that the function $(\tau,x)\mapsto\mathbb{E}_{\tau,x}\left[Y(t)\right]$ is continuous, functions of the form $(\tau, x) \mapsto \mathbb{E}_{\tau,x}\left[f\left(t, u\left(t, X(t)\right), Z_M(t)\right)\right]$ have to be continuous, whenever the following mappings

$$\begin{aligned}
&(\tau, x) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T |u(s, X(s))|^2 \, ds\right] \quad \text{and} \\
&(\tau, x) \mapsto \mathbb{E}_{\tau,x}\left[\langle M, M\rangle(T) - \langle M, M\rangle\right]
\end{aligned} \tag{6.1}$$

represent finite and continuous functions. Comparison theorems enable us to compare solutions if these solutions can be compared at their endpoints. In the proof of these comparison theorems we introduced a new martingale: see formula (6.3). They also serve to prove that solutions to BSDE's often are viscosity solutions: see e.g. Theorem 6.3.

6.1 Comparison theorems

As an introduction to the present section we insert a comparison theorem. This theorem will also be used to establish the fact that solutions to semi-linear BSDE's are in fact viscosity solutions. In the following theorem the measure $\mathbb{P}$ could be one of the probability measures $\mathbb{P}_{0,x}$, $x \in E$. If the interval $[\tau, T]$ is taken instead of $[0, T]$ then $\mathbb{P}$ could also be one of the measures $\mathbb{P}_{\tau,x}$, and, of course, $\mathcal{F}_T$ should be replaced with $\mathcal{F}_T^\tau$. Recall that the space $\mathcal{M}_{\mathrm{AC}}^{2,t}$ is explained in Definition 5.7.

Theorem 6.1. *Suppose that $Y(T) = \xi \leqslant \xi' = Y'(T)$ $\mathbb{P}$-a.s., and $f(t,x,y,z) \leqslant f'(t,x,y,z)$ almost everywhere. Then $Y(t) \leqslant Y'(t)$, $0 \leqslant t \leqslant T$, $\mathbb{P}$-a.s., provided that there exists a martingale $N(t)$ such that the quadratic covariation process $t \mapsto \langle N, M' - M\rangle(t)$ satisfies*

$$f'\left(t, X(t), Y(t), Z_{M'}(t)\right) - f'\left(t, X(t), Y(t), Z_M(t)\right) = \frac{d}{dt}\left\langle N, M' - M\right\rangle(t). \tag{6.2}$$

If moreover $Y(0) = Y'(0)$, then $Y(t) = Y'(t)$, $0 \leqslant t \leqslant T$, $\mathbb{P}$-a.s. Moreover, if either $\mathbb{P}(\xi < \xi') > 0$ or $f(t, y, Z_M(t)) < f'(t, y, Z_M(t))$, $(y, Z_M(t)) \in \mathbb{R} \times \mathcal{M}_{\mathrm{AC}}^{2,t}$, on a set of positive $dt \times d\mathbb{P}$ measure, then $Y(0) < Y'(0)$.

In fact for the martingale $N(t)$ in (6.2) we may choose:

$$\begin{aligned}
N(t) = \int_0^t & \frac{f'\left(s, X(s), Y(s), Z_{M'}(s)\right) - f'\left(s, X(s), Y(s), Z_M(s)\right)}{\dfrac{d}{ds}\left\langle M' - M, M' - M\right\rangle(s)} \\
& \left(dM'(s) - dM(s)\right),
\end{aligned} \tag{6.3}$$

where the derivative

$$\frac{d}{ds}\left\langle M'-M,M'-M\right\rangle(s)$$

stands for the Radon-Nikodym derivative of the quadratic variation process $t\mapsto\langle M'-M,M'-M\rangle(t)$ at $t=s$ (relative to the Lebesgue measure). For more explanation see Definition 5.7 and Lemma 5.1. In the following proposition we collect some properties of the martingale $t\mapsto N(t)$. Among other things it says that the process $t\mapsto N(t)$ is well-defined and continuous provided the martingale $t\mapsto M'(t)-M(t)$ is continuous. It is assumed that there exists a constant C' such that

$$\begin{aligned}&\left|f'\left(s,X(s),Y(s),Z_{M'}(s)\right)-f'\left(s,X(s),Y(s),Z_M(s)\right)\right|^2\\&\leqslant\left(C'\right)^2\frac{d}{ds}\left\langle M'-M,M'-M\right\rangle(s),\quad 0\leqslant s\leqslant T.\end{aligned}\tag{6.4}$$

Proposition 6.1. *Suppose that the processes $X(s)$, $Y(s)$, $M'(s)$, and $M(s)$ are such that (6.4) is satisfied for the constant C'. In addition suppose that the process $M'-M$ is a martingale belonging to $\mathcal{M}^2([0,T],\mathbb{P})$ with the property that the quadratic variation process $s\mapsto\langle M'-M,M'-M\rangle(s)$ is absolutely continuous with respect to the Lebesgue measure. Then the process $t\mapsto N(t)$ is a martingale which is well-defined, and also belongs to $\mathcal{M}^2([0,T],\mathbb{P})$. The following inequality is satisfied:*

$$\langle N,N\rangle(t)-\langle N,N\rangle(s)\leqslant\left(C'\right)^2(t-s).\tag{6.5}$$

The quadratic variation process $t\mapsto\langle N,N\rangle(t)$ is absolutely continuous relative to the Lebesgue measure. Its Radon-Nikodym derivative $\frac{d}{ds}\langle N,N\rangle(s)$ satisfies

$$\frac{d}{ds}\langle N,N\rangle(s)=\frac{\left|f'\left(s,X(s),Y(s),Z_{M'}(s)\right)-f'\left(s,X(s),Y(s),Z_M(s)\right)\right|^2}{\frac{d}{ds}\left\langle M'-M,M'-M\right\rangle(s)}.\tag{6.6}$$

In the notation of Definition 5.6 the martingale $M'-M$ belongs to the space $\mathcal{M}^2_{\mathrm{AC}}(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R})=\mathcal{M}^2_{\mathrm{AC}}([0,T],\mathbb{P})$.

Let $s\mapsto M_1(s)$ and $s\mapsto M_2(s)$ be two martingales with quadratic variation processes $\langle M_1,M_1\rangle$ and $\langle M_2,M_2\rangle$ respectively. Let the Doléans measures $Q_j:\mathcal{F}^0_T\otimes\mathcal{B}_{[0,T]}\to[0,\infty]$, $j=1,2$ be determined by

$$Q_j(A\times(a,b])=\mathbb{E}\left[\mathbf{1}_A\left(\langle M_j,M_j\rangle(b)-\langle M_j,M_j\rangle(a)\right)\right],\tag{6.7}$$

with $A\in\mathcal{F}^0_T$, $0\leqslant a\leqslant b\leqslant T$, $j=1,2$. In addition, let $s\mapsto f_1(s)$ and $s\mapsto f_2(s)$ be predictable process which belong to $L^2\left(\Omega,\mathcal{F}^0_T\otimes\mathcal{B}_{[0,T]},Q_1\right)$

and $L^2\left(\Omega, \mathcal{F}_T^0 \otimes \mathcal{B}_{[0,T]}, Q_2\right)$ respectively. In the proof of Proposition 6.1 we need the following equality:

$$\left\langle \int_0^{(\cdot)} f_1(s) dM_1(s), \int_0^{(\cdot)} f_2(s) dM_2(s) \right\rangle (t) = \int_0^t f_1(s) f_2(s) d \left\langle M_1, M_2 \right\rangle (s) \tag{6.8}$$

where $t \in [0,T]$. A definition of Doléans measure like (6.7), and an equality like (6.8) are given in books on martingale theory, like [Williams (1991)].

Proof. [Proof of Proposition 6.1.] Equality (6.8) in Proposition 6.1 yields

$$\begin{aligned}
&\langle N, N \rangle (t) \\
&= \int_0^t \frac{\left| f'\left(s, X(s), Y(s), Z_{M'}(s)\right) - f'\left(s, X(s), Y(s), Z_M(s)\right)\right|^2}{\left(\frac{d}{ds}\left\langle M' - M, M' - M \right\rangle (s)\right)^2} \\
&\qquad d\left\langle M' - M, M' - M\right\rangle (s) \\
&= \int_0^t \frac{\left| f'\left(s, X(s), Y(s), Z_{M'}(s)\right) - f'\left(s, X(s), Y(s), Z_M(s)\right)\right|^2}{\left(\frac{d}{ds}\left\langle M' - M, M' - M \right\rangle (s)\right)^2} \\
&\qquad \frac{d}{ds}\left\langle M' - M, M' - M\right\rangle (s)\, ds \\
&= \int_0^t \frac{\left| f'\left(s, X(s), Y(s), Z_{M'}(s)\right) - f'\left(s, X(s), Y(s), Z_M(s)\right)\right|^2}{\frac{d}{ds}\left\langle M' - M, M' - M \right\rangle (s)}\, ds.
\end{aligned} \tag{6.9}$$

The equality in (6.5) follows from (6.9). Combining the equality in (6.4) and (6.9) results in the inequality in (6.5). The inequality in (6.6) follows from (6.5).

If the martingale $s \mapsto (M'(s) - M(s))$ is continuous, then so is the martingale $s \mapsto N(s)$ which is obtained as a stochastic integral relative to $d\left(M' - M\right)(s)$. This assertion also follows from Itô calculus for martingales: see e.g. [Williams (1991)].

This completes the proof of Proposition 6.1. □

Proof. [Proof of Theorem 6.1.] Following [Pardoux (1998a)] we introduce the process $\alpha(t)$, $0 \leqslant t \leqslant T$, by $\alpha(t) = 0$ if $Y(t) = Y'(t)$, and

$$\begin{aligned}
&\alpha(t) \\
&= \left(Y'(t) - Y(t)\right)^{-1}\left(f'\left(t, X(t), Y'(t), Z_{M'}(t)\right) - f'\left(t, X(t), Y(t), Z_{M'}(t)\right)\right)
\end{aligned} \tag{6.10}$$

if $Y(t) \neq Y'(t)$. Then $\alpha(t) \leqslant C_1$ $\mathbb{P}$-almost surely. We also introduce the following processes:

$$U(t) = f'\left(t, X(t), Y(t), Z_M(t)\right) - f\left(t, X(t), Y(t), Z_M(t)\right); \tag{6.11}$$

$$\overline{Y}(t) = Y'(t) - Y(t); \tag{6.12}$$
$$\overline{M}(t) = M'(t) - M(t); \tag{6.13}$$
$$\overline{\xi} = \overline{Y}(T) = Y'(T) - Y(T) = \xi' - \xi. \tag{6.14}$$

In terms of $\alpha(t)$, $\overline{\xi}$, $U(t)$, and the martingales $N(t)$ and $\overline{M}(t)$ the adapted process $\overline{Y}(t)$ satisfies the following backward integral equation:

$$\begin{aligned}&\overline{Y}(t) - \overline{\xi} \\ &= \int_t^T \alpha(s)\overline{Y}(s)ds + \int_t^T U(s)ds - \overline{M}(T) + \overline{M}(t) + \left\langle N, \overline{M}\right\rangle(T) - \left\langle N, \overline{M}\right\rangle(t).\end{aligned} \tag{6.15}$$

From Itô calculus and (6.15) it then follows that

$$\begin{aligned}\overline{Y}(t) &= \overline{Y}(T)e^{\int_t^T \alpha(\tau)d\tau - \frac{1}{2}\langle N,N\rangle(T) + \frac{1}{2}\langle N,N\rangle(t) + N(T) - N(t)} \\ &+ \int_t^T e^{\int_t^s \alpha(\tau)d\tau - \frac{1}{2}\langle N,N\rangle(s) + \frac{1}{2}\langle N,N\rangle(t) + N(s) - N(t)}(U(s)ds - d\overline{M}(s) - \overline{Y}(s)dN(s)).\end{aligned} \tag{6.16}$$

Since the process $\overline{Y}(t)$ is adapted and since Itô integrals with respect to martingales with bounded integrands are martingales the equality in (6.16) implies:

$$\begin{aligned}\overline{Y}(t) = \mathbb{E}\Bigg[&\overline{Y}(T)e^{\int_t^T \alpha(\tau)d\tau - \frac{1}{2}\langle N,N\rangle(T) + \frac{1}{2}\langle N,N\rangle(t) + N(T) - N(t)} \\ &+ \int_t^T e^{\int_t^s \alpha(\tau)d\tau - \frac{1}{2}\langle N,N\rangle(T) + \frac{1}{2}\langle N,N\rangle(t) + N(s) - N(t)}U(s)ds \;\Big|\; \mathcal{F}_t\Bigg].\end{aligned} \tag{6.17}$$

Since by hypothesis $\overline{Y}(T) \geqslant 0$ and $U(s) \geqslant 0$ for all $s \in [0,T]$, the equality in (6.17) implies $\overline{Y}(t) \geqslant 0$. The other assertions also follow from representation (6.17). This completes the proof of Theorem 6.1. □

The following result can be proved along the same lines as Theorem 6.1. It will be used in the proof of Theorem 6.3 with

$$V(s) = \dot{\varphi}(s, X(s)) + L(s)\varphi(s,\cdot)(X(s)),$$

with $Y(s) = u(s, X(s))$ and $Y'(s) = \varphi(s, X(s))$. In fact the arguments in the proof of Theorem 2.4 of [Pardoux (1998a)] inspired our proof of the following theorem.

Theorem 6.2. *Fix $(t,x) \in [0,T) \times E$ and fix a stopping time τ such that $t < \tau \leqslant T$. Let $V(s)$ be a progressively measurable process such that $\mathbb{E}_{t,x}\left[\int_t^\tau |V(s)|\, ds\right] < \infty$. Let (Y,M) and $(Y',M') \in \mathcal{S}^2([t,T], \mathbb{P}_{t,x}, \mathbb{R}) \times$*

$\mathcal{M}^2([t,T],\mathbb{P}_{t,x},\mathbb{R})$ *satisfy the following type of backward stochastic integral equations:*

$$Y(s)=Y(\tau)+\int_s^\tau f(\rho,X(\rho),Y(\rho),Z_M(\rho))\,d\rho+M(s)-M(\tau)\quad \textit{and}$$
$$Y'(s)=Y'(\tau)+\int_s^\tau V(\rho)d\rho+M'(s)-M'(\tau)$$

for $t\leqslant s<\tau$. *Suppose that* $Y(\tau)\leqslant Y'(\tau)$ *and*

$$f\left(s,X(s),Y'(s),Z_{M'}(s)\right)\leqslant V(s),\quad t\leqslant s\leqslant\tau.$$

Then $Y(s)\leqslant Y'(s)$, $t\leqslant s\leqslant T$. *If*

$$f\left(s,X(s),Y'(s),Z_{M'}(s)\right)<V(s)$$

on a subset of $[t,\tau)\times\Omega$ *of strictly positive* $ds\times\mathbb{P}$*-measure, then* $Y(t)<Y'(t)$.

Proof. Define the stochastic process $f'(s,X(s),y,z)$ by

$$f'(s,X(s),y,z)=f(s,X(s),y,z)+V(s)-f\left(s,X(s),Y'(s),Z_{M'}(s)\right).$$

The arguments for the proof of Theorem 6.1 now apply with the martingale $N(s)$, $t\leqslant s\leqslant T$, given by

$$N(s)=\int_t^{s\wedge\tau}\frac{f(\rho,X(\rho),Y(\rho),Z_{M'}(\rho))-f(\rho,X(\rho),Y(\rho),Z_M(\rho))}{\frac{d}{d\rho}\left\langle M'-M,M'-M\right\rangle(\rho)}\left(dM'(\rho)-dM(\rho)\right), \tag{6.18}$$

and the process $\alpha(s)$, $t\leqslant s\leqslant T$, defined by $\alpha(s)=0$ if $Y(s)=Y'(s)$, and

$$\alpha(s) \tag{6.19}$$
$$=\left(Y'(t)-Y(t)\right)^{-1}\left(f\left(s,X(s),Y'(s),Z_{M'}(s)\right)-f\left(s,X(s),Y(s),Z_{M'}(s)\right)\right)$$

if $Y(t)\neq Y'(t)$. The other relevant processes are:

$$U(s)=V(s)-f'\left(s,X(s),Y'(s),Z_{M'}(s)\right); \tag{6.20}$$
$$\overline{Y}(s)=Y'(s)-Y(s); \tag{6.21}$$
$$\overline{M}(s)=M'(s)-M(s); \tag{6.22}$$
$$\overline{\xi}=\overline{Y}(\tau)=Y'(\tau)-Y(\tau)=\xi'-\xi. \tag{6.23}$$

The remaining reasoning follows the lines of the proof of Theorem 6.1.

This completes the proof of Theorem 6.2. □

Remark 6.1. If $Y(s) = u(s, X(s))$, u is "smooth", and $u(t,x)$ satisfies (5.84), which is the same as (6.36) below, then $Y(s)$ satisfies (5.85), and vice versa. If $f(s,x,y,z)$ only depends on $y \in \mathbb{R}$, then, by the occupation formula,

$$\begin{aligned}\int_t^T g(Y(s))\, Z(s)(Y,Y)\, ds &= \int_t^T g(Y(s))\, d\langle Y(\cdot), Y(\cdot)\rangle (s) \\ &= \int_{\mathbb{R}} \left(L_T^y(Y) - L_t^y(Y)\right) g(y) dy,\end{aligned}$$

where dy is the Lebesgue measure, and $L_t^y(Y)$ is the (density of the) local time of the process $Y(t)$. If $g \equiv 1$ and $Y(s) = u(s, X(s))$, then (5.85) is also equivalent to the following assertion: the process

$$\exp\left(Y(s) - Y(T) - \int_s^T \left(f(\tau, X(\tau), Y(\tau), Z(\tau)(\cdot, Y)) - \frac{1}{2}\langle Y, Y\rangle (\tau)\right) d\tau\right),$$

$t_0 < t \leqslant s \leqslant T$, is a local backward (exponential) $\mathbb{P}_{t,x}$-martingale (for every $T > t > t_0$). The function f depends on $x \in E$, $s \in (t_0, T]$, $y \in \mathbb{R}$, and on the square gradient operator $(f_1, f_2) \mapsto \Gamma_1(f_1, f_2)$, or, more generally, on the covariance mapping $(Y_1, Y_2) \mapsto \langle Y_1, Y_2\rangle (s)$ of the local semi-martingales $Y_1(s)$ and $Y_2(s)$. In order to introduce boundary conditions it is required to insert in equation (5.85) a term of the form

$$\int_t^T h(X(s), s, Y(s), Z(s)(\cdot, Y))\, dA(s),$$

where $A(s)$ is a process which is locally of bounded variation, and which only increases when e.g. $X(s)$ hits the boundary. To be more precise the equality in (5.85) should be replaced with:

$$\begin{aligned}Y(t) - Y(T) &- \int_t^T f(s, X(s), Y(s), Z(s)(\cdot, Y))\, ds \\ &- \int_t^T h(X(s), s, Y(s), Z(s)(\cdot, Y))\, dA(s) = M(t) - M(T). \qquad (6.24)\end{aligned}$$

We hope to come back on this and similar problems in future work. In order to be sure about uniqueness and existence of solutions we probably will need some Lipschitz and linear growth conditions on the function f and some boundedness condition on φ. For more details on backward stochastic differential equations see e.g. [Pardoux and Peng (1990)] and [Pardoux (1998a)].

6.2 Viscosity solutions

The main result in this section is Theorem 6.3. We begin with some formal definitions.

Definition 6.1. Fix $t_0 \in [0,T]$, and let

$$\begin{aligned} F : &C\left([t_0,T]\times E,\mathbb{R}\right)\times C\left([t_0,T]\times E,\mathbb{R}\right)\times C\left([t_0,T]\times E,\mathbb{R}\right) \\ &\times \mathcal{L}\left(C^{(0,1)}\left([t_0,T]\times E,\mathbb{R}\right), C\left([t_0,T]\times E,\mathbb{R}\right)\right) \\ &\to C\left([t_0,T]\times E,\mathbb{R}\right) \end{aligned}$$

be a function with the following property. If (t,x) is any point in $[t_0,T]\times E$, then for all functions φ and ψ belonging to $C\left([t_0,T]\times E,\mathbb{R}\right)$, for which the 4 functions

$$(s,y)\mapsto \dot{\varphi}(s,y), \quad (s,y)\mapsto L(s)\varphi(s,\cdot)(y), \tag{6.25}$$

$$(s,y)\mapsto \dot{\psi}(s,y), \quad \text{and} \quad (s,y)\mapsto L(s)\psi(s,\cdot)(y) \tag{6.26}$$

belong to $C_b\left([t_0,T]\times E,\mathbb{R}\right)$, for which the operators $g\mapsto \nabla^L_\varphi(g)$ and $g\mapsto \nabla^L_\psi(g)$ are $\mathcal{T}_\beta$-continuous mappings from $D(\Gamma_1)$ to $C_b\left([t_0,T]\times E\right)$, and which are such that in case

$$\dot{\varphi}(t,x)=\dot{\psi}(t,x),\ \Gamma_1(\varphi-\psi,\varphi-\psi)(t,x)=0,\ L(t)\varphi(t,x)\leqslant L(t)\psi(t,x), \text{and}$$
$$\varphi(t,x)=\psi(t,x) \tag{6.27}$$

it follows that

$$F\left(\dot{\varphi},L\varphi,\varphi,\nabla^L_\varphi\right)(t,x)\leqslant F\left(\dot{\psi},L\psi,\psi,\nabla^L_\psi\right)(t,x).$$

Here we wrote

$$\dot{\varphi}=\frac{\partial\varphi}{\partial t}, \quad L\varphi(t,x)=\left[L(t)\varphi(t,\cdot)\right](x), \quad \text{and} \quad \nabla^L_\varphi g(t,x)=\Gamma_1(\varphi,g)(t,x).$$

Of course, similarly notions are in vogue for the function ψ. It is noticed that

$$\Gamma_1(\varphi-\psi,\varphi-\psi)(t,x)=0$$

if and only if the equality

$$\nabla^L_\varphi f(t,x)=\nabla^L_\psi f(t,x) \quad \text{holds for all } f\in C^{(0,1)}\left([0,T]\times E,\mathbb{R}\right). \tag{6.28}$$

The proof of this assertion uses the inequality

$$\left|\Gamma_1(\varphi-\psi,f)(t,x)\right|^2\leqslant \Gamma_1(\varphi-\psi,\varphi-\psi)(t,x)\Gamma_1(f,f)(t,x) \tag{6.29}$$

together with the identity $\nabla^L_{\varphi-\psi}(f)(t,x) = \Gamma_1(\varphi-\psi,f)(t,x)$. If $f = \varphi - \psi$ we have equality in (6.29). An example of such a function F is:

$$F(\varphi_1,\varphi_2,\varphi_3,\chi)(t,x) = \varphi_1(t,x) + \varphi_2(t,x) + f(t,x,\varphi_3(t,x),\chi(t,x)), \tag{6.30}$$

where $\chi(t,x)$ is the linear functional $g \mapsto \chi(g)(t,x)$. A *viscosity sub-solution* for the equation

$$F\left(\dot{w}, Lw, w, \nabla^L_w\right)(t,x) = 0, \quad w(T,x) = g(x) \tag{6.31}$$

is a continuous function w with the following properties. First of all $w(T,x) \leqslant g(x)$, and if $\varphi : [t_0,T] \times E \to \mathbb{R}$ is any "smooth function" (i.e. all three functions $\dot{\varphi}$, $L\varphi$, φ are continuous and the linear mapping $\psi \mapsto \nabla^L_\varphi \psi = \Gamma_1(\psi,\varphi)$ is continuous as well) $\Gamma_1(\varphi,\varphi)$, $L(s)\varphi$ belong to $C([t_0,T] \times E, \mathbb{R})$), and if (t,x) is any point in $[t_0,T) \times E$ where the function $w - \varphi$ vanishes and attains a (local) maximum, then

$$F\left(\dot{\varphi}, L\varphi, w, \nabla^L_\varphi\right)(t,x) \geqslant 0. \tag{6.32}$$

The function w is a *super-solution* for equation (6.31) if $w(T,x) \geqslant g(x)$, and if for any "smooth" function φ with the property that the function $w - \varphi$ vanishes and attains a (local) minimum at any point $(t,x) \in [t_0,T) \times E$, then

$$F\left(\dot{\varphi}, L\varphi, w, \nabla^L_\varphi\right)(t,x) \leqslant 0. \tag{6.33}$$

If a function w satisfies (6.32) as well as (6.33) then w is called a *viscosity solution* to equation (6.31).

The definition of the space $D(\Gamma_1)$ was given in 5.3. The following result says essentially speaking that solutions to BSDE's in (6.35) and viscosity solutions to equation (5.84), which is the same as (6.36) below, are intimately related in the sense that $u(t,x) = \mathbb{E}_{t,x}[Y(t)]$, and conversely $Y(t) = u(t,X(t))$. As in Section 5.1 the family of operators $L(s)$, $0 \leqslant s \leqslant T$, generates a Markov process:

$$\left\{(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}), (X(t) : T \geqslant t \geqslant 0), (\vee_t, T \geqslant t \geqslant 0), (E, \mathcal{E})\right\}. \tag{6.34}$$

Theorem 6.3. *Let the ordered pair* $\begin{pmatrix} Y(t) \\ M(t) \end{pmatrix} = \begin{pmatrix} u(t,X(t)) \\ M(t) \end{pmatrix}$ *be a solution to the BSDE:*

$$Y(s) = Y(T) + \int_s^T f(\rho, X(\rho), Y(\rho), Z_M(\rho))\, d\rho + M(s) - M(T). \tag{6.35}$$

Then the function $u(t,x)$ defined by $u(t,x) = \mathbb{E}_{t,x}\left[Y(t)\right]$ is a viscosity solution to the following equation

$$\begin{cases} \dfrac{\partial u}{\partial s}(s,x) + L(s)u\left(s,x\right) + f\left(s,x,u(s,x),\nabla_u^L\left(s,x\right)\right) = 0; \\ u(T,x) = \varphi\left(T,x\right), \quad x \in E, \end{cases} \tag{6.36}$$

provided that the function $u(t,x)$ is continuous.

Notice that the equation in (6.36) is the same as the one in (5.84): see Remark 6.1.

Proof. Let the function $\varphi(s,y)$ be "smooth" and suppose that (t,x) is point in $[0,T) \times E$ where the function $u - \varphi$ vanishes and attains a local maximum. This means that there exists a subset of the form $[t, t+\varepsilon] \times U$, where U is an open neighborhood of x such that

$$\sup_{(s,y)\in[t,t+\varepsilon]\times U} \left(u(s,y) - \varphi(s,y)\right) = u(t,x) - \varphi(t,x) = 0.$$

We have to show that

$$\frac{\partial}{\partial t}\varphi\left(t,x\right) + L(t)\varphi\left(t,\cdot\right)\left(x\right) + f\left(t,x,u\left(t,x\right),\nabla_\varphi^L(t,x)\right) \geqslant 0, \tag{6.37}$$

where in (6.32) we have chosen

$$F\left(\dot{\varphi}, L\varphi, u, \nabla_\varphi^L\right)(t,x) = \dot{\varphi}(t,x) + L(t)\varphi(t,\cdot)(x) + f\left(t,x,u(t,x),\nabla_\varphi^L(t,x)\right). \tag{6.38}$$

Assume to arrive at a contradiction that the expression in (6.37) is strictly less than zero:

$$\frac{\partial}{\partial t}\varphi\left(t,x\right) + L(t)\varphi\left(t,\cdot\right)\left(x\right) + f\left(t,x,u\left(t,x\right),\nabla_\varphi^L(t,x)\right) < 0. \tag{6.39}$$

Upon shrinking $\varepsilon > 0$ and the open subset U we may and do assume that for all $(s,y) \in [t, t+\varepsilon] \times U$ the inequality

$$\frac{\partial}{\partial s}\varphi\left(s,y\right) + L(s)\varphi\left(s,\cdot\right)\left(y\right) + f\left(s,y,u\left(s,y\right),\nabla_\varphi^L(s,y)\right) < 0 \tag{6.40}$$

holds. Define the stopping τ by $\tau = \inf\left\{s \geqslant t : X(s) \notin U\right\} \wedge (t+\varepsilon)$. From (5.84) we have:

$$\begin{aligned} u(t,X(t)) = u\left(\tau,X(\tau)\right) + \int_t^\tau f\left(\rho,X(\rho),u\left(\rho,X(\rho)\right),Z_M(\rho)\right)d\rho \\ + M(t) - M(\tau). \end{aligned} \tag{6.41}$$

Let $M_\varphi(s)$ be the martingale associated to the function φ as in Proposition 5.3. Then

$$\varphi\left(t,X\left(t\right)\right) - \varphi\left(\tau,X\left(\tau\right)\right)$$

$$-\int_t^\tau \left(\frac{\partial}{\partial s}\varphi\left(s, X(s)\right) + L(s)\varphi\left(s,\cdot\right)\left(X(s)\right)\right) ds + M_\varphi\left(t\right) - M_\varphi\left(\tau\right).$$

From the definition of the stopping time τ it follows that $u\left(\tau, X\left(\tau\right)\right) \leqslant \varphi\left(\tau, X\left(\tau\right)\right)$. An application of Theorem 6.2 with

$$V(s) = \dot{\varphi}(s, X(s)) + L(s)\varphi(s,\cdot)(X(s)),$$

with $Y(s) = u\left(s, X(s)\right)$ and $Y'(s) = \varphi\left(s, X(s)\right)$ then shows $u\left(t, X(t)\right) < \varphi\left(t, X(t)\right)$ $\mathbb{P}_{t,x}$-almost surely. Since

$$u(t,x) = \mathbb{E}_{t,x}\left[u\left(t, X(t)\right)\right] \quad \text{and also} \quad \varphi(t,x) = \mathbb{E}_{t,x}\left[\varphi\left(t, X(t)\right)\right]$$

this leads to a contradiction. This means that our assumption (6.39) is false, and hence the function $u(t,x)$ is a viscosity sub-solution to equation (5.84) which is the same as (6.36). In the same manner one shows that $u(t,x)$ is also a viscosity super-solution to (5.84).

Altogether this completes the proof of Theorem 6.3. □

The following proposition says that solutions to the equation (5.84), which is the same as (6.36), are automatically continuous provided that the underlying Markov process is strong Feller: see the equalities in (6.42) below. For the notion of the strong Feller property see e.g. Definitions 2.5 and 2.16.

Proposition 6.2. *Let the pair (Y, M) be a solution to equation (6.35) in Theorem 6.3. Suppose that the pair (Y, M) belongs to the space*

$$\mathcal{S}^2_{\text{loc,unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}\right) \times \mathcal{M}^2_{\text{loc,unif}}\left(\Omega, \mathcal{F}^\tau_T, \mathbb{P}_{\tau,x}; \mathbb{R}\right)$$

(see Definitions 5.4 and 5.5). In addition, suppose that the Markov process in (6.34) is strong Feller. Then the function $(t,x) \mapsto u(t,x) := \mathbb{E}_{t,x}\left[Y(t)\right]$ is continuous on $[0,T] \times E$. This is a consequence of the strong Feller property and the following equalities:

$$\begin{aligned} u(t,x) &= \mathbb{E}_{t,x}[u(T, X(T))] + \int_t^T \mathbb{E}_{t,x}[f(s, X(s), Y(s), Z_M(s))]ds \qquad (6.42) \\ &= \mathbb{E}_{t,x}[u(T, X(T))] + \int_t^T \mathbb{E}_{t,x}\big[\mathbb{E}_{s,X(s)}[f(s, X(s), Y(s), Z_M(s))]\big]ds. \end{aligned}$$

Definition 6.2. Let $\{Y(t) : t \in [\tau, T]\}$ be a process in $L^1\left(\Omega, \mathcal{F}, \mathbb{P}_{\tau,x}\right)$ which is adapted relative to a filtration $\left(\mathcal{F}^\tau_t\right)_{\tau \leqslant t \leqslant T}$. Then the process $\{Y(t) : t \in [\tau, T]\}$ is said to be of class (DL) if the collection

$$\{Y(S) : \tau \leqslant S \leqslant T,\ S \text{ stopping time}\}$$

is uniformly integrable.

Notice that an increasing process in $L^1(\Omega,\mathcal{F},\mathbb{P})$ is automatically of class (DL), and that the same is true for a martingale. In addition, notice that that in our case the process $Y(t)$, $0\leqslant t\leqslant T$, which satisfies

$$Y(t)=Y(T)+\int_t^T f\left(s,X(s),Y(s),Z_M(s)\right)\,ds+M(t)-M(T), \quad (6.43)$$

where the pair (Y,M) belongs to $\mathcal{S}^2\times\mathcal{M}^2\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x}\right)$ is automatically of class (DL) in the space $L^1\left(\Omega,\mathcal{F}_T^\tau,\mathbb{P}_{\tau,x}\right)$. The reason being that a martingale is automatically of class (DL), and the same is true for a process of the form $t\mapsto\int_\tau^t f\left(s,X(s),Y(s),Z_M(s)\right)\,ds$, $\tau\leqslant t\leqslant T$.

In the proof we will employ a technique which is also used in the proof of the Doob-Meyer decomposition theorem. It states that a local right-continuous sub-martingale $\widetilde{Y}(t)$ of class (DL) can be written in the form

$$\widetilde{Y}(t)=M(t)+A(t)$$

where $t\mapsto M(t)$ is a right-continuous local martingale, and $t\mapsto A(t)$ is a predictable increasing process. For details see e.g. [Protter (2005)] theorems 12 and 13 in Chapter 3. For another account see [Karatzas and Shreve (1991b)] Theorem 4.10. Another proof can be found in [Rao (1969)]. In [van Neerven (2004)] Van Neerven gives a detailed account of the corresponding result in [Karatzas and Shreve (1991b)]. In addition, in the proof of the Doob-Meyer decomposition theorem Van Neerven uses the following version of the Dunford-Pettis theorem.

Theorem 6.4 (Dunford-Pettis). *If $(Y_n)_{n\in\mathbb{N}}$ is uniformly integrable sequence of random variables, then there exists an integrable random variable Y and a subsequence $(Y_{n_k})_{k\in\mathbb{N}}$ such that weak-$\lim_{k\to\infty}Y_{n_k}=Y$, i.e., for all bounded random variables ξ the following equality holds:*

$$\lim_{k\to\infty}\mathbb{E}\left[\xi Y_{n_k}\right]=\mathbb{E}\left[\xi Y\right].$$

For a proof of this version of the Dunford-Pettis theorem the reader is referred to [Kallenberg (2002)]. From general arguments in integration theory and functional analysis, it then follows that the variable Y can be written as the $\mathbb{P}$-almost sure limit of appropriately chosen convex combinations of the sequence $\{Y_{n_k}:k\geqslant\ell\}$, and this for all $\ell\in\mathbb{N}$. In other words there exists a sequence $\widetilde{Y}_\ell=\sum_{k=\ell}^{N_\ell}\alpha_{\ell,k}Y_{n_k}$, $\ell\in\mathbb{N}$, in $L^1(\Omega,\mathcal{F},\mathbb{P})$ with $\alpha_{\ell,k}\geqslant 0$, and for which $\sum_{k=\ell}^{N_\ell}\alpha_{\ell,k}=1$, L^1-$\lim_{\ell\to\infty}\widetilde{Y}_\ell=Y$, and $\lim_{\ell\to\infty}\widetilde{Y}_\ell=Y$ $\mathbb{P}$-almost surely.

Proof. Proof of Proposition 6.2 By the strong Feller property it suffices to show that the process $\rho \mapsto f\left(\rho, X(\rho), u\left(\rho, X(\rho)\right), Z_M(\rho)\right)$ only depends on the pair $(\rho, X(\rho))$. In other words we have to show that the functional $Z_M(\rho)$ only depends on $(\rho, X(\rho))$. We will verify this claim. Therefore we introduce the processes

$$
\begin{aligned}
Y_j(t) &= \mathbb{E}_{t,X(t)}\left[Y\left(\frac{\lceil 2^j t\rceil}{2^j}\wedge T\right)\right] \quad \text{and} \\
A_j(t) &= \sum_{0\leqslant k<2^j t} \mathbb{E}_{k2^{-j},X(k2^{-j})}\left[Y\left(\frac{k+1}{2^j}\wedge T\right)-Y\left(\frac{k}{2^j}\wedge T\right)\right], \qquad (6.44)
\end{aligned}
$$

$j \in \mathbb{N}$, $t \in [0,T]$. Fix $0 \leqslant t_1 < t_2 \leqslant T$. From (6.44) we see that the increment $A_j(t_2) - A_j(t_1)$ is measurable relative to the σ-field generated by $X\left(k2^{-j}\right)$, $t_1 \leqslant k2^{-j} < t_2$, $k \in \mathbb{N}$. Next, let $(\tau, x) \in [0,T) \times E$. Since $Y(s) = u(s, X(s))$, $0 \leqslant s \leqslant T$, we are eligible to apply the Markov property to infer that $\mathbb{P}_{\tau,x}$-almost surely

$$
\begin{aligned}
Y_j(t) &= \mathbb{E}_{\tau,x}\left[Y\left(\frac{\lceil 2^j t\rceil}{2^j}\wedge T\right) \,\Big|\, \mathcal{F}_t^\tau\right] \quad \text{and} \\
A_j(t) &= A_j(\tau) + \sum_{2^j\tau\leqslant k<2^j t} \mathbb{E}_{\tau,x}\left[Y\left(\frac{k+1}{2^j}\wedge T\right)-Y\left(\frac{k}{2^j}\wedge T\right) \,\Big|\, \mathcal{F}_{k2^{-j}}^\tau\right].
\end{aligned}
\tag{6.45}
$$

Next we show that the process $t \mapsto Y_j(t) - A_j(t) + A_j(\tau)$ is a $\mathbb{P}_{\tau,x}$-martingale. Let $0 \leqslant t_1 < t_2 \leqslant T$, and notice that the variables $Y_j(t_1)$ and $A_j(t_1) - A_j(\tau)$ are $\mathcal{F}_{t_1}^\tau$-measurable. We employ (6.44) and (6.45) to obtain

$$
\begin{aligned}
&\mathbb{E}_{\tau,x}\left[Y_j(t_2) - A_j(t_2) + A_j(\tau) \,\big|\, \mathcal{F}_{t_1}^\tau\right] - Y_j(t_1) + A_j(t_1) - A_j(\tau) \\
&= \mathbb{E}_{\tau,x}\left[Y_j(t_2) - Y_j(t_1) - A_j(t_2) + A_j(t_1) \,\big|\, \mathcal{F}_{t_1}^\tau\right] \\
&= \mathbb{E}_{\tau,x}\Bigg[\mathbb{E}_{\tau,x}\left[Y\left(\frac{\lceil 2^j t_2\rceil}{2^j}\wedge T\right) \,\Big|\, \mathcal{F}_{t_2}^\tau\right] - \mathbb{E}_{\tau,x}\left[Y\left(\frac{\lceil 2^j t_1\rceil}{2^j}\wedge T\right) \,\Big|\, \mathcal{F}_{t_1}^\tau\right] \\
&\qquad - \sum_{2^j t_1\leqslant k<2^j t_2} \mathbb{E}_{\tau,x}\left[Y\left(\frac{k+1}{2^j}\wedge T\right)-Y\left(\frac{k}{2^j}\wedge T\right) \,\Big|\, \mathcal{F}_{k2^{-j}}^\tau\right] \,\Big|\, \mathcal{F}_{t_1}^\tau\Bigg]
\end{aligned}
$$

(tower property of conditional expectations)

$$
\begin{aligned}
&= \mathbb{E}_{\tau,x}\Bigg[Y\left(\frac{\lceil 2^j t_2\rceil}{2^j}\wedge T\right) - Y\left(\frac{\lceil 2^j t_1\rceil}{2^j}\wedge T\right) \\
&\qquad - \sum_{2^j t_1\leqslant k<2^j t_2} \left[Y\left(\frac{k+1}{2^j}\wedge T\right)-Y\left(\frac{k}{2^j}\wedge T\right)\right] \,\Big|\, \mathcal{F}_{t_1}^\tau\Bigg]
\end{aligned}
$$

$$= \mathbb{E}_{\tau,x}\left[0 \mid \mathcal{F}_{t_1}^{\tau}\right] = 0. \tag{6.46}$$

From (6.46) it follows that for every pair $(\tau,x) \in [0,T) \times E$ the processes $t \mapsto Y_j(t) - A_j(t) + A_j(\tau)$, $j \in \mathbb{N}$, are $\mathbb{P}_{\tau,x}$-martingales relative to the filtration $(\mathcal{F}_t^{\tau})_{t\in[\tau,T]}$. Put $M_j(t) = Y_j(t) - A_j(t)$. Then the process $t \mapsto M_j(t) - M_j(\tau)$, $t \in [\tau,T]$, is a $\mathbb{P}_{\tau,x}$-martingale, and

$$Y_j(t) - Y_j(\tau) = A_j(t) - A_j(\tau) + M_j(t) - M_j(\tau). \tag{6.47}$$

In (6.47) we let j tend to ∞, and if necessary, we pass to a subsequence, to obtain

$$Y(t) - Y(\tau) = \int_{\tau}^{t} f(s, X(s), Y(s), Z_M(s))\, ds + M(t) - M(\tau), \tag{6.48}$$

where in $L^1(\Omega, \mathcal{F}_t^{\tau}, \mathbb{P}_{\tau,x})$ and $\mathbb{P}_{\tau,x}$-almost surely the following equalities hold

$$\int_{\tau}^{t} f(s, X(s), Y(s), Z_M(s))\, ds = \lim_{n\to\infty} \sum_{k=n}^{N_n} \alpha_{n,k}\left(A_{j_k}(t) - A_{j_k}(\tau)\right), \quad \text{and}$$

$$M(t) - M(\tau) = \lim_{n\to\infty} \sum_{k=n}^{N_n} \alpha_{n,k}\left(M_{j_k}(t) - M_{j_k}(\tau)\right), \tag{6.49}$$

for certain real numbers $\alpha_{n,k} \geqslant 0$ which satisfy $\sum_{k=n}^{N_n} \alpha_{n,k} = 1$. For all this see the comments following Theorem 6.4. It follows that $\mathbb{P}_{\tau,x}$-almost surely, the variables

$$\int_{t_1}^{t_2} f(s, X(s), Y(s), Z_M(s))\, ds, \quad \tau \leqslant t_1 < t_2 \leqslant T, \tag{6.50}$$

are $\mathcal{F}_{t_2}^{t_1}$-measurable. Consequently, for almost every $s \in [\tau,T]$, the variable $f(s, X(s), Y(s), Z_M(s))$ is almost surely $\mathbb{P}_{\tau,x}$-measurable relative to $\sigma(s, X(s))$. Since the paths of the process X are continuous from the right it follows that for almost all $s \in [0,T]$ the variable $f(s, X(s), Y(s), Z_M(s))$ is $\mathcal{F}_{s+}^{t}$-measurable for all $0 \leqslant t < s$. If $0 \leqslant t < s \leqslant T$ by the strong Markov property relative to the filtration $\left(\mathcal{F}_{s+}^{t}\right)_{s\in[t,T]}$ (see Theorem 2.9) we then have

$$\begin{aligned}
&\mathbb{E}_{t,x}\left[\mathbb{E}_{s,X(s)}\left[f(s, X(s), Y(s), Z_M(s))\right]\right] \\
&= \mathbb{E}_{t,x}\left[\mathbb{E}_{t,x}\left[f(s, X(s), Y(s), Z_M(s)) \mid \mathcal{F}_{s+}^{t}\right]\right] \\
&= \mathbb{E}_{t,x}\left[f(s, X(s), Y(s), Z_M(s))\right],
\end{aligned} \tag{6.51}$$

and

$$\mathbb{E}_{s,X(s)}\left[f(s, X(s), Y(s), Z_M(s))\right]$$

$$
\begin{aligned}
&= \mathbb{E}_{t,x}\left[f\left(s, X(s), Y(s), Z_M(s)\right) \mid \mathcal{F}_s^t \right] \\
&= \mathbb{E}_{t,x}\left[f\left(s, X(s), Y(s), Z_M(s)\right) \mid \mathcal{F}_{s+}^t \right] \\
&= f\left(s, X(s), Y(s), Z_M(s)\right), \quad \mathbb{P}_{t,x}\text{-almost surely.}
\end{aligned} \tag{6.52}
$$

From (6.35) and (6.52) we infer

$$
\begin{aligned}
Y(t) = Y(T) + \int_t^T \mathbb{E}_{s,X(s)}\left[f\left(s, X(s), Y(s), Z_M(s)\right)\right] ds \\
+ M(t) - M(T),
\end{aligned} \tag{6.53}
$$

and hence by (6.51) from (6.53) we get

$$
\begin{aligned}
u(t,x) &= \mathbb{E}_{t,x}\left[Y(t)\right] = \mathbb{E}_{t,x}\left[u\left(T, X(T)\right)\right] \\
&\quad + \int_t^T \mathbb{E}_{t,x}\left[\mathbb{E}_{s,X(s)}\left[f\left(s, X(s), Y(s), Z_M(s)\right)\right]\right] ds.
\end{aligned} \tag{6.54}
$$

As a consequence, the strong Feller property implies that the function

$$
(t,x) \mapsto \mathbb{E}_{t,x}\left[\mathbb{E}_{s,X(s)}\left[f\left(s, X(s), Y(s), Z_M(s)\right)\right]\right], \tag{6.55}
$$

$0 \leqslant t \leqslant s \leqslant T$, $x \in E$, is continuous. As a consequence, from (6.54) and (6.55) we infer that the function $(t,x) \mapsto u(t,x)$ is continuous.

This conclusion completes the proof of Proposition 6.2. □

6.3 Backward stochastic differential equations in finance

In [Crandall *et al.* (1984)] the authors M.G. Crandall, L.C. Evans, and P.L. Lions study properties of viscosity solutions of Hamilton-Jacobi equations. In [Pardoux (1998b)] E. Pardoux uses viscosity solutions in the study of backward stochastic differential equations and semi-linear parabolic equations. In [El Karoui *et al.* (1997)] and in [El Karoui and Quenez (1997)] the authors employ backward stochastic equations to study American option pricing. We like to give an introduction to this kind of stochastic differential equations and the corresponding parabolic partial differential equations. As a rule the operator L generates a d-dimensional diffusion. For instance, if $L = \frac{1}{2}\Delta$, then the corresponding diffusion is Brownian motion. To some extent a solution to a BSDE corresponding to a semilinear parabolic partial differential equation generalizes the (classical) Feynman-Kac formula. We also mention that Nelson [Nelson (1967)] was perhaps the first to consider backward stochastic differential equations. In the linear case Bismut [Bismut (1973, 1978)] also considered backward stochastic differential equations. Most of the material presented in this section is taken from [El Karoui and Quenez (1997)] and [El Karoui *et al.* (1997)]. See the papers in [El Karoui and Mazliak (1997)] as well. First we describe a model of assets and hedging strategies. There is a non-risky asset (the money market or bond) $S^0(t)$, and there are n risky assets $S^j(t)$, $1 \leqslant j \leqslant n$. The process $S^0(t)$ satisfies the differential equation

$$dS^0(t) = S^0(t)r(t)dt, \text{ where } r(t) \text{ is the short time interest rate.}$$

The other assets satisfy a linear stochastic differential equation (SDE) of the form

$$dS^j(t) = S^j(t)\left[b^j(t)dt + \sum_{k=1}^{n} \sigma_{jk}(t)dW^k(t)\right],$$

which is driven by a standard Wiener process $W(t) = \left(W^1(t), \dots, W^n(t)\right)^*$, defined on a filtered space $\left(\Omega, (\mathcal{F}_t)_{0\leqslant t\leqslant T}, \mathbb{P}\right)$. It is assumed that $(\mathcal{F}_t)_{0\leqslant t\leqslant T}$ is generated by the Wiener process. Generally speaking the coefficients $r(t)$, $b^j(t)$, $\sigma_{jk}(t)$ are supposed to be bounded predictable processes with values in $\mathbb{R}$. We also write $\sigma_j(t) = (\sigma_{jk}(t))_{k=1}^n$. The matrix $\left[\sigma_{jk}(t)\right]_{j,k=1}^n$ is called the volatility matrix. To ensure the absence of arbitrage opportunities in the market, it is assumed that there exists an n-dimensional bounded predictable vector process $\vartheta(t)$ such that

$$b(t) - r(t)\mathbf{1} = \sigma(t)\vartheta(t), \quad dt \otimes \mathbb{P}\text{-almost surely.}$$

The vector $\mathbf{1}$ is the column vector, which is constant 1, and $\vartheta(t)$ is called the risk premium vector. It is assumed that $\sigma(t)$ has full rank. Consider a small investor, whose actions do not affect the market prices, and who can decide at time $t \in [0, T]$ what amount of the wealth $V(t)$ to invest in the j-th stock, $1 \leqslant j \leqslant n$. Of course his decisions are only based on the current information $\mathcal{F}_t$; i.e. $\pi(t) = \left(\pi^1(t), \dots, \pi^n(t)\right)^*$, and $\pi^0(t) = V(t) - \sum_{j=1}^n \pi^j(t)$ are predictable processes. The process $\pi(t)$ is called the portfolio process. The existence of such a *risk process* $\vartheta(t)$ guarantees that the model is *arbitrage free*. Let us make this precise by beginning with some definitions.

Definition 6.3.

(a) A progressively measurable $\mathbb{R}^n$-valued process

$$\pi = \left\{(\pi_1(t), \dots, \pi_n(t))^* : 0 \leqslant t \leqslant T\right\}$$

with the property

$$\int_0^t \left|\pi^*(t)\sigma(t)\right|^2 dt + \int_0^T \left|\pi^*(t)\left(b(t) - r(t)\mathbf{1}\right)\right| dt < \infty, \quad \mathbb{P}\text{-almost surely}$$

is called a *portfolio process*.

(b) Put $\gamma(t) = \exp\left(-\int_0^t r(\tau)d\tau\right)$, and define for a given portfolio $\pi(t)$ the process $M^\pi(t)$ by

$$M^\pi(t) = \int_0^t \gamma(s)\pi^*(s)\left[\sigma(s)dW(s) + (b(s) - r(s)\mathbf{1})\ ds\right], \quad 0 \leqslant t \leqslant T.$$

The process $M^\pi(t)$ is called the *discounted gains process*. A portfolio $\pi(t)$ is called *tame* if there exists a real constant q^π such that $\mathbb{P}$-almost surely $M^\pi(t) \geqslant q^\pi$, $0 \leqslant t \leqslant T$.

(c) A tame portfolio $\pi(t)$ that satisfies

$$\mathbb{P}\left[M^\pi(T) \geqslant 0\right] = 1, \quad \text{and} \quad \mathbb{P}\left[M^\pi(T) > 0\right] > 0,$$

is called an *arbitrage opportunity* (or *"free lunch"*). A market $\mathcal{M}$ is called *arbitrage free* if no such portfolios exist in it.

The following theorem shows the relevance of the existence of a risk process $\vartheta(t)$.

Theorem 6.5.

(i) If the market $\mathcal{M}$ is arbitrage-free, then there exists a progressively measurable process $\vartheta : [0,T] \times \Omega \to \mathbb{R}^n$, called the market price or price of risk (or price of relative risk) process, such that

$$b(t) - r(t)\mathbf{1} = \sigma(t)\vartheta(t), \quad 0 \leqslant t \leqslant T,\ \mathbb{P}\text{-almost surely.}$$

(ii) Conversely, if such a price of risk process exists and satisfies, in addition to the above requirements,

$$\int_0^T |\vartheta(t)|^2\, dt < \infty, \quad \mathbb{P}\text{-almost surely}, \quad \text{and} \tag{6.56}$$

$$\mathbb{E}\left[\exp\left(-\int_0^T \vartheta^*(t)dW(t) - \frac{1}{2}\int_0^T |\vartheta(t)|^2\, dt\right)\right] = 1, \tag{6.57}$$

then $\mathcal{M}$ is arbitrage free.

From Novikov's condition (see Proposition 3.5.12 in [Karatzas and Shreve (1991a)]), it follows that conditions (6.56) and (6.57) are satisfied if

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T |\vartheta(t)|^2\, dt\right)\right] < \infty;$$

in particular this is the case if $|\vartheta(t)|$ is uniformly bounded in $(t,\omega) \in [0,T] \times \Omega$. For Novikov's condition see Theorem 1.6 and its Corollary 1.3 in Chapter 1. It is noticed that under the condition (6.57) the process

$W(t) + \int_0^t \vartheta^*(s)\mathbf{1}ds$ is a Brownian motion with respect to the martingale measure $\mathbb{Q}$ which has Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \exp\left(-\int_0^T \vartheta^*(t)dW(t) - \frac{1}{2}\int_0^T |\vartheta(t)|^2 \,dt\right).$$

For more details the reader is referred to [Karatzas (1997)], and to [Karatzas and Shreve (1998)]. Another valuable source of information is Kleinert [Kleinert (2003)], Chapter 20. Following Harrison and Pliska [Harrison and Pliska (1981)] a strategy $(V(t), \pi(t))$ is called self-financing if the wealth process $V(t) = \sum_{j=0}^n \pi^j(t)$ obeys the equality

$$V(t) = V(0) + \int_0^t \sum_{j=1}^n \pi^j(s)\frac{dS^j(s)}{S^j(s)},$$

or, equivalently, if it satisfies the linear stochastic differential equation

$$\begin{aligned} dV(t) &= r(t)V(t)dt + \pi^*(t)\left(b(t) - r(t)\mathbf{1}\right)dt + \pi^*(t)\sigma(t)dW(t) \\ &= r(t)V(t)dt + \pi^*(t)\sigma(t)\left[dW(t) + \vartheta(t)dt\right]. \end{aligned} \tag{6.58}$$

Often the left side of (6.58) contains a term $dK(t)$, where the process $K(t)$ is, adapted, increasing and right-continuous, with $K(0) = 0$, $K(T) < \infty$, $\mathbb{P}$-almost surely. The process is called the *cumulative consumption process*. A pair $(V(t), \pi(t))$ satisfying (6.58) is called a self-financing trading strategy. There exists a one to one correspondence between the pairs $(x, \pi(t))$ and pairs $(V(t), \pi(t))$ with $V(0) = x$ and which satisfy (6.58).

Definition 6.4. A *hedging strategy against a contingent claim* $\xi \in \mathbb{L}^2$ is a self-financing strategy $(V(t), \pi(t))$ such that $V(T) = \xi$ with

$$\mathbb{E}\left[\int_0^T |\sigma^*(t)\pi(t)|^2 \,dt\right] < \infty.$$

Theorem 6.6. *An attainable square integrable contingent claim ξ is replicated by a unique hedging strategy $(V(t), \pi(t))$; i.e. there exists a unique solution $(V(t), \pi(t))$ to equation (6.58) such that $V(T) = \xi$.*

The following theorem elaborates on this statement.

Theorem 6.7. *Any square integrable contingent claim is attainable; i.e. the market is complete. In other words, for every square integrable random variable ξ there exists a unique pair $(X(t), \pi(t))$ such that $\mathbb{E}\left[\int_0^T |\sigma^*(t)\pi(t)|^2 \,dt\right] < \infty$, and such that*

$$dX(t) = r(t)X(t)dt + \pi^*(t)\sigma(t)\left(\vartheta(t)\,dt + dW(t)\right), \quad X(T) = \xi. \tag{6.59}$$

The process $X(t)$ represents the price of the claim at time t, given by the closed formula $X(t) = \mathbb{E}\left[H^t(T)\xi \mid \mathcal{F}_t\right]$, where $H^t(s)$, $t \leqslant s \leqslant T$, is the deflator process, starting at time t such that

$$dH^t(s) = -H^t(s)\left[r(s)ds + \vartheta^*(s)dW(s)\right]; \quad H^t(t) = 1. \tag{6.60}$$

Remark 6.2. Suppose that the process $t \mapsto X(t)$ satisfies equation (6.59). By Itô's calculus it follows that the process $\left\{H^0(t)X(t) : 0 \leqslant t \leqslant T\right\}$ is a stochastic integral such that

$$d\left(H^0(\cdot)X(\cdot)\right)(t) = H^0(t)\left\{\pi^*(t) - X(t)\vartheta^*(t)\right\}dW(t).$$

Classical results about solutions to the linear SDE (6.60) with bounded coefficients yield the (uniform) boundedness of the martingale $H^0(t)$ in $\mathbb{L}^2$; moreover the process $\left(H^0(t)X(t) : 0 \leqslant t \leqslant T\right)$ is uniformly integrable. It follows that

$$H^0(t)X(t) = \mathbb{E}\left[H^0(T)\xi \mid \mathcal{F}_t\right], \quad \text{or, equivalently,} \quad X(t) = \mathbb{E}\left[H^t(T)\xi \mid \mathcal{F}_t\right].$$

The closed form of the deflator process,

$$H^t(s) = \exp\left(-\left\{\int_t^s r(\tau)d\tau + \int_t^s \vartheta^*(\tau)dW(\tau) + \frac{1}{2}\int_t^s \left|\vartheta(\tau)\right|^2 d\tau\right\}\right),$$

leads to a more classical formulation of the contingent claim:

$$\begin{aligned} X(t) & \\ &= \mathbb{E}\left[\exp\left(-\left\{\int_t^T r(\tau)d\tau + \int_t^T \vartheta^*(\tau)dW(\tau) + \frac{1}{2}\int_t^T \left|\vartheta(\tau)\right|^2 d\tau\right\}\right)\xi \mid \mathcal{F}_t\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r(\tau)d\tau\right)\xi \mid \mathcal{F}_t\right], \end{aligned} \tag{6.61}$$

where $\exp\left(-\int_t^T r(\tau)d\tau\right)$ is the discounted factor over the time interval $[0,T]$ and the measure $\mathbb{Q}$ is the risk-adjusted probability measure defined by the Radon-Nikodym derivative with respect to $\mathbb{P}$:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\left\{\int_0^T \vartheta^*(\tau)dW(\tau) + \frac{1}{2}\int_0^T \left|\vartheta(\tau)\right|^2 d\tau\right\}\right).$$

Proof. [Proof of Theorem 6.7.] First we prove *uniqueness*. Let the pair $(X(t), \pi(t))$, where $X(t)$ is adapted and $\pi(t)$ is predictable, satisfy equation (6.59). Let the process $H^0(t)$ satisfy the differential equation as exhibited in (6.60). Then

$$d\left(H^0(\cdot)X(\cdot)\right)(t) = H^0(t)\left\{\pi^*(t) - X(t)\vartheta^*(t)\right\}dW(t).$$

As explained in the previous Remark 6.2, it follows that

$$X(t) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r(\tau)d\tau\right)\xi \mid \mathcal{F}_t\right].$$

This shows that that the process $X(t)$ is uniquely determined. But once $X(t)$ is uniquely determined, then the same is true for the process $\pi(t)$. Corollary 1.7 in Chapter 10 implies that the process $W(t) + \int_0^t \vartheta^*(s)ds$ is Brownian motion with respect to the measure $\mathbb{Q}$. Moreover, the process $X(t) - \int_0^t r(\tau) X(\tau)d\tau = \int_0^t \pi^*(s)\sigma(s)\left(dW(s) + \vartheta(s)ds\right)$, where the process $t \mapsto W(t) + \int_0^t \vartheta(s)ds$ is a Brownian motion with respect to the measure $\mathbb{Q}$. Let $(X_1(t), \pi_1(t))$ and $(X_2(t), \pi_2(t))$ be two solutions to the equation in (6.59). Then

$$\begin{aligned}
&X_1(T) - X_1(t) - \int_t^T r(\tau)X_1(\tau)d\tau = X_2(T) - X_2(t) - \int_t^T r(\tau)X_2(\tau)d\tau \\
&= \xi - \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r(\tau)d\tau\right)\xi \mid \mathcal{F}_t\right] \\
&\quad - \int_t^T r(\tau)\mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_\tau^T r(s)ds\right)\xi \mid \mathcal{F}_t\right]d\tau \\
&= \xi - \int_t^T \pi_1^*(s)\sigma(s)\left(dW(s) + \vartheta(s)ds\right) \\
&= \xi - \int_t^T \pi_2^*(s)\sigma(s)\left(dW(s) + \vartheta(s)ds\right).
\end{aligned}$$

Hence,

$$\int_t^T \left(\pi_1^*(s)\sigma(s) - \pi_2^*(s)\right)\left(dW(s) + \vartheta(s)ds\right) = 0, \quad 0 \leqslant t \leqslant T. \tag{6.62}$$

Thus $\mathbb{E}_{\mathbb{Q}}\left[\int_t^T \left|\pi_1(\tau) - \pi_2(\tau)\right|^2 d\tau\right] = 0$, and consequently the equality $\pi_1(t) = \pi_2(t)$ holds $\lambda \times \mathbb{Q}$-almost surely. Here we wrote λ for the Lebesgue measure on $\mathbb{R}$. Since the $\mathbb{Q}$-negligible sets coincide with $\mathbb{P}$-negligible sets, we get $\pi_1(t) = \pi_2(t)$ for $\lambda \times \mathbb{P}$-almost all $(t, \omega) \in [0, T] \times \Omega$.

Next we prove the existence. Define the process $Y(t)$, $0 \leqslant t \leqslant T$, by

$$Y(t) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_0^T r(\tau)\,d\tau\right)\xi \mid \mathcal{F}_t\right].$$

The process $t \mapsto Y(t)$ is a $\mathbb{P}_{\mathbb{Q}}$-martingale, and since the processes $t \mapsto W(t) + \int_0^t \vartheta(s)ds$ is a $\mathbb{P}_{\mathbb{Q}}$-Brownian motion, there exists by a martingale representation theorem a predictable process $\widetilde{\pi}(t)$ such that

$$Y(T) - Y(t) = \int_t^T \widetilde{\pi}^*(s)\left(dW(s) + \vartheta(s)ds\right). \tag{6.63}$$

From (6.63) we easily infer that

$$dY(t) = \widetilde{\pi}^*(t) \left(dW(t) + \vartheta(t)dt\right).$$

Next, put

$$\pi^*(t) = \exp\left(\int_0^t r(\tau)d\tau\right) \widetilde{\pi}^*(t)\sigma(t)^{-1} \quad \text{and} \quad X(t) = \exp\left(\int_0^t r(\tau)d\tau\right) Y(t).$$

Then we have $X(T) = \xi$, and

$$dY(t) = \exp\left(-\int_0^t r(\tau)d\tau\right) \pi^*(t)\sigma(t) \left(dW(t) + \vartheta(t)dt\right),$$

and hence

$$\begin{aligned} dX(t) &= r(t)X(t)dt + \exp\left(\int_0^t r(\tau)d\tau\right) dY(t) \\ &= r(t)X(t)dt \\ &\quad + \exp\left(\int_0^t r(\tau)d\tau\right) \exp\left(-\int_0^t r(\tau)d\tau\right) \pi^*(t)\sigma(t) \left(dW(t) + \vartheta(t)dt\right) \\ &= r(t)X(t)dt + \pi^*(t)\sigma(t) \left(dW(t) + \vartheta(t)dt\right). \end{aligned} \tag{6.64}$$

This proves the existence of a solution to equation (6.59).

Altogether this completes the proof of Theorem 6.7. □

For more information on the martingale representation theorem in relation to hedging strategies in financial mathematics see e.g. [Shreve (2004)]. For a proof of the martingale representation theorem see e.g. [Protter (2005)].

6.4 Some related remarks

In this section we will explain the relevance of backward stochastic differential equations (BSDEs). We will also mention that Bismut was the first to discuss BSDEs [Bismut (1978)], and [Bismut (1981b)]. Of course BSDEs were popularized by Pardoux and coworkers; see e.g. [Pardoux and Peng (1990); Pardoux and Zhang (1998); Pardoux (1998a, 1999)]. The first paper in which a solution to a BSDE is linked to a non-linear Feynman-Kac formula is [Peng (1991)]. The BSDEs discussed in Chapters 5 and 6 have as input a Markov process which could be a solution to a Stochastic Differential Equation, and therefore these BSDEs could be considered as generalizations of forward-backward stochastic differential equations. In the more classical context such equations are treated in the book by Ma-Yong [Ma and Yong (1999)]. Other relevant work is done by Lejay [Lejay

(2002, 2004)]. There is also a link with control theory: see e.g. Yong and Zhou [Yong and Zhou (1999)]. For the close connection between BSDEs and hedging strategies in financial mathematics the reader is referred to e.g. [El Karoui *et al.* (1997)], and [El Karoui and Quenez (1997)]. Another paper related to obstacles, and therefore also to hedging strategies, is the reference [Karoui *et al.* (1997)]. For some more explanation the reader is also referred to §6 in [Van Casteren (2002)]. An important area of mathematics and its applications where backward problems play a central role is control theory: see e.g. [Soner (1997)]. In the finite-dimensional setting the paper [Crandall *et al.* (1992a)] is very relevant for understanding the notion of viscosity solutions. Classical results on viscosity solutions can also be found in [Jensen (1989)]. Not necessary continuous viscosity solutions also play a central role in applied fields like dislocation theory, see e.g. [Barles *et al.* (2008)] and [Barles (1993)]. As remarked in Chapter 5 for a recent paper in which the martingale approach is used to treat forward-backward stochastic differential equations we refer the reader to [Ma *et al.* (2008)].

Chapter 7

The Hamilton-Jacobi-Bellman equation and the stochastic Noether theorem

In this chapter we prove that the Lagrangian action, which may be phrased in terms of a non-linear Feynman-Kac formula, coincides under rather generous hypotheses with the unique viscosity solution to the Hamilton-Jacobi-Bellman equation: see Theorem 7.1. The method of proof is based on martingale theory and Jensen inequality. A version of the stochastic Noether theorem is proved, as well as its complex companion: see Theorems 7.5 and 7.6 further on this chapter. The proofs of these Noether theorems are cumbersome and require a dextrous calculation. Whereas in the other chapters of the book we use the notation $L(s)$, $0 \leqslant s \leqslant T$, or $L(s)$, $0 \leqslant s \leqslant T$, to indicate the generator of a diffusion or a Markov process, in the present chapter we will use the family of operators $-K_0(s)$, $0 \leqslant s \leqslant T$, to indicate such a family. In physical terms such an operator family $K_0(s)$, $0 \leqslant s \leqslant T$, is in notation closer to a Hamiltonian than the operator family $L(s)$, $s \in [0, T]$.

7.1 Introduction

We start this chapter by pointing out that Zambrini and coworkers [Albeverio *et al.* (2006a,b); Chung and Zambrini (2001); Thieullen and Zambrini (1997a,c,b,d); Zambrini (1998b,a)] have kind of a transition scheme to go from classical stochastic calculus (with non-reversible processes) to physical real time (reversible) quantum mechanics and vice versa. An important tool in this connection is the so-called Noether theorem. In fact, in Zambrini's words, reference [Zambrini (1998a)] contains the first concrete application of this theorem. In [Zambrini (1998a)] the author formulates a theorem like Theorem 7.1 below, he also uses so-called "Bernstein diffusions" (see e.g. [Cruzeiro and Zambrini (1991)]) for the "Euclidean Born interpretation" of quantum mechanics. The Bernstein diffusions are related to solutions of $\left(\frac{\partial}{\partial t} - \left(K_0 \dot{+} V\right)\right) \eta(t,x) = 0$, and of $\left(\frac{\partial}{\partial t} + K_0 \dot{+} V\right) \eta^*(t,x) = 0$. In the present paper we prove a version of the stochastic Noether theorem in terms of the carré du champ operator and ideas from stochastic control: see Theorem 7.5, which should be compared with Theorem 2.4 in [Zambrini (1998a)]. The operator K_0 generates a *diffusion* in the following sense: for every C^∞-function $\Phi : \mathbb{R}^\nu \to \mathbb{R}$, with $\Phi(0,\dots,0) = 0$, the following identity is valid:

$$K_0\left(\Phi(f_1,\dots,f_n)\right) = \sum_{j=1}^{n} \frac{\partial \Phi}{\partial x_j}(f_1,\dots,f_n)\, K_0 f_j - \frac{1}{2}\sum_{j,k=1}^{n} \frac{\partial^2 \Phi}{\partial x_j \partial x_k}(f_1,\dots,f_n)\, \Gamma_1(f_j,f_k) \tag{7.1}$$

for all functions $f_1,\dots,f_n$ in a rich enough algebra of functions $\mathcal{A}$, contained in the domain of the generator K_0, as described below. The condition $\Phi(0,\dots,0) = 0$ will be omitted in case the function $\mathbf{1}$ belongs to the domain of the operator K_0. Throughout this chapter we will assume that the operator $K_0 = K_0(t)$, $t \in [0,T]$, is a space-time operator. Compare all this with Definition 5.1 and the comments following it.

7.1.0.1 *Hypotheses on the generator and the algebra* $\mathcal{A}$

We will assume that the constant functions belong to $D(K_0)$, and that $K_0\mathbf{1} = 0$. The algebra $\mathcal{A}$ has to be "large" enough. To be specific, we assume that the operator K_0 is a space-time operator with domain in $C_b([0,T]\times E)$, and that $\mathcal{A}$ is a core for the operator K_0, which means that the $\mathcal{T}_\beta$-closure of its graph $\{(\varphi, K_0\varphi) : \varphi \in \mathcal{A}\}$ is again the graph of $\mathcal{T}_\beta$-closed operator, which we keep denoting by K_0. In addition, it is assumed that $\mathcal{A}$ is stable under composition with C^∞-functions of several variables, that vanish at the origin. Moreover, in order to obtain some nice results a rather technical condition is required: whenever $(f_n : n \in \mathbb{N})$ is a sequence in $\mathcal{A}$ that converges to f with respect to the $\mathcal{T}_\beta$-topology in $C_b([0,T]\times E) \times C_b([0,T]\times E)$ and whenever $\Phi : \mathbb{R} \to \mathbb{R}$ is a C^∞-function with bounded derivatives of all orders (including the order 0), then one may extract a subsequence $(\Phi(f_{n_k}) : k \in \mathbb{N})$ that converges to $\Phi(f)$ in $C_b([0,T]\times E)$, whereas the sequence $(K_0\Phi(f_{n_k}) : k \in \mathbb{N})$ converges in $C_b([0,T]\times E)$. In fact it would be no restriction to assume that $\Phi(0) = 0$, because we assume that the constant functions belong $D(K_0)$ and $K_0\mathbf{1} = 0$. So we can always replace Φ by $\Phi - \Phi(0)$. Notice that all functions of the form $e^{-\psi}f$, ψ, $f \in \mathcal{A}$, belong to $\mathcal{A}$. Also notice that the required properties of $\mathcal{A}$ depend on the generator K_0. In fact we will assume that the algebra $\mathcal{A}$ is also large enough for all operators of the form $f \mapsto e^{\psi}K_0\left(e^{-\psi}f\right)$, where ψ belongs to $\mathcal{A}$. In addition, we assume that $\mathbf{1} \in D(K_0)$, and that $K_0\mathbf{1} = 0$. The operator K_0 is supposed to be $\mathcal{T}_\beta$-closed when viewed as an operator acting on functions in $C_b([0,T] \times E)$.

Remark 7.1. Let ds be the Lebesgue measure on $[0,T]$. If there exists a reference measure m on the Borel field $\mathcal{E}$ of E, and if we want to work in the L^p-spaces $L^p([0,T] \times E, ds \times m)$, $1 \leqslant p < \infty$, then it is assumed that K_0 has dense domain in $L^p([0,T] \times E, ds \times m)$, for each $1 \leqslant p < \infty$. In addition, it is assumed that $\mathcal{A}$ is a subalgebra of $D(K_0)$ which possesses the following properties (cf. [Bakry (1994)]). Its is dense in $L^p([0,T] \times E, ds \times m)$ for all $1 \leqslant p < \infty$ and it is a core for K_0, provided K_0 is considered as a densely defined operator in such a space. The latter means that the algebra $\mathcal{A}$ consists of functions in $D(K_0)$ viewed as an operator in $L^p([0,T] \times E, ds \times m)$.

The same is true for the space $C_b([0,T] \times E) = C_b([0,T] \times E, \mathbb{C})$, but then relative to the strict topology. In addition, it is assumed that $\mathcal{A}$ is stable under composition with C^∞-functions of several variables, that do not necessarily vanish at the origin. Moreover, as indicated above in order to obtain some nice results a more technical condition is required. Whenever $(f_n : n \in \mathbb{N})$ is a sequence in $\mathcal{A}$ that converges to f with respect to the graph norm of K_0 (in $L^2([0,T] \times E, ds \times m)$) and whenever $\Phi : \mathbb{R} \to \mathbb{R}$ is a C^∞-function, vanishing at 0, with bounded derivatives of all orders (including the order 0), then there exists a subsequence $(\Phi(f_{n_k}) : k \in \mathbb{N})$ that converges to $\Phi(f)$ in $C_b([0,T] \times E)$, whereas the sequence $(K_0\Phi(f_{n_k}) : k \in \mathbb{N})$ converges in $C_b([0,T] \times E)$ and also in $L^1(E,m)$ to $K_0\Phi(f)$.

7.1.0.2 *Some additional comments*

From (7.1) we see that

$$-e^{\psi}K_0\left(e^{-\psi}f\right) = \left(K_0\psi + \frac{1}{2}\Gamma_1(\psi,\psi)\right)f - K_0 f - \Gamma_1(\psi,f), \text{ and} \tag{7.2}$$

$$K_0(\varphi\psi) = (K_0\varphi)\psi + \varphi(K_0\psi) - \Gamma_1(\varphi,\psi) \tag{7.3}$$

for φ, $\psi \in \mathcal{A}$, and $f \in D(K_0)$. For the notion of the squared gradient operator (carré du champ opérateur) see equality (7.7). The operator K_0 acts on the space and time variable, and the squared gradient operator Γ_1 only acts on the space variable; its action depends on the time-coordinate. The symbol D_1 stands for the operator $D_1 = \frac{\partial}{\partial t}$. Fix $T > t_0 \geqslant 0$. In the remainder of the present chapter we work in a continuous function spaces like $C_b((t_0, T] \times E)$ and sometimes in $C((t_0, T] \times E)$. If we write $D(D_1 - K_0)$ for the domain of the operator $D_1 - K_0$, then the corresponding space should be specified. In fact the space $C_b((t_0, T] \times E)$ is endowed with the strict topology $\mathcal{T}_\beta$, and also with that of uniform convergence. The operator $D_1 - K_0$ is considered as the generator of the semigroup $\{S(\rho) : \rho \geqslant 0\}$ defined by

$$\begin{aligned} S(\rho)f(\tau, x) &= P(\tau, (\rho + s) \wedge T) f((\rho + s) \wedge T, \cdot)(x) \\ &= \mathbb{E}_{\tau,x}\left[f((\rho + s) \wedge T, X((\rho + s) \wedge T))\right]. \end{aligned} \tag{7.4}$$

Here $\{P(s, t) : 0 \leqslant s \leqslant t \leqslant T\}$ is the Feller propagator generated by the operator $-K_0$: see Definition 2.8 and also Definition 2.7. The formula in (7.4) is the same as (3.90) in Chapter 3. Then it follows that for $t + \rho \leqslant T$ we have

$$S(\rho)P(\tau + \rho, t + \rho) f(\tau, x) = P(\tau, t + \rho) f(t + \rho, \cdot)(x) = S(t + \rho)f(\tau, x) \tag{7.5}$$

where $f \in C_b(E)$. Notice that in (7.5) the operator $S(\rho)$ acts on the function $(s, y) \mapsto P(\tau + \rho, t + \rho) f(s, \cdot)(y)$ and that $S(t + \rho)$ acts on the function $(s, y) \mapsto f(y)$. The process

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), (X(t) : T \geqslant t \geqslant 0), (\vee_t : T \geqslant t \geqslant 0), (E, \mathcal{E})\right\} \tag{7.6}$$

is the strong Markov process generated by $-K_0$; it is supposed to have continuous paths. In the space $C(E)$ the operator K_0 is considered as a local operator in the sense that a function $f \in C(E)$ belongs to its domain if there exists a function $g \in C((\tau, T) \times E)$ such that for every open subset U of E together with every compact subset K of U we have

$$\lim_{h \downarrow 0} \sup_{(\tau,x) \in [0,T-h] \times K} \left| g(\tau, x) - \frac{f(\tau, x) - \mathbb{E}_{\tau,x}[f(\tau+h, X(\tau+h)) : \tau_U > \tau+h]}{h} \right|$$
$$= 0.$$

Here τ_U is the first exit time from U: $\tau_U = \inf\{t > 0 : X(t) \in E \backslash U\}$. We write $g = -K_0 f$. From Proposition 1.6 in [Demuth and van Casteren (2000)] page 9 it follows that the constant function $\mathbf{1}$ belongs to the domain of K_0 and that $K_0 \mathbf{1} = 0$, provided K_0 is time-independent.

Remark 7.2. In a more classical context in e.g. L^p-spaces the operator K_0 can often be considered as a differential operator in "distributional" sense. In a physical context the operators $K_0(s)$, $s \in [0,T]$, are considered as self-adjoint operators in $L^2(E,m)$. It is noticed that there exists a close relationship between the viscous Burgers' equation (in an open subset of $\mathbb{R}^d$)

$$-\frac{\partial U}{\partial t} + U \cdot \nabla U - \frac{1}{2}\Delta U = \nabla V,$$

and the Hamilton-Jacobi-Bellman equation. If we write the vector field U in the form $U = \nabla\varphi$, then the function φ satisfies

$$-\frac{\partial \varphi}{\partial t} + \frac{1}{2}\nabla\varphi \cdot \nabla\varphi - \frac{1}{2}\Delta\varphi = V + \text{ constant}.$$

7.2 The Hamilton-Jacobi-Bellman equation and its solution

In this section we will mainly be concerned with the Hamilton-Jacobi-Bellman equation as exhibited in equation (7.12) below. We have the following result for generators of diffusions: it refines Theorem 2.4 in [Zambrini (1998a)]. Its proof is contained in the proof of Theorem 7.3. We begin by inserting a definition.

Definition 7.1. Fix a function $v : (t_0, T] \times E \to \mathbb{R}$ in $D(D_1 - K_0)$, where, as above, $D_1 = \dfrac{\partial}{\partial t}$ is differentiation with respect to t. Let the process

$$\left\{(\Omega, \mathcal{F}, \mathbb{P}_{t,x}), ((q_v(t), t) : t \geqslant 0), (\vee_t : t \geqslant 0), \left(\mathbb{R}^+ \times E, \mathcal{B}_{\mathbb{R}^+} \otimes \mathcal{E}\right)\right\}$$

be the Markov process generated by the operator $-K_v + D_1$, where K_v is defined by $K_v(f)(t,x) = K_0 f(t,x) + \Gamma_1(v,f)(t,x)$. Here, $\mathcal{B}_{\mathbb{R}^+}$ denotes the Borel field of $\mathbb{R}^+$, and by $\Gamma_1(v,f)(t,x)$ we mean

$$\begin{aligned}
&\Gamma_1(v,f)(t,x) \qquad (7.7)\\
&= \lim_{s \downarrow t} \frac{1}{s-t}\mathbb{E}_{t,x}\left[(v(s,X(s)) - v(t,X(t)))\,(f(X(s),s) - f(t,X(t)))\right].
\end{aligned}$$

It is also believed that the following version of the Cameron-Martin formula is true. For all finite n-tuples $t_1, \dots, t_n$ in $(0,\infty)$ the identity (7.9) is valid:

$$\begin{aligned}
&\mathbb{E}_{t,x}^{M_{v,t}}\left[\prod_{j=1}^{n} f_j\left(t_j + t, X\left(t_j + t\right)\right)\right] \qquad (7.8)\\
&= \mathbb{E}_{t,x}\left[\exp\left(-\frac{1}{2}\int_t^T \Gamma_1(v,v)(\tau, X(\tau))\,d\tau - M_{v,t}(T)\right)\right.
\end{aligned}$$

$$\times \prod_{j=1}^{n} f_j \left(t_j + t, X\left(t_j + t\right)\right)\Bigg]$$

$$= \mathbb{E}_{t,x}\left[\prod_{j=1}^{n} f_j \left(t_j + t, q_v\left(t_j + t\right)\right)\right] \tag{7.9}$$

where the $\mathbb{E}_{t,x}$-martingale $M_{v,t}(s)$, $s \geqslant t$, is given by

$$M_{v,t}(s) = v\left(s, X(s)\right) - v\left(t, X(t)\right) + \int_t^s \left(-\frac{\partial}{\partial \tau} + K_0\right) v\left(\tau, X(\tau)\right) d\tau. \tag{7.10}$$

Its quadratic variation part $\langle M_{v,t}\rangle(s) := \langle M_{v,t}, M_{v,t}\rangle(s)$ is given by

$$\langle M_{v,t}\rangle(s) = \int_t^s \Gamma_1(v,v)\left(\tau, X(\tau)\right) d\tau. \tag{7.11}$$

The equality in (7.8) serves as a definition of the measure $\mathbb{P}_{t,x}^{M_{v,t}}(\cdot)$, and the equality in (7.9) is a statement.

The formula in (7.11) is explained in (the proof of) Proposition 5.3. Next we formulate a theorem in which we use the notation introduced in Definition 7.1. The next theorem is the same as Theorem 5.8 with $K_0(s)$ instead of $-L(s)$.

Theorem 7.1. *Let $\chi : (t_0, T] \times E \to [0, \infty]$ be a function such that*

$$\mathbb{E}_{t,x}^{M_{v,t}}\left[\left|\log \chi\left(T, X(T)\right)\right|\right], \quad v \in D\left(D_1 - K_0\right)$$

is finite for $t_0 < t \leqslant T$. Here $T > t_0 \geqslant 0$ are fixed times and

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), \left(\vee_t : T \geqslant t \geqslant 0\right), \left(E, \mathcal{E}\right)\right\}$$

is the strong Markov process generated by the operator family $-K_0(s)$, $0 \leqslant s \leqslant T$. Let S_L be a solution to the following Riccati type equation. This equation is called the Hamilton-Jacobi-Bellman equation. For $t_0 < s \leqslant T$ and $x \in E$ the following identity is true:

$$\begin{cases} -\dfrac{\partial S_L}{\partial s}(s,x) + \dfrac{1}{2}\Gamma_1\left(S_L, S_L\right)(s,x) + K_0(s) S_L(s,x) - V(s,x) = 0; \\ S_L(T,x) = -\log \chi(T,x), \quad x \in E. \end{cases} \tag{7.12}$$

Then for any real valued $v \in D\left(D_1 - K_0\right)$ the following inequality is valid:

$$S_L(t,x) \leqslant \mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T \left(\frac{1}{2}\Gamma_1(v,v) + V\right)(\tau, X(\tau)) d\tau\right] - \mathbb{E}_{t,x}^{M_{v,t}}\left[\log \chi(T, X(T))\right], \tag{7.13}$$

and equality is attained for the "Lagrangian action" $v = S_L$.

By definition $E_{t,x}[Y]$ is the expectation, conditioned at $X(t) = x$, of the random variable Y which is measurable with respect to the information from the future: i.e. with respect to $\sigma\{X(s) : s \geqslant t\}$. The measure $\mathbb{P}_{t,x}^{M_{v,t}}$ is defined in equality (7.8) below. Put $\eta_\chi(t,x) = \exp(-S_L(t,x))$, where S_L satisfies (7.12). From (7.1) it follows that $\left(\frac{\partial}{\partial t} - \left(K_0 \dot{+} V\right)\right)\eta_\chi(t,x) = 0$, provided that $K_0\mathbf{1}(t,x) = 0$ for all $(t,x) \in [0,T] \times E$. The proof of Theorem 7.1 can be found in [Van Casteren (2001)]; Theorem 7.1 is superseded by the second inequality in assertion (i) of Theorem 7.3.

Next, let $\chi : [t,T] \times E \to [0,\infty]$ be as in Theorem 7.1. In what follows we write $D_1 = \frac{\partial}{\partial t}$. We also write $D_1\varphi = \dot{\phi}$. What is the relationship between the following expressions?

$$\sup_{\Phi \in D(D_1 - K_0)} \left\{\Phi(t,x) : -\dot{\Phi} + K_0\Phi + \frac{1}{2}\Gamma_1(\Phi,\Phi) \leqslant V,\ \Phi(T,\cdot) \leqslant -\log\chi(T,\cdot)\right\}; \tag{7.14}$$

$$-\log \mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V(\sigma, X(\sigma))\,d\sigma\right)\chi(T, X(T))\right]; \tag{7.15}$$

$$\begin{aligned}\inf_{\Phi \in D(D_1 - K_0)} &\left\{\mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T \left(\frac{1}{2}\Gamma_1(v,v) + V\right)(\tau, X(\tau))d\tau\right]\right. \\ &\left. - \mathbb{E}_{t,x}^{M_{v,t}}\left[\log\chi(T, X(T))\right]\right\};\end{aligned} \tag{7.16}$$

$$\inf_{\Phi \in D(D_1 - K_0)} \left\{\Phi(t,x) : -\dot{\Phi} + K_0\Phi + \frac{1}{2}\Gamma_1(\Phi,\Phi) \geqslant V,\ \Phi(T,\cdot) \geqslant -\log\chi(T,\cdot)\right\}. \tag{7.17}$$

In order that everything works appropriately we need the following definition and lemma.

Definition 7.2. The potential $V : [0,T] \times E \to \mathbb{R}$ satisfies the Myadera perturbation condition, provided that

$$\begin{aligned}&\limsup_{s \downarrow 0} \sup_{(\tau,x) \in [0,T-s] \times E} \mathbb{E}_{\tau,x}\left[\int_\tau^{s+\tau} V_-(\rho, X(\rho))\,d\rho\right] \\ &= \limsup_{s \downarrow 0} \sup_{(\tau,x) \in [0,T-s] \times E} \int_\tau^{\tau+s} P(\tau,\rho)\,V_-(\rho,\cdot)(x)d\rho < 1.\end{aligned} \tag{7.18}$$

For more information on Myadera perturbations the reader is referred to e.g. Räbiger, *et al.* [Räbiger *et al.* (1996, 2000)].

Lemma 7.1. *Suppose that*

$$\alpha := \limsup_{s\downarrow 0} \sup_{(\tau,x)\in[0,T-s]\times E} \int_{\tau}^{\tau+s} P(\tau,\rho) V_-(\rho,\cdot)(x) d\rho < 1. \tag{7.19}$$

Then

$$\sup_{(\tau,x)\in[0,T]\times E} \mathbb{E}_{\tau,x}\left[\exp\left(\int_{\tau}^{T} V_-(\rho, X(\rho))\, d\rho\right)\right] < \infty. \tag{7.20}$$

Proof. Choose $n \in \mathbb{N}$ so large that

$$\alpha_n := \sup_{(\tau,x)\in[0,T]\times E} \int_{\tau}^{(n\tau+T)/(n+1)} P(\tau,\rho) V_-(\rho,\cdot)(x) d\rho < 1. \tag{7.21}$$

By (7.18) such a choice is possible. For $\tau \in [0,T]$ fixed we choose a subdivision of the interval $[\tau, T]$ in such a way that

$$\tau = \tau_0 < \tau_1 < \cdots < \tau_n < \tau_{n+1} = T, \quad \text{where } \tau_j = \frac{n+1-j}{n+1}\tau + \frac{j}{n+1}T.$$

Notice that $\tau_{k+1} - \tau_k = (T-\tau)/(n+1) \leqslant T/(n+1)$. Then by the Markov property we have

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\exp\left(\int_{\tau}^{T} V_-(\rho, X(\rho))\, d\rho\right)\right] \\
&= \mathbb{E}_{\tau,x}\left[\prod_{j=0}^{n} \exp\left(\int_{\tau_j}^{\tau_{j+1}} V_-(\rho, X(\rho))\, d\rho\right)\right] \\
&= \mathbb{E}_{\tau,x}\left[\prod_{j=0}^{n-1} \exp\left(\int_{\tau_j}^{\tau_{j+1}} V_-(\rho, X(\rho))\, d\rho\right)\right. \\
&\qquad \left.\times \mathbb{E}_{\tau_n, X(\tau_n)}\left[\exp\left(\int_{\tau_n}^{\tau_{n+1}} V_-(\rho, X(\rho))\, d\rho\right)\right]\right]
\end{aligned}$$

(by induction)

$$\leqslant \prod_{k=0}^{n} \sup_{y\in E} \mathbb{E}_{\tau_k,y}\left[\exp\left(\int_{\tau_k}^{\tau_{k+1}} V_-(\rho, X(\rho))\, d\rho\right)\right]. \tag{7.22}$$

We also have

$$\begin{aligned}
&\mathbb{E}_{\tau_k,y}\left[\exp\left(\int_{\tau_k}^{\tau_{k+1}} V_-(\rho, X(\rho))\, d\rho\right)\right] \\
&= 1 + \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \mathbb{E}_{\tau_k,y}\left[\left(\int_{\tau_k}^{\tau_{k+1}} V_-(\rho, X(\rho))\, d\rho\right)^{\ell}\right]
\end{aligned}$$

$$= 1 + \sum_{\ell=1}^{\infty} \mathbb{E}_{\tau_k,y}\left[\int_{\tau_k<\rho_1<\dots<\rho_\ell<\tau_{k+1}}\int \prod_{j=1}^{\ell} V_-\left(\rho_j, X\left(\rho_j\right)\right) d\rho_\ell \dots d\rho_1\right]$$

(again Markov property)

$$= 1 + \sum_{\ell=1}^{\infty} \mathbb{E}_{\tau_k,y}\left[\int_{\tau_k<\rho_1<\dots<\rho_{\ell-1}}\int \prod_{j=1}^{\ell-1} V_-\left(\rho_j, X\left(\rho_j\right)\right)\right.$$

$$\left.\mathbb{E}_{\rho_{\ell-1},X(\rho_{\ell-1})}\left[\int_{\rho_{\ell-1}}^{\tau_{k+1}} V_-\left(\rho_\ell, X\left(\rho_\ell\right)\right) d\rho_\ell\right] d\rho_{\ell-1}\dots d\rho_1\right]$$

$$\leqslant \sum_{\ell=0}^{\infty}\left(\sup_{(\rho,z)\in[\tau_k,\tau_{k+1}]\times E} \mathbb{E}_{\rho,z}\left[\int_{\rho}^{\tau_{k+1}} V_-\left(s, X(s)\right) ds\right]\right)^{\ell}$$

(notice the inequality $\tau_{n+1} \leqslant \rho + (T-\tau)/(n+1)$)

$$\leqslant \sum_{\ell=0}^{\infty} \alpha_n^{\ell} = \frac{1}{1-\alpha_n}, \tag{7.23}$$

where in the final step of (7.23) we used (7.21). From (7.22) and (7.23) we obtain (7.20).

This completes the proof of Lemma 7.1. □

We also have to insert the standard Feynman-Kac formula, and its properties related to the strict topology. In addition, we have to discuss matters like stability and consistency of families of Kato-type or Myadera potentials. More precisely, let $(V_k)_{k\in\mathbb{N}}$ be a sequence of potentials which satisfies, uniformly in k, a condition like (7.19). Under what consistency (or convergence) conditions are we sure that the corresponding perturbed evolutions $\{P_{V_k}(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$, $k \in \mathbb{N}$, converges to an evolution of the form $\{P_V(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$. In addition, we want this convergence to behave in such a way that the operators $P_V(s,t)$, $0 \leqslant s \leqslant t \leqslant T$, assign bounded continuous functions to bounded continuous functions, provided the same is true for each of the operators $P_{V_k}(s,t)$, $k \in \mathbb{N}$, $0 \leqslant s \leqslant t \leqslant T$.

Theorem 7.2. *Let the Feller evolution $\{P(s,t) : \tau \leqslant s \leqslant t \leqslant T\}$ be the transition probabilities of the Markov process in (7.6). Let $V : [0,T] \times E \to \mathbb{R}$ be a Myadera type potential function with the following properties:*

(i) *Its negative part satisfies (7.19).*

(ii) *For every k, $\ell \in \mathbb{N}$, and $f \in C_b([0,T] \times E)$, the function*

$$(\tau, x, t) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^{(\tau+t)\wedge T} V_{k,\ell}(\rho, X(\rho)) f(\rho, X(\rho))\, d\rho\right]$$

is continuous. Here $V_{k,\ell} = (V \wedge \ell) \vee (-k)$.

(iii) *The following equalities hold for all compact subsets K of E:*

$$\lim_{\ell\to\infty} \sup_{(\tau,x)\in[0,T]\times K} \mathbb{E}_{\tau,x}\left[\int_\tau^T 0 \vee (V-\ell)(\rho, X(\rho))\, d\rho\right] = 0, \quad \textit{and}$$

$$\lim_{k\to\infty} \sup_{(\tau,x)\in[0,T]\times K} \mathbb{E}_{\tau,x}\left[\int_\tau^T 0 \vee (-V-k)(\rho, X(\rho))\, d\rho\right] = 0. \tag{7.24}$$

(iv) *The function V satisfies*

$$\sup_{(\tau,x)\in[0,T]\times E} \mathbb{E}_{\tau,x}\left[\int_\tau^T |V(\rho, X(\rho))|\, d\rho\right] < \infty.$$

Then the functions

$$(\tau, x, t) \mapsto \mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V(\rho, X(\rho))\, d\rho\right) f(X(t))\right], \quad f \in C_b(E), \tag{7.25}$$

are bounded continuous functions.

Remark 7.3. Suppose that the functions in (7.24) are continuous; i.e. suppose that for every $k \in \mathbb{N}$ the functions

$$(\tau, x) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T 0 \vee (V-k)(\rho, X(\rho))\, d\rho\right] \quad \text{and}$$

$$(\tau, x) \mapsto \mathbb{E}_{\tau,x}\left[\int_\tau^T 0 \vee (-V-k)(\rho, X(\rho))\, d\rho\right]$$

are continuous. Then (iii) is a consequence of (iv). From (iv) it follows that the pointwise limits in (7.24) are zero. By Dini's lemma this convergence occurs uniformly on compact subsets of $[0,T] \times E$. Also observe that the limits in (7.24) decrease monotonically with increasing ℓ and k respectively.

Proof. [Proof of Theorem 7.2.] Let $f \in C_b(E)$ be such that $\|f\|_\infty \leqslant 1$. First we notice that $-V_- \leqslant V_{k,\ell} \leqslant V_+$, and hence $|V - V_{k,\ell}| \leqslant |V|$. It follows that

$$\mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V(\rho, X(\rho))\, d\rho\right) f(X((\tau+t)\wedge T))\right]$$

$$
\begin{aligned}
&- \mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho, X(\rho)\right) d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right] \\
&= \int_0^1 \mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} \left\{(1-s)V\left(\rho, X(\rho)\right) + sV_{k,\ell}\left(\rho, X(\rho)\right)\right\} d\rho\right)\right. \\
&\qquad \left.\int_\tau^{(\tau+t)\wedge T} \left(V - V_{k,\ell}\right)\left(\rho, X(\rho)\right) d\rho f\left(X\left((\tau+t)\wedge T\right)\right)\right] ds,
\end{aligned}
$$

and hence

$$
\begin{aligned}
&\left|\mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V\left(\rho, X(\rho)\right) d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right]\right. \\
&\qquad \left. -\mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho, X(\rho)\right) d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right]\right| \\
&\leqslant \int_0^1 \mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} \left\{(1-s)V\left(\rho, X(\rho)\right) + sV_{k,\ell}\left(\rho, X(\rho)\right)\right\} d\rho\right)\right. \\
&\qquad \left.\left|\int_\tau^{(\tau+t)\wedge T} \left(V - V_{k,\ell}\right)\left(\rho, X(\rho)\right) d\rho\right|\right] ds \left\|f\right\|_\infty \\
&\leqslant \mathbb{E}_{\tau,x}\left[\exp\left(\int_\tau^{(\tau+t)\wedge T} V_-\left(\rho, X(\rho)\right) d\rho\right)\right. \\
&\qquad \left.\left|\int_\tau^{(\tau+t)\wedge T} \left(V - V_{k,\ell}\right)\left(\rho, X(\rho)\right) d\rho\right|\right] \\
&\leqslant \left(\mathbb{E}_{\tau,x}\left[\exp\left(\frac{2m+2}{2m+1}\int_\tau^{(\tau+t)\wedge T} V_-\left(\rho, X(\rho)\right) d\rho\right)\right]\right)^{(2m+1)/(2m+2)} \\
&\qquad \left(\mathbb{E}_{\tau,x}\left[\left|\int_\tau^{(\tau+t)\wedge T} \left(V - V_{k,\ell}\right)\left(\rho, X(\rho)\right) d\rho\right|^{2m+2}\right]\right)^{1/(2m+2)}. \qquad (7.26)
\end{aligned}
$$

In (7.26) we choose m so large that

$$
\sup_{(s,y)\in[0,T]\times E} \mathbb{E}_{s,y}\left[\exp\left(\frac{2m+2}{2m+1}\int_\tau^{(\tau+t)\wedge T} V_-\left(\rho, X(\rho)\right) d\rho\right)\right] < \infty. \quad (7.27)
$$

From Lemma 7.1 it follows that such a choice of m is possible: see (7.19) and (7.21). From the Markov property we infer

$$
\frac{1}{(2m+2)!}\mathbb{E}_{\tau,x}\left[\left|\int_\tau^{(\tau+t)\wedge T} \left(V - V_{k,\ell}\right)\left(\rho, X(\rho)\right) d\rho\right|^{2m+2}\right]
$$

$$
\begin{aligned}
&\leqslant \frac{1}{(2m+2)!}\mathbb{E}_{\tau,x}\left[\int_{\tau}^{(\tau+t)\wedge T}\left|\left(V-V_{k,\ell}\right)(\rho,X(\rho))\right|^{2m+2}d\rho\right]\\
&=\mathbb{E}_{\tau,x}\left[\int\limits_{\tau<\rho_1<\dots<\rho_{2m+2}<(\tau+t)\wedge T}\int\prod_{j=1}^{2m+2}\left|(V-V_{k,\ell})(\rho_j,X(\rho_j))\right|\,d\rho_{2m+2}\dots d\rho_1\right]\\
&=\mathbb{E}_{\tau,x}\left[\int_{\tau<\rho_1<\dots<\rho_{2m+1}<(\tau+t)\wedge T}\int\prod_{j=1}^{2m+1}\left|\left(V-V_{k,\ell}\right)\left(\rho_j,X\left(\rho_j\right)\right)\right|\right.\\
&\qquad\times\mathbb{E}_{\rho_{2m+1},X(\rho_{2m+1})}\left[\int_{\rho_{2m+1}}^{(\tau+t)\wedge T}\left|\left(V-V_{k,\ell}\right)\left(\rho_{2m+2},X\left(\rho_{2m+2}\right)\right)\right|\,d\rho_{2m+2}\right]\\
&\qquad\qquad\left. d\rho_{2m+1}\dots d\rho_1\right]\\
&\leqslant\mathbb{E}_{\tau,x}\left[\int\limits_{\tau<\rho_1<\dots<\rho_{2m+1}<(\tau+t)\wedge T}\int\prod_{j=1}^{2m+1}\left|(V-V_{k,\ell})(\rho_j,X(\rho_j))\right|d\rho_{2m+1}\dots d\rho_1\right]\\
&\qquad\sup_{(s,y)\in[\tau,(\tau+t)\wedge T]\times E}\mathbb{E}_{s,y}\left[\int_{s}^{(\tau+t)\wedge T}\left|\left(V-V_{k,\ell}\right)\left(\rho,X\left(\rho\right)\right)\right|\,d\rho\right]
\end{aligned}
$$

(use induction)

$$
\begin{aligned}
&\leqslant\mathbb{E}_{\tau,x}\left[\int_{\tau}^{(\tau+t)\wedge T}\left|\left(V-V_{k,\ell}\right)\left(\rho_1,X\left(\rho_1\right)\right)\right|\,d\rho_1\right]\\
&\qquad\sup_{(s,y)\in[\tau,(\tau+t)\wedge T]\times E}\left(\mathbb{E}_{s,y}\left[\int_{s}^{(\tau+t)\wedge T}\left|\left(V-V_{k,\ell}\right)\left(\rho,X\left(\rho\right)\right)\right|\,d\rho\right]\right)^{2m+1}\\
&\leqslant\mathbb{E}_{\tau,x}\left[\int_{\tau}^{(\tau+t)\wedge T}\left|\left(V-V_{k,\ell}\right)\left(\rho_1,X\left(\rho_1\right)\right)\right|\,d\rho_1\right]\\
&\qquad\sup_{(s,y)\in[\tau,(\tau+t)\wedge T]\times E}\left(\mathbb{E}_{s,y}\left[\int_{s}^{(\tau+t)\wedge T}\left|V\left(\rho,X\left(\rho\right)\right)\right|\,d\rho\right]\right)^{2m+1}\\
&\leqslant\left(\mathbb{E}_{\tau,x}\left[\int_{\tau}^{(\tau+t)\wedge T}0\vee\left(V-\ell\right)\left(\rho_1,X\left(\rho_1\right)\right)\,d\rho_1\right]\right.\\
&\qquad\left.+\mathbb{E}_{\tau,x}\left[\int_{\tau}^{(\tau+t)\wedge T}0\vee\left(-V-k\right)\left(\rho_1,X\left(\rho_1\right)\right)\,d\rho_1\right]\right)
\end{aligned}
$$

$$\times \sup_{(s,y)\in[\tau,(\tau+t)\wedge T]\times E}\left(\mathbb{E}_{s,y}\left[\int_s^{(\tau+t)\wedge T}\left|V\left(\rho,X\left(\rho\right)\right)\right|\,d\rho\right]\right)^{2m+1}. \quad (7.28)$$

From (7.26), (7.27), (7.28), assumptions (iii) and (iv) it follows that, uniformly on compact subsets of $[0,T]\times E$, the following equality holds:

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V\left(\rho,X(\rho)\right)d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right]\\
&= \lim_{k\to\infty}\lim_{\ell\to\infty}\mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho,X(\rho)\right)d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right].
\end{aligned} \quad (7.29)$$

In order to finish the proof of Theorem 7.2 we need to establish the continuity of the function

$$(\tau,x,t)\mapsto \mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho,X(\rho)\right)d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right]. \quad (7.30)$$

By expanding the exponential in (7.30), using the Markov property together with assumption (ii) the continuity of the function in (7.30) follows. More precisely, we have

$$\begin{aligned}
&\mathbb{E}_{\tau,x}\left[\exp\left(-\int_\tau^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho,X(\rho)\right)d\rho\right) f\left(X\left((\tau+t)\wedge T\right)\right)\right]\\
&= \sum_{n=0}^\infty \frac{(-1)^n}{n!}\mathbb{E}_{\tau,x}\left[\left(\int_0^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho,X(\rho)\right)d\rho\right)^n f\left(X(\tau+t\wedge T)\right)\right]\\
&= \mathbb{E}_{\tau,x}\left[f\left(X(\tau+t\wedge T)\right)\right]\\
&\quad+\sum_{n=1}^\infty(-1)^n\int_{\tau<\rho_1<\dots<\rho_n<(\tau+t)\wedge T}\int\\
&\qquad \mathbb{E}_{\tau,x}\left[\prod_{j=1}^n V_{k,\ell}\left(\rho_j,X\left(\rho_j\right)\right) f\left(X(\tau+t\wedge T)\right)\right]d\rho_n\dots d\rho_1
\end{aligned}$$

(Markov property)

$$\begin{aligned}
&= \mathbb{E}_{\tau,x}\left[f\left(X(\tau+t\wedge T)\right)\right]\\
&\quad+\sum_{n=1}^\infty(-1)^k\int_{\tau<\rho_1<\dots<\rho_{n-1}<(\tau+t)\wedge T}\int \mathbb{E}_{\tau,x}\left[\prod_{j=1}^{n-1} V_{k,\ell}\left(\rho_j,X\left(\rho_j\right)\right)\right.
\end{aligned}$$

$$\mathbb{E}_{\rho_{n-1},X(\rho_{n-1})}\left[\left[\int_{\rho_{n-1}}^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho_n, X\left(\rho_n\right)\right) d\rho_n\, f\left(X(\tau+t)\wedge T\right)\right]\right]$$
$$d\rho_{n-1}\ldots d\rho_1. \tag{7.31}$$

Notice that by assumption (ii) the function

$$(\rho,t,y)\mapsto \mathbb{E}_{\rho,y}\left[\int_{\rho_{n-1}}^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho_n, X\left(\rho_n\right)\right)\, d\rho_n\, f\left(X\left((\tau+t)\wedge T\right)\right)\right] \tag{7.32}$$
$$= \mathbb{E}_{\rho,y}\left[\int_{\rho_{n-1}}^{(\tau+t)\wedge T} V_{k,\ell}\left(\rho_n, X\left(\rho_n\right)\right) \mathbb{E}_{\rho_n,X(\rho_n)}\left[f\left(X\left((\tau+t)\wedge T\right)\right)\right] d\rho_n\right].$$

By induction with respect to n it follows that each term in the right-hand side of (7.31) is continuous. The series in (7.31) being uniformly convergent yields the continuity of the functions in (7.25).

This concludes the proof of Theorem 7.2. □

7.3 The Hamilton-Jacobi-Bellman equation and viscosity solutions

A result which is somewhat more general than Theorem 7.1 reads as follows. As above, we work in the space $C_b\left((t_0,T]\times E\right)$, where $T>t_0\geqslant 0$ is fixed. The fact that the non-linear Feynman-Kac formula (7.33) yields a viscosity solution to the HJB-equation in (7.12) is proved by analytic means: see the proof of assertion (iii) below. In the semi-linear case this kind result was established by means of a stopping time argument: see the proof of Theorem 6.3 in Chapter 6. In fact using a stopping time argument yields a more refined result; one gets local rather than global inequalities.

Theorem 7.3.

(i) *The following inequalities are valid:*

$$\sup_{\Phi\in D(D_1-K_0)}\left\{\Phi(t,x):\ -\dot{\Phi}+K_0\Phi+\frac{1}{2}\Gamma_1\left(\Phi,\Phi\right)\leqslant V,\right.$$
$$\left.\Phi(T,\cdot)\leqslant -\log\chi\left(T,\cdot\right)\right\}$$
$$\leqslant -\log\mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V\left(\sigma,X(\sigma)\right)d\sigma\right)\chi\left(T,X(T)\right)\right]$$

$$\leqslant \inf_{v\in D(D_1-K_0)}\left\{\mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T\left(\frac{1}{2}\Gamma_1\left(v,v\right)+V\right)(\tau,X(\tau))d\tau\right]\right.$$

$$\left.-\mathbb{E}_{t,x}^{M_{v,t}}\left[\log\chi\left(T,X(T)\right)\right]\right\}$$

$$\leqslant \inf_{\Phi\in D(D_1-K_0)}\left\{\Phi(t,x):\ -\dot{\Phi}+K_0\Phi+\frac{1}{2}\Gamma_1\left(\Phi,\Phi\right)\geqslant V,\right.$$

$$\left.\Phi(T,\cdot)\geqslant-\log\chi\left(T,\cdot\right)\right\}.$$

(ii) *If the function S_L defined by the non-linear Feynman-Kac formula*

$$S_L(t,x)=-\log\mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V\left(\sigma,X(\sigma)\right)d\sigma\right)\chi\left(T,X(T)\right)\right] \tag{7.33}$$

belongs to $D\left(D_1-K_0\right)$, then the above 4 quantities are equal. Moreover the function S_L satisfies the Hamilton-Jacobi-Bellman equation (7.12). The same is true if the expressions in (7.14) and in (7.17) are equal.

(iii) *In general the function in (7.33) is a viscosity solution of the Hamilton-Jacobi-Bellman equation (7.12). This means that if $(t,x)\in(t_0,T]\times E$ is given and if $\varphi\in D\left(D_1-K_0\right)$ has the property that*

$$\left[S_L-\varphi\right](t,x)=\sup\left\{\left[S_L-\varphi\right](s,y):(s,y)\in[t,T]\times E\right\},$$

then

$$\left[-\dot{\varphi}+K_0\varphi\right](t,x)+\frac{1}{2}\Gamma_1\left(\varphi,\varphi\right)(t,x)\leqslant V(t,x). \tag{7.34}$$

It also means that if (t,x) belongs to $(t_0,T]\times E$ and if $\varphi\in D\left(D_1-K_0\right)$ has the property that

$$\left[S_L-\varphi\right](t,x)=\inf\left\{\left[S_L-\varphi\right](s,y):(s,y)\in[t,T]\times E\right\},$$

then

$$\left[-\dot{\varphi}+K_0\varphi\right](t,x)+\frac{1}{2}\Gamma_1\left(\varphi,\varphi\right)(t,x)\geqslant V(t,x). \tag{7.35}$$

(iv) *If for all $(t,x)\in(t_0,T]\times E$ the expression*

$$\mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V\left(\sigma,X(\sigma)\right)d\sigma\right)\chi\left(T,X(T)\right)\right],$$

is strictly positive, then the following equality is valid:

$$-\log\mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V\left(\sigma,X(\sigma)\right)d\sigma\right)\chi\left(T,X(T)\right)\right] \tag{7.36}$$

$$= \inf_{v\in D(D_1-K_0)} \left\{ \mathbb{E}_{t,x}^{M_{v,t}} \left[\int_t^T \left(\frac{1}{2}\Gamma_1(v,v) + V \right)(\tau, X(\tau))d\tau \right] - \mathbb{E}_{t,x}^{M_{v,t}} \left[\log\chi(T, X(T)) \right] \right\}. \tag{7.37}$$

(v) *Let S be a viscosity solution to* (7.12)*. Suppose that for every* $(t,x) \in (t_0,T] \times E$ *there exist functions* φ_1 *and* $\varphi_2 \in D(K_0)$ *such that*

$$(S-\varphi_1)(t,x) = \sup_{y\in E, T>s>t} (S-\varphi_1)(s,y), \text{ and} \tag{7.38}$$

$$(S-\varphi_2)(t,x) = \inf_{y\in E, T>s>t} (S-\varphi_2)(s,y). \tag{7.39}$$

Then $S = S_L$. More precisely, in the presence of (7.39) *and* (7.38) *the 4 quantities in assertion* (i) *are equal.*

Notice that the formula in (7.33) is the same as formula (5.33) in Chapter 5. The main difference is notational: in Chapter 5 and the other chapters we write $L(s)$ instead of $-K_0(s)$. The notation $K_0 = \{K_0(s) : 0 \leqslant s \leqslant T\}$ refers to a self-adjoint unperturbed (or free) Hamiltonian, which is often written as H_0, which usually is given by $H_0 = -\dfrac{\hbar^2}{2m}\Delta$. The Schrödinger equation is then given by $(H_0 + V)\psi = i\hbar\dfrac{\partial\psi}{\partial t}$. Here V stands for a potential function, which belongs to a certain Kato type class. In mathematics Planck's normalized constant $\hbar$ and the particle mass m are often set equal to 1.

Remark 7.4. It would be nice to have explicit, and easy to check, conditions on the function V which guarantee the strict positivity of the expression

$$\mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V(\sigma, X(\sigma))\,d\sigma\right),\ X(T) \in B\right],$$

where B is any compact subset of E. Another problem which poses itself is the following. What can be done if in equation (7.12) the expression $\Gamma_1(S_L, S_L)$ is replaced with $(\Gamma_1(S_L,S_L))^p$, $p > 0$. If $0 < p < 1$, then the equation probably can be treated by the use of *branching processes*: see e.g. [Etheridge (2000)] or [Dawson and Perkins (1999)].

Remark 7.5. Another point of concern is the Novikov condition which is required to be sure that processes of the form

$$t \mapsto \exp\left(-M(t) - \frac{1}{2}\langle M, M\rangle(t)\right) \quad \text{and} \tag{7.40}$$

$$t \mapsto \exp\left(-M(t) - \frac{1}{2}\langle M, M\rangle(t)\right)(M(t) + \langle M.M\rangle(t)) \tag{7.41}$$

are martingales. The Novikov condition reads as follows. Let $M(t)$ be a martingale, and suppose that $\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M, M\rangle(t)\right)\right]$ is finite for all $t \geqslant 0$. Then the process in (7.40) is a martingale. So, strictly speaking, we have to assume in the sequel that the Novikov condition is satisfied: i.e. all the expectations ($x \in E$, $t_0 \leqslant t < s \leqslant T$)

$$\mathbb{E}_{t,x}\left[\exp\left(\frac{1}{2}\int_t^s \Gamma_1(\varphi, \varphi)\big(\tau, X(\tau)\big)d\tau\right)\right]$$

are supposed to be finite; otherwise we will only get local martingales. For more details on the Novikov condition see e.g. [Revuz and Yor (1999)], Corollary 1.16, page 309. Novikov's condition is also treated in Theorem 1.6 and its Corollary 1.3 in Chapter 1.

Remark 7.6. Another problem is about the uniqueness of the viscosity solution of equation (7.12). In order to address this problem we use a technique, which is related to the methods used in [Dynkin and Kuznetsov (1996b)] p. 26 ff, and [Dynkin and Kuznetsov (1996a)], p. 1969 ff. Among other things we tried the method of "doubling the number of variables" as advertised in [Evans (1998)] page 547, but it did not work out so far. We also tried (without success) the jet bundle technique in [Crandall *et al.* (1992b)]. To be precise we use a martingale technique combined with sub- and super-solutions: see assertion (v) of Theorem 7.3.

First we insert the following proposition.

Proposition 7.1.

(i) *The operator $D_1 - K_0 - V$ extends to a generator of a semigroup $\exp(s(D_1 - K_0 - V))$, $s \geqslant 0$, given by*

$$\begin{aligned}&\exp(s(D_1 - K_0 - V))\,\Phi(t,x)\\ &= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V(\tau, X(\tau))\,d\tau\right)\Phi(s+t, X(s+t))\right].\end{aligned} \tag{7.42}$$

(ii) *Let the function $S_L(t,x)$ be given by*

$$S_L(t,x) = -\log \mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V(\tau, X(\tau))\,d\tau\right)\chi(T, X(T))\right].$$

Then the following identity is valid

$$[\exp(s(D_1 - K_0 - V))\exp(-S_L)](t,x) = \exp(-S_L(t,x)). \tag{7.43}$$

(iii) *Let* $\Psi : (t_0, T] \times E \to \mathbb{R}$ *be a function belonging to* $D(D_1 - K_0)$, *and let* $V_0 : (t_0, T] \times E \to \mathbb{R}$ *be a function for which* ($(s,y) \in (t_0, t] \times E$)

$$\Psi(s,y) = -\log \mathbb{E}_{s,y}\left[\exp\left(-\int_s^t V_0\big(\tau, X(\tau)\big)d\tau - \Psi\big(t, X(t)\big)\right)\right]. \tag{7.44}$$

Then $-D_1\Psi + K_0\Psi + \frac{1}{2}\Gamma_1(\Psi,\Psi) = V_0$ *on* $(t_0, t] \times E$.

(iv) *Conversely, let* $\Psi : (t_0, T] \times E \to \mathbb{R}$ *be a function belonging to the space* $D(D_1 - K_0)$, *and put* $V_0 = -D_1\Psi + K_0\Psi + \frac{1}{2}\Gamma_1(\Psi,\Psi)$. *Then the equality in (7.44) holds.*

Remark 7.7. Suppose that the Feller propagator $\{P(s,t) : 0 \leqslant s \leqslant t \leqslant T\}$ has an integral kernel $p_0(s,x;t,y)$, which is continuous on

$$\{(\tau, x; t, y) \in [0,T] \times E \times [0,T] \times E : 0 \leqslant \tau < t \leqslant T\}, \tag{7.45}$$

and hence, for $f : E \to [0,\infty)$ any bounded Borel measurable function, we have $P(\tau,t) f(x) = \int_E p_0(\tau,x;t,y) f(y) dm(y)$, where m is a non-negative Radon measure on E. Instead of $dm(y)$ we write dy most of the time. Define the measures $\mu_{\tau,x}^{t,y}$ on the σ-field generated by $X(\tau)$, $\tau < t \leqslant T$ by

$$\mu_{\tau,x}^{t,y}(A) = \mathbb{E}_{\tau,x}\left[p_0\left(s, X(s); t, y\right)\mathbf{1}_A\right],$$

where A belongs to the σ-field generated by $X(\rho)$, $\tau \leqslant \rho < s$, with $s \in (\tau, t)$ fixed. By the $\mathbb{P}_{\tau,x}$-martingale property of the process $s \mapsto p_0(s, X(s); t, y)$, $\tau \leqslant s < t$, the measure $\mu_{\tau,x}^{t,y}$ is well defined and can be extended to the σ-field generated by $X(s)$, $\tau \leqslant s < t$. The latter can be done via the classical Kolmogorov extension theorem: see §3.1.7. The integral kernel of the operator $\exp(s(D_1 - K_0 - V))$ is given by the Feynman-Kac formula:

$$\begin{aligned} &\exp\left(s\left(D_1 - K_0 - V\right)\right)(x;t,y) \\ &= \int \exp\left(-\int_t^{s+t} V(\rho, X(\rho))d\rho\right) d\mu_{t,x}^{s+t,y} \\ &= \lim_{t'\uparrow t} \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t'} V(\rho, X(\rho))d\rho\right) p\left(s+t', X\left(s+t'\right); s+t, y\right)\right]. \end{aligned} \tag{7.46}$$

The following argument shows this claim. Let $f \geqslant 0$ be a bounded Borel measurable function. Then we have for $t' < t$:

$$\begin{aligned} &\int_E \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t'} V(\rho, X(\rho))d\rho\right) p(s+t', X(s+t'); s+t, y)\right] f(y)\, dy \\ &= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t'} V(\rho, X(\rho))d\rho\right) \int_E p(s+t', X(s+t'); s+t, y) f(y) dy\right] \end{aligned}$$

$$= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t'} V(\rho, X(\rho))d\rho\right)\mathbb{E}_{s+t',X(s+t')}\left[f\left(X(s+t)\right)\right]\right]$$

(Markov property)

$$= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t'} V(\rho, X(\rho))d\rho\right) f\left(X(s+t)\right)\right]. \tag{7.47}$$

By taking limits as $t' \uparrow t$ in the first and last term in (7.47) our claim follows. Under appropriate conditions on V, the integral kernel of the operator

$$\exp\left(s\left(D_1 - K_0 - V\right)\right)$$

is again continuous on the space mentioned in (7.45). In fact if the function V is bounded we obtain by expanding the exponential and using the martingale property of the process $\rho \mapsto p\left(\rho, X(\rho); t, y\right)$, $0 \leqslant \rho < t$:

$$\begin{aligned}
&\exp\left(s\left(D_1 - K_0 - V\right)\right)(x; t, y) && (7.48)\\
&= p\left(t, x; s+t, y\right) + \sum_{k=1}^{\infty}(-1)^k \int_{t<\rho_1<\cdots<\rho_k<s+t}\int\int_E \cdots \int_E \\
&\quad \prod_{j=1}^{k} p\left(\rho_{j-1}, y_{j-1}; \rho_j, y_j\right) V\left(\rho_j, y_j\right) p\left(\rho_k, y_k; s+t, y\right)\, dy_k \dots dy_1 d\rho_k \dots d\rho_1.
\end{aligned}$$

In (7.48) we wrote $\rho_0 = t$ and $y_0 = x$. Suppose that the function $(s, t, x) \mapsto p\left(t, x; s+t, y\right)$, $0 \leqslant t < s+t < T$, x, $y \in E$, is continuous. From the representation in (7.48) we see that each term in the right-hand side of (7.48) is continuous. Uniform convergence on compact subsets then yields the continuity of the left-hand side in (7.48). The proof of the following theorem is left as an exercise for the reader.

Theorem 7.4. *Suppose that function $(s, t, x) \mapsto p\left(t, x; s+t, y\right)$, $0 \leqslant t < s+t < T$, x, $y \in E$, is continuous. In addition, suppose that*

$$\lim_{h\downarrow 0}\sup_{0\leqslant\tau\leqslant T-h}\sup_{x,y\in E}\int_\tau^t p\left(\tau, x; \rho, z\right)\left|V(\rho, z)\right| p\left(\rho, z; \tau+h, y\right)\, dz = 0. \tag{7.49}$$

Then the integral kernel $(s, t, x, y) \mapsto \exp\left(s\left(D_1 - K_0 - V\right)\right)(x; t, y)$, $s > 0$, $t \in [0, T]$, x, $y \in E$, is continuous.

Details for time-independent functions V and time-homogenous Markov processes on second countable locally compact spaces can be found in e.g. [Demuth and van Casteren (2000)].

Remark 7.8. Let $\Phi(t,x)$ be a bounded continuous function which satisfies the following conditions:

(1) The function $(t,x) \mapsto V(t,x)\Phi(t,x)$ is continuous;
(2) The function $(t,x) \mapsto K_0\Phi(t,x) = K_0(t)\Phi(t,\cdot)(x)$ is continuous;
(3) The function $t \mapsto \Phi(t,x)$ is continuously differentiable for every $x \in E$.

Then the process

$$\begin{aligned} s \mapsto \exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right)\, d\tau\right) \Phi\left(s+t, X(s+t)\right) - \Phi\left(t, X(t)\right) \\ + \int_t^{s+t} \exp\left(-\int_t^{\tau+t} V\left(\rho, X(\rho)\right)\, d\rho\right) \\ \times \left(K_0(\tau) + V\left(\tau, X(\tau)\right) - \frac{\partial}{\partial\tau}\right) \Phi\left(\tau, X(\tau)\right)\, d\tau \end{aligned} \tag{7.50}$$

is a $\mathbb{P}_{t,x}$-martingale relative to the filtration $\left\{\mathcal{F}^t_{s+t} : 0 \leqslant s \leqslant T-t\right\}$. This assertion is a consequence of the fact that the operator family $-K_0(\tau)$, $0 \leqslant \tau \leqslant T$, generates the Markov process in (7.55), and the fact that the operator $D_1 - K_0 - V$ extends to a generator of the semigroup defined by (7.42).

Proof. [Proof of Proposition 7.1.] (i) Let s_1 and s_2 be positive real numbers, and let Φ be a non-negative Borel measurable function defined on $[0,\infty) \times E$. Then we have:

$$\begin{aligned} &\left[\exp\left(s_1\left(D_1 - K_0 - V\right)\right)\exp\left(s_2\left(D_1 - K_0 - V\right)\right)\Phi\right](t,x) \\ &= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s_1+t} V\left(\tau, X(\tau)\right) d\tau\right)\right. \\ &\qquad \left.\left\{\exp\left(s_2\left(D_1 - K_0 - V\right)\right)\Phi\left(s_1+t, X\left(s_1+t\right)\right)\right\}\right] \\ &= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s_1+t} V\left(\tau, X(\tau)\right) d\tau\right)\mathbb{E}_{s_1+t, X(s_1+t)}\right. \\ &\qquad \left.\left\{\exp\left(-\int_{s_1+t}^{s_2+s_1+t} V\left(\tau, X(\tau)\right) d\tau\right)\Phi\left(s_2+s_1+t, X\left(s_2+s_1+t\right)\right)\right\}\right] \end{aligned}$$

(Markov property)

$$\begin{aligned} &= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s_1+t} V\left(\tau, X(\tau)\right) d\tau\right)\right. \\ &\qquad \left.\exp\left(-\int_{s_1+t}^{s_2+s_1+t} V\left(\tau, X(\tau)\right) d\tau\right)\Phi\left(X(s_2+s_1+t), s_2+s_1+t\right)\right] \end{aligned}$$

$$= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s_2+s_1+t} V\left(\tau, X(\tau)\right) d\tau\right) \Phi\left(X(s_2+s_1+t), s_2+s_1+t\right)\right]$$
$$= \left[\exp\left((s_1+s_2)\left(D_1 - K_0 - V\right)\right)\Phi\right](t,x). \tag{7.51}$$

Next we show that the generator of the semigroup given by the formula in (7.42) extends the operator $D_1 - K_0 - V$. More precisely we will prove that that

$$\lim_{s\downarrow 0} \frac{1}{s}\left\{\mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right) \Phi\left(X(s+t), s+t\right)\right] - \Phi(t,x)\right\}$$
$$= \left(\frac{\partial}{\partial t} - K_0(t) - V(t,x)\right)\Phi(t,x). \tag{7.52}$$

Here Φ is a function which belongs to the intersections of the domains of D_1 (i.e. the time derivative), K_0 (i.e. for each $t \in [0,T]$ the function $x \mapsto \Phi(t,x)$ belongs to the domain of $K_0(t)$, and the function $(t,x) \mapsto K_0(t)\Phi(t,x)$ is continuous), and the function $(t,x) \mapsto V(t,x)\Phi(t,x)$ is continuous as well. The expression

$$\mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right) \Phi\left(X(s+t), s+t\right)\right] - \Phi(t,x) \tag{7.53}$$

can be rewritten as follows:

$$\mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right) \Phi\left(X(s+t), s+t\right)\right] - \Phi(t,x)$$
$$= \mathbb{E}_{t,x}\left[\left(\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right) - 1\right) \Phi\left(X(s+t), s+t\right)\right]$$
$$+ \mathbb{E}_{t,x}\left[\Phi\left(X(s+t), s+t\right)\right] - \Phi(t,x). \tag{7.54}$$

Since the operator family $-K_0(\tau)$, $0 \leqslant \tau \leqslant T$, generates the Markov process (see (7.6))

$$\left\{\left(\Omega, \mathcal{F}_T^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : T \geqslant t \geqslant 0\right), \left(\vee_t : T \geqslant t \geqslant 0\right), \left(E, \mathcal{E}\right)\right\} \tag{7.55}$$

we know that

$$\mathbb{E}_{t,x}\left[\Phi\left(X(s+t), s+t\right)\right]$$
$$= \Phi(t,x) + \int_0^s \mathbb{E}_{t,x}\left[\left(D_1 - K_0\right)\Phi\left(\tau+t, X(\tau+t)\right)\right] d\tau. \tag{7.56}$$

From (7.54) and (7.56) we infer the equality in (7.52).

Next we prove assertion (ii):

$$\left[\exp\left(s\left(D_1 - K_0 - V\right)\right)\exp\left(-S_L\right)\right](t,x)$$

$$
\begin{aligned}
&= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right) \exp\left(-S_L\left(s+t, X(s+t)\right)\right)\right] \\
&= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right)\right. \\
&\qquad \left.\mathbb{E}_{s+t,X(s+t)}\left\{\exp\left(-\int_{s+t}^{T} V\left(\tau, X(\tau)\right) d\tau\right) \chi\left(T, X(T)\right)\right\}\right]
\end{aligned}
$$

(Markov property)

$$
\begin{aligned}
&= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right)\right. \\
&\qquad \left.\exp\left(-\int_{s+t}^{T} V\left(\tau, X(\tau)\right) d\tau\right) \chi\left(T, X(T)\right)\right] \\
&= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{s+t} V\left(\tau, X(\tau)\right) d\tau\right)\right. \\
&\qquad \left.\exp\left(-\int_{s+t}^{T} V\left(\tau, X(\tau)\right) d\tau\right) \chi\left(T, X(T)\right)\right] \\
&= \mathbb{E}_{t,x}\left[\exp\left(-\int_t^{T} V\left(\tau, X(\tau)\right) d\tau\right) \chi\left(T, X(T)\right)\right] \\
&= \exp\left(-S_L(t,x)\right). \qquad (7.57)
\end{aligned}
$$

This proves assertion (ii).

(iii) From (7.1) and the proof of assertion (ii) of Proposition 7.1 it follows that

$$
\begin{aligned}
&- D_1\Psi + K_0\Psi + \frac{1}{2}\Gamma_1\left(\Psi, \Psi\right) = e^{\Psi}\left(D_1 - K_0\right) e^{-\Psi} \\
&= e^{\psi}\left(D_1 - K_0 - V_0\right) e^{-\Psi} + V_0 \\
&= e^{\Psi} \lim_{s\downarrow s} \frac{1}{s}\left(\exp\left(s\left(D_1 - K_0 - V_0\right)\right) - I\right) e^{-\Psi} + V_0 \\
&= V_0, \qquad (7.58)
\end{aligned}
$$

where we used the invariance $\exp\left(s\left(D_1 - K_0 - V_0\right)\right) e^{-\Psi}(\tau, x) = e^{-\Psi}(\tau, x)$, $0 < s < t - \tau$. This proves assertion (iii).

(iv) We write

$$
\begin{aligned}
&\exp\left(s\left(D_1 - K_0 - V_0\right)\right) e^{-\Psi} - e^{-\Psi} \\
&= \int_0^s \left(\rho\left(D_1 - K_0 - V_0\right)\right)\left(D_1 - K_0 - V_0\right) e^{-\Psi}\, d\rho
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^s \exp\left(\rho\left(D_1 - K_0 - V_0\right)\right) e^{-\Psi} e^{\Psi} \left(D_1 - K_0 - V_0\right) e^{-\Psi}\, d\rho \\
&= \int_0^s \exp\left(\rho\left(D_1 - K_0 - V_0\right)\right) e^{-\Psi} \left(e^{\Psi}\left(D_1 - K_0\right) e^{-\Psi} - V_0\right)\, d\rho \\
&= \int_0^s \exp\left(\rho\left(D_1 - K_0 - V_0\right)\right) e^{-\Psi} \left(V_0 - V_0\right)\, d\rho = 0. \qquad (7.59)
\end{aligned}
$$

The equality in (7.44) is a consequence of (7.59).

Altogether this shows assertion (iv) and completes the proof of Proposition 7.1. □

Proof. [Proof of Theorem 7.3.] (i) The first and the final inequality in (i) follow from the non-linear Feynman-Kac formula. For $\Phi \in D\left(D_1 - K_0\right)$ we have with $V_\Phi = -\dot{\Phi} + K_0\Phi + \frac{1}{2}\Gamma_1\left(\Phi, \Phi\right)$:

$$
\Phi(t,x) = -\log \mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V_\Phi\left(\tau, X(\tau)\right) d\tau\right) \Phi\left(T, X(T)\right)\right].
$$

The second inequality of (i) is a consequence of Jensen inequality, and should be compared with the arguments in [Zambrini (1998a)], who used ideas from Fleming and Soner: see Chapter VI in [Fleming and Soner (1993)]. Another relatively recent source of information is Chapter 8 in [Bressan and Piccoli (2007)]. The reader is also referred to [Sheu (1984)] and to [Van Casteren (2001)]. The inequality we have in mind is the following one:

$$
-\log \mathbb{E}_{t,x}^{M_{v,t}}\left[\exp\left(-\varphi\right)\right] \leqslant \mathbb{E}_{t,x}^{M_{v,t}}\left[\varphi\right], \qquad (7.60)
$$

with equality only if φ is constant $\mathbb{P}_{t,x}$-almost surely. We apply (7.60) to the random variable $\varphi = \varphi_v$, given by

$$
\varphi_v = -\int_t^T \left[\frac{1}{2}\Gamma_1\left(v, v\right) + V\right]\left(\tau, X(\tau)\right) d\tau - M_{v,t}\left(T\right) - \log\chi\left(X\left(T\right)\right). \qquad (7.61)
$$

We also notice that the following processes are $\mathbb{P}_{t,x}$-martingales on the interval $[t, T]$:

$$
\exp\left(-\frac{1}{2}\left\langle M_{v,t}\right\rangle(s) - M_{v,t}(s)\right) \quad \text{and} \qquad (7.62)
$$

$$
\exp\left(-\frac{1}{2}\left\langle M_{v,t}\right\rangle(s) - M_{v,t}(s)\right)\left(\left\langle M_{v,t}\right\rangle(s) + M_{v,t}(s)\right). \qquad (7.63)
$$

By the Jensen inequality we have

$$
\mathbb{E}_{t,x}^{M_{v,t}}\left[\frac{1}{2}\left\langle M_{v,t}\right\rangle(T) + \int_t^T V\left(\tau, X(\tau)\right) d\tau - \log\chi\left(T, X(T)\right)\right]
$$

(the process in (7.63) is a $\mathbb{P}_{t,x}^{M_{v,t}}$-martingale)

$$\mathbb{E}_{t,x}^{M_{v,t}}\left[-\frac{1}{2}\langle M_{v,t}\rangle(T)-M_{v,t}(T)+\int_t^T V(\tau,X(\tau))\,d\tau-\log\chi(T,X(T))\right]$$

(here we apply Jensen inequality)

$$\geqslant -\log\mathbb{E}_{t,x}^{M_{v,t}}\left[e^{\frac{1}{2}\langle M_{v,t}\rangle(T)+M_{v,t}(T)-\int_t^T V(\tau,X(\tau))d\tau+\log\chi(T,X(T))}\right]$$

(definition of the probability measure $\mathbb{E}_{t,x}^{M_{v,t}}$)

$$= -\log\mathbb{E}_{t,x}\left[\exp\left(-\int_t^T V(\tau,X(\tau))\,d\tau+\log\chi(T,X(T))\right)\right].$$

(ii) The assertion in (ii) immediately follows from (i).

(iii) Let (t,x) belong to $(t_0,T]\times E$, and let φ be as in (7.34). Then we have

$$\begin{aligned}
&\left[-\dot{\varphi}+K_0\varphi-V\right](t,x)+\frac{1}{2}\Gamma_1(\varphi,\varphi)(t,x)\\
&= e^{\varphi(t,x)}\left[(D_1-K_0-V)\,e^{-\varphi}\right](t,x)\\
&= \exp(\varphi(t,x))\lim_{s\downarrow 0}\frac{1}{s}\left[(\exp(sD_1-sK_0-sV)-I)\exp(-\varphi)\right](t,x)\\
&= e^{\varphi(t,x)}\liminf_{s\downarrow 0}\frac{1}{s}\Big(\left[\exp(sD_1-sK_0-sV)\left\{e^{S_L-\varphi}\right\}e^{-S_L}\right](t,x)\\
&\qquad - e^{S_L(t,x)-\varphi(t,x)}e^{-S_L(t,x)}\Big)\\
&= e^{\varphi(t,x)}\liminf_{s\downarrow 0}\frac{1}{s}\Bigg(\left[\exp(sD_1-sK_0-sV)\left\{e^{S_L-\varphi}\right\}e^{-S_L}\right](t,x)\\
&\qquad -\left\{\sup_{(\sigma,y)\in[t,T]\times E} e^{S_L(s,y)-\varphi(s,y)}\right\}e^{-S_L(t,x)}\Bigg)\\
&\leqslant e^{\varphi(t,x)}\liminf_{s\downarrow 0}\frac{1}{s}\\
&\quad\Bigg(\left[\exp(sD_1-sK_0-sV)\left\{\sup_{(\sigma,y)\in[t,T]\times E}\left\{e^{S_L-\varphi}\right\}(\sigma,y)\right\}e^{-S_L}\right](t,x)\\
&\qquad -\left\{\sup_{(\sigma,y)\in[t,T]\times E} e^{S_L(\sigma,y)-\varphi(\sigma,y)}\right\}e^{-S_L(t,x)}\Bigg)
\end{aligned}$$

$$\leqslant \sup\left\{\left\{e^{S_L-\varphi}\right\}(\sigma,y) : (\sigma,y)\in[t,T]\times E\right\}\exp\left(\varphi(t,x)\right)$$
$$\liminf_{s\downarrow 0}\frac{1}{s}\left(\left[\exp\left(sD_1-sK_0-sV\right)e^{-S_L}\right](t,x)-e^{-S_L(t,x)}\right)$$
$$= e^{S_L(t,x)-\varphi(t,x)}e^{\varphi(t,x)}\cdot 0 = 0. \tag{7.64}$$

The latter equality follows because, for $s > 0$ the equality

$$\left[\exp\left(sD_1-sK_0-sV\right)\exp\left(-S_L\right)\right](t,x) = \exp\left(-S_L(t,x)\right)$$

is valid: see Proposition 7.1, assertion (ii).

The reverse inequality (7.35) follows in a similar manner.

(iv) In view of assertion (i) we only have to prove that the expression in (7.37) is less than equal to the one in (7.36). To this end we consider $(v\in D\left(D_1-K_0\right))$

$$\mathbb{E}_{t,x}^{M_{v,t}}\left[\frac{1}{2}\int_t^T\Gamma_1\left(v,v\right)\left(\tau,X(\tau)\right)d\tau+\int_t^T V\left(\tau,X(\tau)\right)d\tau-\log\chi\left(T,X(T)\right)\right]$$
$$= \mathbb{E}_{t,x}^{M_{v,t}}\left[\frac{1}{2}\left\langle M_{v,t}\right\rangle(T)+\int_t^T V\left(\tau,X(\tau)\right)d\tau-\log\chi\left(T,X(T)\right)\right]$$

(the process in (7.63) is a martingale)

$$= \mathbb{E}_{t,x}^{M_{v,t}}\left[-\frac{1}{2}\left\langle M_{v,t}\right\rangle(T)-M_{v,t}\left(T\right)+\int_t^T V\left(\tau,X(\tau)\right)d\tau-\log\chi\left(T,X(T)\right)\right]$$

(definition of the martingale $s\mapsto M_{v,t}(s)$)

$$= \mathbb{E}_{t,x}^{M_{v,t}}\left[-\frac{1}{2}\int_t^T\Gamma_1\left(v,v\right)\left(\tau,X(\tau)\right)d\tau-v\left(T,X(T)\right)+v\left(t,X(t)\right)\right.$$
$$\left.-\int_t^T\left(-D_1+K_0\right)v\left(\tau,X(\tau)\right)d\tau+\int_t^T V\left(\tau,X(\tau)\right)d\tau-\log\chi\left(T,X(T)\right)\right]$$
$$= \mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T\left\{\left(D_1-K_0\right)v\left(\tau,X(\tau)\right)-\frac{1}{2}\Gamma_1\left(v,v\right)\left(\tau,X(\tau)\right)\right\}d\tau\right.$$
$$\left.+\int_t^T V\left(\tau,X(\tau)\right)d\tau+v\left(t,X(t)\right)-v\left(T,X(T)\right)-\log\chi\left(T,X(T)\right)\right]$$

($(t,X(t))=(t,x)$, $\mathbb{P}_{t,x}$-almost surely)

$$= v(t,x)+\mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T\left\{\left(D_1-K_0\right)v-\frac{1}{2}\Gamma_1\left(v,v\right)\left(\tau,X(\tau)\right)\right\}d\tau\right.$$

$$
\begin{aligned}
&\quad + \int_t^T V\left(\tau, X(\tau)\right) d\tau - v\left(T, X(T)\right) - \log \chi\left(T, X(T)\right)\Bigg] \\
&= v(t,x) \\
&\quad + \mathbb{E}_{t,x}^{M_{v,t}}\left[\int_t^T \exp\left(v\left(\tau, X(\tau)\right)\right)\left[\left(-D_1 + K_0 + V\right)\exp(-v)\right]\left(\tau, X(\tau)\right) d\tau \right.\\
&\quad \left. - v\left(T, X(T)\right) - \log \chi\left(T, X(T)\right)\right] \\
&= v(t,x) + \mathbb{E}_{t,x}\left[\exp\left(-\frac{1}{2}\left\langle M_{v,t}\right\rangle(T) - M_{v,t}(T) + \int_t^T V\left(\tau, X(\tau)\right) d\tau\right)\right. \\
&\quad \exp\left(-\int_t^T V\left(\tau, X(\tau)\right) d\tau\right) \\
&\quad \int_t^T \exp\left(v\left(\tau, X(\tau)\right)\right)\left[\left(-D_1 + K_0 + V\right)\exp(-v)\right]\left(\tau, X(\tau)\right) d\tau \\
&\quad \left. - v\left(T, X(T)\right) - \log \chi\left(T, X(T)\right)\right]
\end{aligned}
$$

(apply the equality in (7.2) with $f = \mathbf{1}$)

$$
\begin{aligned}
&= v(t,x) \\
&+ \mathbb{E}_{t,x}\left[\exp\left(\int_t^T \exp\left(v\left(\tau, X(\tau)\right)\right)\left[\left(-D_1 + K_0 + V\right)\exp\left(-v\right)\right]\left(\tau, X(\tau)\right) d\tau\right)\right. \\
&\quad \exp\left(v\left(t, X(t)\right) - v\left(T, X(T)\right)\right)\exp\left(-\int_t^T V\left(\tau, X(\tau)\right) d\tau\right) \\
&\quad \int_t^T \exp\left(v\left(\tau, X(\tau)\right)\right)\left[\left(-D_1 + K_0 + V\right)\exp(-v)\right]\left(\tau, X(\tau)\right) d\tau \\
&\quad \left. - v\left(T, X(T)\right) - \log \chi\left(T, X(T)\right)\right]. \qquad (7.65)
\end{aligned}
$$

Choose $w \in D\left(D_1 - K_0\right)$ and define for $s > 0$ the function v_s by

$$
\exp\left(-v_s\right) = \frac{1}{s}\int_0^s \exp\left(\sigma\left(D_1 - K_0 - V\right)\right)\exp(-w)\, d\sigma.
$$

Then

$$
\exp(v_s)(-D_1 + K_0 + V)\exp(-v_s) = \frac{(I - \exp(s(D_1 - K_0 - V)))\exp(-w)}{\int_0^s \exp(\sigma(D_1 - K_0 - V))\exp(-w) d\sigma}.
$$

So from (7.65) we obtain for w in the domain of $D_1 - K_0$ and $s > 0$ the equality

$$\mathbb{E}_{t,x}^{M_{v_s,t}}\left[\frac{1}{2}\int_t^T \Gamma_1\left(v_s,v_s\right)(\tau,X(\tau))\,d\tau + \int_t^T V\left(\tau,X(\tau)\right)d\tau - \log\chi\left(T,X(T)\right)\right]$$
$$= v_s(t,x)$$
$$+ \mathbb{E}_{t,x}\left[\exp\left(\int_t^T \frac{\left[\left(I-\exp\left(s\left(D_1-K_0-V\right)\right)\right)\exp(-w)\right](\tau,X(\tau))}{\int_0^s\left[\exp\left(\sigma\left(D_1-K_0-V\right)\right)\exp(-w)\right](\tau,X(\tau))\,d\sigma}d\tau\right)\right.$$
$$\exp\left(v_s\left(t,X(t)\right)-v_s\left(T,X(T)\right)\right)\exp\left(-\int_t^T V\left(\tau,X(\tau)\right)d\tau\right)$$
$$\int_t^T \frac{\left[\left(I-\exp\left(s\left(D_1-K_0-V\right)\right)\right)\exp(-w)\right](\tau,X(\tau))}{\int_0^s\left[\exp\left(\sigma\left(D_1-K_0-V\right)\right)\exp(-w)\right](\tau,X(\tau))\,d\sigma}d\tau$$
$$\left.- v_s\left(T,X(T)\right) - \log\chi\left(T,X(T)\right)\right]. \tag{7.66}$$

Upon letting $w \in D\left(D_1 - K_0\right)$ tend to the function S_L in an appropriate manner, we obtain by invoking Proposition 7.1 the inequality

$$\inf\left\{\mathbb{E}_{t,x}^{M_{v,t}}\left[\frac{1}{2}\int_t^T \Gamma_1\left(\tau,X(\tau)\right)d\tau + \int_t^T V\left(\tau,X(\tau)\right)d\tau - \log\chi\left(T,X(T)\right)\right]:\right.$$
$$\left. v \in D\left(D_1-K_0\right)\right\} \leqslant S_L(t,x).$$

This proves assertion (iv). The "appropriate manner" should be such that $w_n \to S_L$ implies that

$$\mathcal{T}_\beta\text{-}\lim_{n\to\infty} e^{s(D_1-K_0-V)}e^{-w_n} = e^{s(D_1-K_0-V)}e^{-S_L} = e^{-S_L}. \tag{7.67}$$

In order that this procedure works, the semigroup $\left\{e^{s\left(D_1-\left(K_0+V\right)\right)} : s \geqslant 0\right\}$ should be continuous for the strict topology. This is true provided the unperturbed semigroup $\left\{e^{s(D_1-K_0)} : s \geqslant 0\right\}$ is continuous for the strict topology, and the potential function satisfies a Myadera type boundedness condition, as explained in Definition 7.2 and the corresponding Khas'minski lemma 7.1.

(v). Let S be a viscosity solution to (7.12). Here we use a martingale approach together with the idea of germs of a function. We will prove the following inequalities:

$$S(t,x) \leqslant \sup\left\{\varphi_1(t,x) : V_{\varphi_1} \leqslant V,\quad \varphi_1(T,\cdot) \leqslant S_L(T,\cdot)\right\} \tag{7.68}$$

$$\leqslant \inf \{\varphi_2(t,x) : V_{\varphi_2} \geqslant V, \quad \varphi_2(T,\cdot) \geqslant S_L(T,\cdot)\} \leqslant S(t,x), \quad (7.69)$$

where $V_\varphi = -\dot{\varphi} + K_0\varphi + \frac{1}{2}\Gamma_1(\varphi,\varphi)$. In view of assertion (i) in Theorem 7.3 we then infer $S = S_L$. Fix $(t,x) \in (t_0,T] \times E$. Let $\varphi_1 \in D(D_1 - K_0)$ be such that

$$(S-\varphi_1)(t,x) = \sup\{(S-\varphi_1)(s,y) : y \in E,\ T \geqslant s \geqslant t\}.$$

We notice that the processes $M_{\varphi,t}$ and $M_{S_L,t}$, defined by respectively

$$M_{\varphi,t}(s) = \exp\left(-\int_t^s V_\varphi(\tau, X(\tau))\,d\tau - \varphi(s,X(s)) + \varphi(t,X(t))\right), \text{ and}$$

$$M_{S_L,t}(s) = \exp\left(-\int_t^s V(\tau,X(\tau))\ d\tau - S_L(s,X(s)) + S_L(t,X(t))\right),$$

$t \leqslant s \leqslant T$, are $\mathbb{P}_{t,x}$-martingales. The latter assertion follows from the Markov property together with the Feynman-Kac formula: see (7.43), which is also true for V_φ instead of V and φ replacing S_L. Let $\mathbb{P}_{t,x}^{M_{\varphi,t}}$ denote the probability measure defined by $\mathbb{P}_{t,x}^{M_{\varphi,t}}(A) = \mathbb{E}_{t,x}[M_{\varphi,t}(s_2)1_A]$, $s_2 \geqslant s_1$, where A is $\mathcal{F}_{s_1}^t$-measurable. Since S is a viscosity sub-solution we see that $V_{\varphi_1}(t,x) \leqslant V(t,x)$. Fix $\varepsilon > 0$ and choose $\delta > 0$ in such a way that, for some neighborhood U of x in E, the inequality $V_{\varphi_1}(s,y) \leqslant V(s,y) + \frac{1}{2}\varepsilon$ is valid for $(s,y) \in U \times [t,t+\delta]$. Here we use the continuity of $V(s,y)$ in $y = x$ and its right continuity in $s = t$. Then we choose a family of germs of "smooth" functions $(U_\alpha, \varphi_\alpha)$, $\alpha \in \mathcal{A}$, with the following properties:

(a) $\bigcup U_\alpha \supseteq [t,T] \times E$, i.e. the family U_α, $\alpha \in \mathcal{A}$, forms an open cover of the set $[t,T] \times E$;
(b) For every α, $\beta \in \mathcal{A}$, $\varphi_\alpha = \varphi_\beta$ on $U_\alpha \bigcap U_\beta$;
(c) For every $\alpha \in \mathcal{A}$ there exists $(t_\alpha, x_\alpha) \in U_\alpha$ such that $(S-\varphi_\alpha)(s,y) \leqslant (S-\varphi_\alpha)(t_\alpha,x_\alpha)$, for $(s,y) \in U_\alpha$ and $s_\alpha \leqslant s$;
(d) For every $\alpha \in \mathcal{A}$, the inequality $V_{\varphi_\alpha} \leqslant V + \frac{1}{2}\varepsilon$ is valid on U_α;
(e) If (t,x) belongs to U_α, then $(S-\varphi_\alpha)(t,x) \leqslant 0$;
(f) If (T,y) belongs to U_α, then

$$\varphi_\alpha(T,y) \leqslant S(T,y) + \frac{1}{2}\varepsilon(T-t) = S_L(T,y) + \frac{1}{2}\varepsilon(T-t).$$

Since S is a viscosity sub-solution property (d) is in fact a consequence of (c); we will need (d). Then we define the function $\psi_1 : [t,T] \times E \to \mathbb{R}$ as follows $\psi_1(s,y) = \varphi_\alpha(s,y)$, for $(s,y) \in U_\alpha$. Then, on U_α, $V_{\psi_1} = V_{\varphi_\alpha} \leqslant V + \frac{1}{2}\varepsilon$. We write

$$V_1 = V_{\psi_1} = -D_1\psi_1 + K_0\psi_1 + \frac{1}{2}\Gamma_1(\psi_1,\psi_1).$$

By assertion (iii) and (iv) of Proposition 7.1 we have

$$\begin{aligned}\Psi_1^{\varepsilon}(s,y) &:= \psi_1(s,y) - \frac{1}{2}\varepsilon(T-s) - \frac{1}{2}\varepsilon(T-t) \qquad (7.70)\\ &= -\log \mathbb{E}_{s,y}\left[\exp\left(-\int_s^T \left(V_1\big(\tau, X(\tau)\big) - \frac{1}{2}\varepsilon\right) d\tau\right)\right.\\ &\qquad \left.\times \exp\left(-\left(\psi_1\big(T, X(T)\big) - \frac{1}{2}\varepsilon(T-t)\right)\right)\right]\\ &= -\log \mathbb{E}_{s,y}\left[\exp\left(-\int_t^T V_{\Psi_1^{\varepsilon}}\big(\tau, X(\tau)\big) d\tau - \Psi_1^{\varepsilon}\big(T, X(T)\big)\right)\right].\end{aligned}$$

Then

$$\begin{aligned}\Psi_1^{\varepsilon}(s,y) &\leqslant -\log \mathbb{E}_{s,y}\left[\exp\left(-\int_s^T V\big(\tau, X(\tau)\big) d\tau - S_L\big(T, X(T)\big)\right)\right]\\ &= S_L(s,y),\end{aligned}$$

and hence $\psi_1(t,x) \leqslant S_L(t,x) + \varepsilon(T-t)$. By construction we also have $S(t,x) \leqslant \psi_1(t,x)$. Consequently $S(t,x) \leqslant S_L(t,x) + \varepsilon(T-t)$. Since $\varepsilon > 0$ is arbitrary we see $S(t,x) \leqslant S_L(t,x)$. In fact, since $V_{\Psi_1^{\varepsilon}} \leqslant V$, and since $\Psi_1^{\varepsilon}(T,y) \leqslant S_L(T,y)$, we see that

$$S(t,x) \leqslant \sup\left\{\varphi_1(t,x) : V_{\varphi_1} \leqslant V, \quad \varphi_1(T,\cdot) \leqslant S_L(T,\cdot)\right\}.$$

A similar argument shows the inequality

$$S(t,x) \geqslant \inf\left\{\varphi_2(t,x) : V_{\varphi_2} \geqslant V, \quad \varphi_2(T,\cdot) \geqslant S_L(T,\cdot)\right\}.$$

To be precise, again we fix $\varepsilon > 0$, and let $\varphi_2 \in D\left(D_1 - K_0\right)$ be a function such that $\left(S - \varphi_2\right)(t,x) = \inf\left\{\left(S - \varphi_2\right)(s,y) : (s,y) \in [t,T] \times E\right\}$. We choose $\delta > 0$ and a neighborhood U of x in such a way that $V_{\varphi_2}(s,y) \geqslant V(s,y) - \frac{1}{2}\varepsilon$ for $(s,y) \in U \times [t, t+\delta]$. Then we choose a family of germs of "smooth" functions $(U_\alpha, \varphi_\alpha)$, $\alpha \in \mathcal{A}$, with the following properties:

(a) $\bigcup U_\alpha \supseteq [t,T] \times E$, i.e. the family U_α forms an open cover of the set $[t,T] \times E$;
(b) α, $\beta \in \mathcal{A}$ implies $\varphi_\alpha = \varphi_\beta$ on $U_\alpha \bigcap U_\beta$;
(c) For every $\alpha \in \mathcal{A}$ there exists $(t_\alpha, x_\alpha) \in U_\alpha$ such that $\left(S - \varphi_\alpha\right)(s,y) \geqslant \left(S - \varphi_\alpha\right)(t_\alpha, x_\alpha)$, for $(s,y) \in U_\alpha$ and $s_\alpha \leqslant s$;
(d) For every $\alpha \in \mathcal{A}$, the inequality $V_{\varphi_\alpha} \geqslant V - \frac{1}{2}\varepsilon$ is valid on U_α;
(e) If (t,x) belongs to U_α, then $\left(S - \varphi_\alpha\right)(t,x) \geqslant 0$;

(f) If (T, y) belongs to U_α, then

$$\varphi_\alpha(T, y) \geqslant S(T, y) - \frac{1}{2}\varepsilon(T - t) = S_L(T, y) - \frac{1}{2}\varepsilon(T - t).$$

Since S is a viscosity super-solution property (d) is in fact a consequence of (c). Then we define the function $\psi_2 : [t, T] \times E \to \mathbb{R}$ as follows $\psi_2(s, y) = \varphi_\alpha(s, y)$, for $(s, y) \in U_\alpha$. In addition, we write as above

$$V_2 = V_{\psi_2} = -D_1\psi_2 + K_0\psi_1 + \frac{1}{2}\Gamma_1(\psi_2, \psi_2).$$

Then, on U_α, $V_{\psi_2} = V_{\varphi_\alpha} \leqslant V + \frac{1}{2}\varepsilon$. As above, assertions (iii) and (iv) of Proposition 7.1 imply

$$\Psi_2^\varepsilon(s, y) := \psi_2(s, y) + \frac{1}{2}\varepsilon(T - s) + \frac{1}{2}\varepsilon(T - t) \geqslant S_L(s, y). \tag{7.71}$$

By construction we have $S(t, x) \geqslant \psi_2(t, x)$, and hence

$$S_L(t, x) \leqslant \Psi_2^\varepsilon(t, x) \leqslant \psi_2(t, x) + \varepsilon(T - t) \leqslant S(t, x) + \varepsilon(T - t).$$

Since $\varepsilon > 0$ is arbitrary we infer $S_L(t, x) \leqslant S(t, x)$. In fact, since $V_{\Psi_2^\varepsilon} \geqslant V$, and since $\Psi_2^\varepsilon(T, y) \geqslant S_L(T, y)$, we see that

$$S(t, x) \geqslant \sup\left\{\varphi_1(t, x) : V_{\varphi_1} \leqslant V, \quad \varphi_1(T, \cdot) \leqslant S_L(T, \cdot)\right\}.$$

In the mean time we also proved that the 4 quantities in assertion (i) are equal.

This concludes the proof of Theorem 7.3. □

7.4 A stochastic Noether theorem

The following theorem may be called the *stochastic Noether theorem*: cf. [Zambrini (1998a)] Proposition 2.3 and Theorem 2.4. For a discussion and formulation of the classical (deterministic) Noether theorem, which in fact can be considered as the second constant of motion for a mechanical system, the reader is referred to [Thieullen and Zambrini (1997b)], pages 300–302, and [Thieullen and Zambrini (1997d)] page 423. In §7.4.1 we also give a short formulation of this theory.

Theorem 7.5. *Let T be a differentiable function which only depends on time. As above the operator D_1 stands for $D_1 = \dfrac{\partial}{\partial t}$. Suppose that the functions φ, w, and T satisfy the following identities.*

(a) $K_0 f \dfrac{dT}{dt} = K_0\Gamma_1(f,w) - \Gamma_1(K_0 f, w) - \Gamma_1\left(f, \dfrac{\partial w}{\partial t} - \varphi\right)$ *for all functions* $f \in D(K_0 - D_1)$ *for which* $\Gamma_1(K_0 f, w)$ *exists as well.*

(b) $\dfrac{\partial \varphi}{\partial t} - K_0\varphi = \Gamma_1(V,w) + \dfrac{\partial (TV)}{\partial t}$.

Put

$$\mathcal{H}(f) = \frac{\partial f}{\partial t} - \left(K_0 \dot{+} V\right) f, \quad \mathcal{N}(f) = \Gamma_1(f,w) + T\frac{\partial f}{\partial t} - \varphi f, \text{ and}$$

$$\mathcal{D}(f) = \frac{\partial f}{\partial t} - \Gamma_1(\sigma_L, f) - K_0 f. \tag{7.72}$$

Suppose that the function σ_L *satisfies:*

(c) $\left(\mathcal{D}\varepsilon - \dfrac{\partial V}{\partial t}\right) T = \Gamma_1\left(\dfrac{\partial \sigma_L}{\partial t} + \varepsilon, w\right)$,

where $\varepsilon = -K_0\sigma_L - \dfrac{1}{2}\Gamma_1(\sigma_L,\sigma_L) + V$.

Write $n := -\Gamma_1(\sigma_L, w) + \varepsilon T - \varphi$. *The following assertions hold true.*

(i) If $\mathcal{H}(f) = 0$, *then* $\mathcal{H}(\mathcal{N}(f)) = 0$ *as well. More generally:* $\mathcal{H}(\mathcal{N}_0(f)) = \mathcal{N}_0(\mathcal{H}(f))$ *for appropriately chosen functions* f*. So the operators* $\mathcal{H}$ *and* $\mathcal{N}_0$ *commute. For the definition of* $\mathcal{N}_0$ *see (7.74) below.*

(ii) $\mathcal{D}n = 0$.

(iii) The process $t \mapsto n(t, X(t))$ *is a martingale with respect to the probability measures*

$$A \mapsto \mathbb{E}_{t,x_0}\left[\exp\left(-M_{\sigma_L,t_0}(t) - \frac{1}{2}\int_{t_0}^{t} \Gamma_1(\sigma_L,\sigma_L)(s, X(s))ds\right) 1_A\right],$$

where as in (7.10) $M_{f,t_0}(t)$ *is given by*

$$M_{f,t_0}(t) = f(t, X(t)) - f(t_0, X(t_0)) + \int_{t_0}^{t} (K_0 - D_1) f(s, X(s))ds. \tag{7.73}$$

Remark 7.9. The operator $\mathcal{N}_0$ is defined by

$$\mathcal{N}_0(f) = \Gamma_1(f,w) + T\left(K_0 \dot{+} V\right) f - \varphi f. \tag{7.74}$$

The proof of assertion (i) shows that the operators $\mathcal{H}$ and $\mathcal{N}_0$ commute: $\mathcal{H}(\mathcal{N}_0 f) = \mathcal{N}_0(\mathcal{H} f)$, $f \in D(\mathcal{H}) \bigcap D(\mathcal{N}_0)$, $\mathcal{H} f \in D(\mathcal{N}_0)$, and $\mathcal{N}_0 f \in D(\mathcal{H})$.

The following proposition shows a situation where (c) is satisfied.

Proposition 7.2. *Suppose* S_L*, the minimal Lagrangian action, belongs to the domain of* $D_1 - K_0$*. Here* $D_1 = \dfrac{\partial}{\partial t}$*. Set* $\sigma_L = S_L$ *in Theorem 7.5. Then* (c) *is satisfied; more precisely,* $\mathcal{D}\varepsilon = D_1 V$ *and* $D_1\sigma_L + \varepsilon = 0$.

Proof. [Proof of Proposition 7.2.] Notice that

$$\varepsilon = -K_0 S_L - \frac{1}{2}\Gamma_1\left(S_L, S_L\right) + V = -D_1 S_L, \tag{7.75}$$

and hence

$$\begin{aligned}
\mathcal{D}\varepsilon &= -\mathcal{D}\left(D_1 S_L\right) = -\frac{\partial^2 S_L}{\partial t^2} + \Gamma_1\left(S_L, \frac{\partial S_L}{\partial t}\right) + K_0\left(\frac{\partial S_L}{\partial t}\right) \\
&= -\frac{\partial^2 S_L}{\partial t^2} + \frac{1}{2}\frac{\partial}{\partial t}\Gamma_1\left(S_L, S_L\right) + \frac{\partial}{\partial t}K_0\left(S_L\right) \\
&= \frac{\partial}{\partial t}\left(-\frac{\partial S_L}{\partial t} + \frac{1}{2}\Gamma_1\left(S_L, S_L\right) + K_0 S_L\right) = \frac{\partial V}{\partial t}.
\end{aligned} \tag{7.76}$$

Proposition 7.2 easily follows from (7.75) and (7.76). □

The equality in (7.77) below will be used in the proof Theorem 7.5.

Lemma 7.2. *For all appropriate functions f, w, T, and φ the following identity is true:*

$$\begin{aligned}
&-\frac{1}{2}\Gamma_1\left(f,f\right)\frac{\partial T}{\partial t} - \frac{1}{2}\Gamma_1\left(\Gamma_1\left(f,f\right),w\right) + \Gamma_1\left(f,\Gamma_1\left(f,w\right)\right) \\
&= \frac{1}{2}\left(K_0\left(f^2\right)\frac{\partial T}{\partial t} + \Gamma_1\left(K_0\left(f^2\right),w\right) - K_0\Gamma_1\left(f^2,w\right) + \Gamma_1\left(f^2, \frac{\partial w}{\partial t} - \varphi\right)\right) \\
&\quad - f\left(K_0\left(f\right)\frac{\partial T}{\partial t} + \Gamma_1\left(K_0\left(f\right),w\right) - K_0\Gamma_1\left(f,w\right) + \Gamma_1\left(f, \frac{\partial w}{\partial t} - \varphi\right)\right).
\end{aligned} \tag{7.77}$$

Proof. [Proof of Lemma 7.2.] The equality

$$-\frac{1}{2}\Gamma_1\left(f,f\right) = \frac{1}{2}K_0\left(f^2\right) - fK_0 f$$

together with

$$\Gamma_1\left(fK_0 f, w\right) = f\Gamma_1\left(K_0 f, w\right) + K_0 f\Gamma_1\left(f,w\right)$$

yields

$$\begin{aligned}
&-\frac{1}{2}\Gamma_1\left(f,f\right)\frac{\partial T}{\partial t} - \frac{1}{2}\Gamma_1\left(\Gamma_1\left(f,f\right),w\right) + \Gamma_1\left(f,\Gamma_1\left(f,w\right)\right) \\
&= \frac{1}{2}\left\{K_0\left(f^2\right)\frac{\partial T}{\partial t} + \Gamma_1\left(K_0\left(f^2\right),w\right) - K_0\Gamma_1\left(f^2,w\right)\right\} \\
&\quad - f\left\{K_0 f\frac{\partial T}{\partial t} + \Gamma_1\left(K_0 f, w\right) - K_0\Gamma_1\left(f,w\right)\right\} \\
&\quad + \frac{1}{2}K_0\Gamma_1\left(f^2,w\right) - fK_0\Gamma_1\left(f,w\right) - \left(K_0 f\right)\Gamma_1\left(f,w\right) + \Gamma_1\left(f,\Gamma_1\left(f,w\right)\right)
\end{aligned}$$

$\left(\frac{1}{2}K_0\Gamma_1\left(f^2,w\right) = K_0\left(f\Gamma_1\left(f,w\right)\right) = \left(K_0 f\right)\Gamma_1\left(f,w\right) - \Gamma_1\left(f,\Gamma_1\left(f,w\right)\right) + f K_0\Gamma_1\left(f,w\right)\right)$

$$
\begin{aligned}
&= \frac{1}{2}\left\{K_0\left(f^2\right)\frac{\partial T}{\partial t} + \Gamma_1\left(K_0\left(f^2\right),w\right) - K_0\Gamma_1\left(f^2,w\right)\right\} \\
&\quad - f\left\{K_0 f\frac{\partial T}{\partial t} + \Gamma_1\left(K_0 f,w\right) - K_0\Gamma_1\left(f,w\right)\right\}. \qquad (7.78)
\end{aligned}
$$

Since $\Gamma_1\left(f^2,\psi\right) = 2f\Gamma_1\left(f,\psi\right)$, equality (7.77) in Lemma 7.2 follows from (7.78). □

Proof. [Proof of Theorem 7.5.] (i). We calculate:

$$
\begin{aligned}
&\mathcal{H}\left(\mathcal{N}f\right) - \mathcal{N}\left(\mathcal{H}f\right) \\
&= \mathcal{H}\left(\Gamma_1(\omega,f)\right) + \mathcal{H}\left(T\frac{\partial f}{\partial t}\right) - \mathcal{H}\left(\varphi f\right) - \Gamma_1\left(\mathcal{H}f,w\right) - T\frac{\partial}{\partial t}\left(\mathcal{H}f\right) + \varphi\mathcal{H}f \\
&= \frac{\partial}{\partial t}\left(\Gamma_1\left(f,w\right)\right) - \left(K_0\dot{+}V\right)\Gamma_1\left(f,w\right) \\
&\quad \frac{\partial}{\partial t}\left(T\frac{\partial f}{\partial t} - \varphi f\right) - \left(K_0\dot{+}V\right)\left(T\frac{\partial f}{\partial t}\right) + \left(K_0\dot{+}V\right)\left(\varphi f\right) \\
&\quad - \Gamma_1\left(\frac{\partial f}{\partial t},w\right) + \Gamma_1\left(K_0 f,w\right) + \Gamma_1\left(Vf,w\right) \\
&\quad - T\left(\frac{\partial^2 f}{\partial t^2} - \frac{\partial}{\partial t}\left(\left(K_0\dot{+}V\right)f\right)\right) + \varphi\frac{\partial f}{\partial t} - \varphi\left(K_0\dot{+}V\right)f \\
&= \Gamma_1\left(f,\frac{\partial w}{\partial t} - \varphi\right) - K_0\Gamma_1\left(f,w\right) + \Gamma_1\left(K_0 f,w\right) \\
&\quad + \left(\frac{\partial T}{\partial t} - K_0 T\right)\frac{\partial f}{\partial t} + \Gamma_1\left(T,\frac{\partial f}{\partial t}\right) \\
&\quad + \left(-\frac{\partial\varphi}{\partial t} + K_0\varphi + \Gamma_1\left(V,w\right) + T\frac{\partial V}{\partial t}\right)f \\
&= \Gamma_1\left(\frac{\partial w}{\partial t} - \varphi,f\right) - K_0\Gamma_1\left(f,w\right) + \Gamma_1\left(K_0 f,w\right) \qquad (7.79) \\
&\quad + \left(K_0 f\right)\frac{\partial T}{\partial t} - \left(K_0 T\right)K_0 f + \Gamma_1\left(T,K_0 f\right) + \Gamma_1\left(T,Vf\right) \\
&\quad + \left(-\frac{\partial\varphi}{\partial t} + K_0\varphi + \Gamma_1\left(V,w\right) + T\frac{\partial V}{\partial t} + \frac{\partial T}{\partial t}V\right)f
\end{aligned}
$$

(T only depends on t)

$$
= \Gamma_1\left(\frac{\partial w}{\partial t} - \varphi,f\right) - K_0\Gamma_1\left(f,w\right) + \Gamma_1\left(K_0 f,w\right) + \left(K_0 f\right)\frac{\partial T}{\partial t}
$$

$$+\left(-\frac{\partial\varphi}{\partial t}+K_0\varphi+\Gamma_1\left(V,w\right)+T\frac{\partial V}{\partial t}+\frac{\partial T}{\partial t}V\right)f=0, \tag{7.80}$$

where in (7.79) we employed the identity $\frac{\partial f}{\partial t}=\left(K_0\dot{+}V\right)f$. The equality in (7.80) follows by our assumptions (a) and (b).

(ii). We compute

$$\begin{aligned}
&\mathcal{D}(n)=\frac{\partial n}{\partial t}-\Gamma_1\left(\sigma_L,n\right)-K_0 n\\
&=-\frac{\partial}{\partial t}\left(\Gamma_1\left(\sigma_L,w\right)\right)+\frac{\partial}{\partial t}\left(\varepsilon T\right)-\frac{\partial}{\partial t}\varphi\\
&\quad+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)-\Gamma_1\left(\sigma_L,\varepsilon T\right)+\Gamma_1\left(\sigma_L,\varphi\right)\\
&\quad+K_0\left(\Gamma_1\left(\sigma_L,w\right)\right)-K_0\left(\varepsilon T\right)+K_0\varphi\\
&=-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t},w\right)-\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}\right)+\varepsilon\frac{\partial T}{\partial t}+\frac{\partial\varepsilon}{\partial t}T-\frac{\partial\varphi}{\partial t}\\
&\quad+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)-\Gamma_1\left(\sigma_L,\varepsilon\right)T+\Gamma_1\left(\sigma_L,\varphi\right)\\
&\quad+K_0\left(\Gamma_1\left(\sigma_L,w\right)\right)-K_0\left(\varepsilon\right)T+K_0\varphi\\
&=-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t}+V,w\right)-\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}-\varphi\right)+\varepsilon\frac{\partial T}{\partial t}+\frac{\partial\varepsilon}{\partial t}T-\frac{\partial\varphi}{\partial t}\\
&\quad+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)-\Gamma_1\left(\sigma_L,\varepsilon\right)T\\
&\quad+K_0\left(\Gamma_1\left(\sigma_L,w\right)\right)-K_0\left(\varepsilon\right)T+K_0\varphi+\Gamma_1\left(V,w\right)\\
&=-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t}+V,w\right)-\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}-\varphi\right)+V\frac{\partial T}{\partial t}\\
&\quad-\left(K_0\sigma_L+\frac{1}{2}\Gamma_1\left(\sigma_L,\sigma_L\right)\right)\frac{\partial T}{\partial t}+\frac{\partial\varepsilon}{\partial t}T-\frac{\partial\varphi}{\partial t}\\
&\quad+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)-\Gamma_1\left(\sigma_L,\varepsilon\right)T\\
&\quad+K_0\left(\Gamma_1\left(\sigma_L,w\right)\right)-K_0\left(\varepsilon\right)T+K_0\varphi+\Gamma_1\left(V,w\right)\\
&=-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t}+V,w\right)-\Gamma_1\left(\sigma_L,\frac{\partial}{\partial t}w-\varphi\right)+\frac{\partial(VT)}{\partial t}\\
&\quad-\left(K_0\sigma_L+\frac{1}{2}\Gamma_1\left(\sigma_L,\sigma_L\right)\right)\frac{\partial T}{\partial t}+\frac{\partial\left(\varepsilon-V\right)}{\partial t}T-\frac{\partial\varphi}{\partial t}\\
&\quad+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)-\Gamma_1\left(\sigma_L,\varepsilon\right)T\\
&\quad+K_0\left(\Gamma_1\left(\sigma_L,w\right)\right)-K_0\left(\varepsilon\right)T+K_0\varphi+\Gamma_1\left(V,w\right)\\
&=-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t}+\varepsilon,w\right)+\left(\mathcal{D}(\varepsilon)-\frac{\partial V}{\partial t}\right)T-\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}-\varphi\right)\\
&\quad-\frac{\partial\varphi}{\partial t}+K_0\varphi+\Gamma_1\left(V,w\right)+\frac{\partial(VT)}{\partial t}
\end{aligned}$$

$$
\begin{aligned}
&-\left(K_0\sigma_L+\frac{1}{2}\Gamma_1\left(\sigma_L,\sigma_L\right)\right)\frac{\partial T}{\partial t}-\Gamma_1\left(K_0\sigma_L+\frac{1}{2}\Gamma_1\left(\sigma_L,\sigma_L\right),w\right)\\
&+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)+K_0\left(\Gamma_1\left(\sigma_L,w\right)\right)\\
=&-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t}+\varepsilon,w\right)+\left(\mathcal{D}(\varepsilon)-\frac{\partial V}{\partial t}\right)T\\
&-\left(K_0\sigma_L\right)\frac{\partial T}{\partial t}-\Gamma_1\left(K_0\sigma_L,w\right)+K_0\Gamma_1\left(\sigma_L,w\right)-\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}-\varphi\right)\\
&-\frac{\partial\varphi}{\partial t}+K_0\varphi+\Gamma_1\left(V,w\right)+\frac{\partial(VT)}{\partial t}\\
&-\frac{1}{2}\Gamma_1\left(\sigma_L,\sigma_L\right)\frac{\partial T}{\partial t}-\frac{1}{2}\Gamma_1\left(\Gamma_1\left(\sigma_L,\sigma_L\right),w\right)+\Gamma_1\left(\sigma_L,\Gamma_1\left(\sigma_L,w\right)\right)
\end{aligned}
$$

(employ Lemma 7.2 with $f=\sigma_L$)

$$
\begin{aligned}
=&-\Gamma_1\left(\frac{\partial\sigma_L}{\partial t}+\varepsilon,w\right)+\left(\mathcal{D}(\varepsilon)-\frac{\partial V}{\partial t}\right)T\\
&-\left(K_0\sigma_L\right)\frac{\partial T}{\partial t}-\Gamma_1\left(K_0\sigma_L,w\right)+K_0\Gamma_1\left(\sigma_L,w\right)-\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}-\varphi\right)\\
&-\frac{\partial\varphi}{\partial t}+K_0\varphi+\Gamma_1\left(V,w\right)+\frac{\partial(VT)}{\partial t}\\
&+\frac{1}{2}\left(K_0\left(\sigma_L^2\right)\frac{\partial T}{\partial t}+\Gamma_1\left(K_0\left(\sigma_L^2\right),w\right)-K_0\Gamma_1\left(\sigma_L^2,w\right)+\Gamma_1\left(\sigma_L^2,\frac{\partial w}{\partial t}-\varphi\right)\right)\\
&-\sigma_L\left(K_0\left(\sigma_L\right)\frac{\partial T}{\partial t}+\Gamma_1\left(K_0\left(\sigma_L\right),w\right)-K_0\Gamma_1\left(\sigma_L,w\right)+\Gamma_1\left(\sigma_L,\frac{\partial w}{\partial t}-\varphi\right)\right).
\end{aligned}
\tag{7.81}
$$

Substituting the equalities (a), (b) and (c) in (7.81) shows (ii), i.e. $\mathcal{D}(n)=0$.

(iii) Let f be a function in the domain of K_0-D_1. As in equation (7.10) of §7.1 the $\mathbb{E}_{t,x_0}$-martingale $M_{f,t_0}(t)$, $t\geqslant t_0$, is given by the equality in (7.73). Let f and g be two functions in $D\left(K_0-D_1\right)$. Then the quadratic covariation $\left\langle M_{f,t_0},M_{g,t_0}\right\rangle(t)$ of $M_{f,t_0}(t)$ and $M_{g,t_0}(t)$ is given by

$$
\left\langle M_{f,t_0},M_{g,t_0}\right\rangle(t)=\int_{t_0}^{t}\Gamma_1\left(f,g\right)\left(\tau,X(\tau)\right)d\tau. \tag{7.82}
$$

By (ii) $\mathcal{D}(n)=0$, and hence $\left(K_0-D_1\right)n=-\Gamma_1\left(\sigma_L,n\right)$. It follows that

$$
\begin{aligned}
n\left(t,X(t)\right)-n\left(t_0,X\left(t_0\right)\right)&=M_{n,t_0}(t)-\int_{t_0}^{t}\left(K_0-D_1\right)n\left(\tau,X(\tau)\right)d\tau\\
&=M_{n,t_0}(t)+\int_{t_0}^{t}\Gamma_1\left(\sigma_L,n\right)\left(\tau,X(\tau)\right)d\tau.
\end{aligned}
\tag{7.83}
$$

Let f be a function in $D(K_0 - D_1)$. From Itô's formula we obtain:

$$\begin{aligned}
&\exp\left(M_{f,t_0}(t) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(t)\right) n(t, X(t)) - n(t_0, X(t_0)) \\
&= \int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(s)\right) n(s, X(s))\, dM_{f,t_0}(s) \\
&\quad - \frac{1}{2}\int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(s)\right) n(s, X(s)) d\langle M_{f,t_0}, M_{f,t_0}\rangle(s) \\
&\quad + \frac{1}{2}\int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(s)\right) n(s, X(s)) d\langle M_{f,t_0}, M_{f,t_0}\rangle(s) \\
&\quad + \int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(s)\right) dM_{n,t_0}(s) \\
&\quad - \int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(s)\right) (K_0 - D_1)\, n(s, X(s))\, ds \\
&\quad + \int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(s)\right) \Gamma_1(f, n)(s, X(s))\, ds
\end{aligned}$$

(employ (7.83))

$$\begin{aligned}
&= \mathbb{E}_{t,x_0}\text{-martingale} \\
&\quad + \int_{t_0}^t \exp\left(M_{f,t_0}(s) - \frac{1}{2}\langle M_{f,t_0}\rangle(s)\right) \left(\Gamma_1(f, n) + \Gamma_1(\sigma_L, n)\right)(s, X(s))\, ds,
\end{aligned} \tag{7.84}$$

where we wrote $\langle M_{f,t_0}\rangle(s) = \langle M_{f,t_0}, M_{f,t_0}\rangle(s)$. Suppose $\Gamma_1(f + \sigma_L, w) = 0$. From (7.84) it follows that the process

$$\begin{aligned}
t \mapsto &\exp\left(M_{f,t_0}(t) - \frac{1}{2}\langle M_{f,t_0}, M_{f,t_0}\rangle(t)\right) \\
&= \exp\left(M_{f,t_0}(t) - \frac{1}{2}\int_{t_0}^t \Gamma_1(f, f)(s, X(s))\, ds\right) n(t, X(t))
\end{aligned}$$

is a $\mathbb{E}_{t_0,x}$-martingale. So, with $f = -\sigma_L$, assertion (iii) of Theorem 7.5 follows, and completes the proof of Theorem 7.5. □

The following theorem can be considered as a complex version of the Noether theorem: see Theorem 7.5 above and Theorem 3.1 in [Zambrini (1998a)]. It has a physical interpretation: $\mathcal{N}(t)$, defined by

$$\mathcal{N}(t)f = i\Gamma_1(f, w) - T(t)\left(K_0 \dot{+} V\right) f - \varphi f,$$

is called a *Noether observable.*

Theorem 7.6. *Let the functions T, w, and φ be related as in* (a′) *and* (b′) *below:*

(a′) $K_0 f \frac{dT}{dt} = K_0\Gamma_1(f,w) - \Gamma_1(K_0 f, w) + i\Gamma_1\left(f, \frac{\partial w}{\partial t} - \varphi\right)$ *for all functions* f *belonging to* $D(K_0 - D_1)$ *for which* $\Gamma_1(K_0 f, w)$ *makes sense as well.*

(b′) $\frac{\partial \varphi}{\partial t} - iK_0\varphi = -\Gamma_1(V,w) - \frac{\partial(TV)}{\partial t}$.

Then the operators $\mathcal{N}(t)$ *and* $\frac{\partial}{i\partial t} - \left(K_0 \dot{+} V\right)$ *commute.*

Suppose $\int K_0 f dm = 0$, $f \in D(K_0) \bigcap L^1(E,m)$. Then the adjoint $\mathcal{N}(t)^*$ is given by

$$\mathcal{N}(t)^* f = i\Gamma_1(f, \overline{w}) - 2i(K_0\overline{w})f - \overline{T(t)}\left(K_0 \dot{+} V\right) f - \overline{\varphi} f.$$

Hence the self-adjoint operator $\frac{\partial}{i\partial t} - \left(K_0 \dot{+} V\right)$ also commutes with the operators $\mathcal{N}(t) + \mathcal{N}(t)^*$ and $\mathcal{N}(t) - \mathcal{N}(t)^*$.

Proof. Let f be a "smooth enough" function. Then a calculation yields:

$$\begin{aligned}
&\mathcal{N}(t)\left(\frac{\partial}{i\partial t} - \left(K_0 \dot{+} V\right)\right) f - \left(\frac{\partial}{i\partial t} - \left(K_0 \dot{+} V\right)\right)\mathcal{N}(t) f \\
&= i\Gamma_1\left(\frac{\partial f}{i\partial t} - \left(K_0 \dot{+} V\right) f, w\right) - T(t)\left(K_0 \dot{+} V\right)\left(\frac{\partial f}{i\partial t} - \left(K_0 \dot{+} V\right) f\right) \\
&\quad - \varphi\left(\frac{\partial f}{i\partial t} - \left(K_0 \dot{+} V\right) f\right) \\
&\quad - \left(\frac{\partial}{i\partial t} - \left(K_0 \dot{+} V\right)\right)\left(i\Gamma_1(f,w) - T(t)\left(K_0 \dot{+} V\right) f - \varphi f\right) \\
&= \Gamma_1\left(\frac{\partial f}{\partial t}, w\right) - i\Gamma_1(K_0 f, w) - i\Gamma_1(Vf, w) - \frac{1}{i}T(t)K_0\left(\frac{\partial f}{\partial t}\right) \\
&\quad - \frac{1}{i}T(t)V\frac{\partial f}{\partial t} + T(t)\left(K_0 \dot{+} V\right)^2 f - \frac{1}{i}\varphi\frac{\partial f}{\partial t} + \varphi K_0 f + \varphi V f \\
&\quad - \frac{\partial}{\partial t}\Gamma_1(f,w) + \frac{1}{i}\frac{\partial}{\partial t}\left(T(t)\left(K_0 \dot{+} V\right) f\right) + \frac{1}{i}\frac{\partial(\varphi f)}{\partial t} \\
&\quad + i\left(K_0 \dot{+} V\right)\Gamma_1(f,w) - T(t)\left(K_0 \dot{+} V\right)^2 f - K_0(\varphi f) - \varphi V f \\
&= \Gamma_1\left(\frac{\partial f}{\partial t}, w\right) - i\Gamma_1(K_0 f, w) - iV\Gamma_1(f,w) - if\Gamma_1(V,w) \\
&\quad - \frac{1}{i}T(t)K_0\left(\frac{\partial f}{\partial t}\right) - \frac{1}{i}T(t)V\frac{\partial f}{\partial t} - \frac{1}{i}\varphi\frac{\partial f}{\partial t} + \varphi K_0 f \\
&\quad - \Gamma_1\left(\frac{\partial f}{\partial t}, w\right) - \Gamma_1\left(f, \frac{\partial w}{\partial t}\right) + \frac{1}{i}\frac{\partial T(t)}{\partial t}K_0 f + \frac{1}{i}\frac{\partial T(t)}{\partial t}V f
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{i}T(t)\frac{\partial V}{\partial t}f + \frac{1}{i}T(t)K_0\left(\frac{\partial f}{\partial t}\right) + \frac{1}{i}T(t)V\frac{\partial f}{\partial t} \\
&+ \frac{1}{i}\frac{\partial \varphi}{\partial t}f + \frac{1}{i}\varphi\frac{\partial f}{\partial t} + iK_0\Gamma_1(f,w) + iV\Gamma_1(f,w) \\
&- (K_0\varphi) f + \Gamma_1(f,\varphi) - \varphi K_0 f \\
= &-i\Gamma_1(K_0 f, w) - if\Gamma_1(V,w) - \Gamma_1\left(f, \frac{\partial w}{\partial t} - \varphi\right) + \frac{1}{i}\frac{\partial T(t)}{\partial t}K_0 f \\
&+ \frac{1}{i}\frac{\partial (T(t)V)}{\partial t}f + \frac{1}{i}\frac{\partial \varphi}{\partial t}f + iK_0\Gamma_1(f,w) - (K_0\varphi) f \\
= &\frac{1}{i}\frac{\partial T(t)}{\partial t}(K_0 f) - \frac{1}{i}K_0\Gamma_1(f,w) + \frac{1}{i}\Gamma_1(K_0 f, w) - \Gamma_1\left(f, \frac{\partial w}{\partial t} - \varphi\right) \\
&+ \frac{1}{i}f\left\{\Gamma_1(V,w) + \frac{\partial (T(t)V)}{\partial t} + \frac{\partial \varphi}{\partial t} + \frac{1}{i}K_0\varphi\right\}. \qquad (7.85)
\end{aligned}
$$

The result in Theorem 7.6 follows from the assumptions (a′) and (b′). □

Corollary 7.1. *Suppose that the functions w, T (which only depends on t), and ψ (which only depends on the space variable, not on the time t) possess the following properties:*

(a′) The set of functions f for which the equality

$$K_0 f\frac{dT}{dt} = K_0\Gamma_1(f,w) - \Gamma_1(K_0 f, w) + \Gamma_1(f, K_0 w + \psi)$$

makes sense and is valid is dense in the space $L^2(E \times [t_0, T], dm \times dt)$.

(b′) The following equality is valid:

$$\left(\frac{\partial^2}{\partial t^2} + K_0^2\right) w + K_0\psi = -\Gamma_1(V,w) - \frac{\partial (TV)}{\partial t}.$$

Put

$$\mathcal{N}(t)f = i\Gamma_1(f,w) - T(t)\left(K_0 \dot{+} V\right) f - \left(\left(\frac{\partial}{\partial t} + iK_0\right) w + i\psi\right) f,$$

where $f \in D\left(K_0 \dot{+} V\right)$. Then $\mathcal{N}(t)$ commutes with $\dfrac{\partial}{i\partial t} - \left(K_0 \dot{+} V\right)$.

Proof. Set $\varphi = \dfrac{\partial w}{\partial t} + iK_0 w + i\psi$ in Theorem 7.6. Then

$$
\begin{aligned}
&K_0\Gamma_1(f,w) - \Gamma_1(K_0 f, w) + i\Gamma_1\left(f, \frac{\partial w}{\partial t} - \varphi\right) \\
&= K_0\Gamma_1(f,w) - \Gamma_1(K_0 f, w) + \Gamma_1(f, K_0 w + \psi) = K_0 f\frac{dT}{dt}. \qquad (7.86)
\end{aligned}
$$

This shows (a′) of Theorem 7.6. Since $\dfrac{\partial \psi}{\partial t} = 0$, we see that (b′) of Theorem 7.6 is satisfied as well. This proves Corollary 7.1. □

The following proposition isolates the properties of the function w.

Proposition 7.3. *Suppose that the function w has property (a) of Theorem 7.5, or (a′) of Theorem 7.6, or (a′) of its Corollary 7.1. Then, for all functions f, $g \in D(D_1 - K_0)$, the following identity is true:*

$$\Gamma_1(f,g)\frac{dT}{dt} + \Gamma_1(\Gamma_1(f,g),w) = \Gamma_1(\Gamma_1(f,w),g) + \Gamma_1(f,\Gamma_1(g,w)). \tag{7.87}$$

Remark 7.10. Let χ be a smooth enough function. From the proof of Proposition 7.3 it follows that the mapping

$$f \mapsto (K_0 f)\frac{dT}{dt} - K_0(\Gamma_1(f,w)) + \Gamma_1(K_0 f, w) + \Gamma_1(f,\chi)$$

is a *derivation* if and only if (7.87) is satisfied for all functions f and g in a "large enough" algebra of functions belonging to $D(D_1 - K_0)$.

Proof. [Proof of Proposition 7.3.] Let f and g be functions in $D(D_1 - K_0)$ with the property that its product fg also belongs to $D(D_1 - K_0)$. We write

$$\chi = \frac{\partial w}{\partial t} - \varphi, \quad \chi = \frac{1}{i}\left(\frac{\partial w}{\partial t} - \varphi\right), \quad \text{or} \quad \chi = -K_0 w - \psi,$$

as the case may be. Then

$$\begin{aligned}
&K_0(fg)\frac{dT}{dt} - K_0\Gamma_1(fg,w) + \Gamma_1(K_0(fg),w) + \Gamma_1(fg,\chi)\\
&= ((K_0 f)g - \Gamma_1(f,g) + f(K_0 g))\frac{dT}{dt} - K_0(\Gamma_1(f,w)g + f\Gamma_1(g,w))\\
&\quad + \Gamma_1((K_0 f)g - \Gamma_1(f,g) + f(K_0 g),w) + f\Gamma_1(g,\xi) + \Gamma_1(f,\chi)g\\
&= ((K_0 f)g - \Gamma_1(f,g) + f(K_0 g))\frac{dT}{dt}\\
&\quad - (K_0\Gamma_1(f,w))g + \Gamma_1(\Gamma_1(f,w),g) - \Gamma_1(f,w)(K_0 g)\\
&\quad - (K_0 f)\Gamma_1(g,w) + \Gamma_1(f,\Gamma_1(g,w)) - f(K_0\Gamma_1(g,w))\\
&\quad + \Gamma_1(K_0 f,w)g + (K_0 f)\Gamma_1(g,w) - \Gamma_1(\Gamma_1(f,g),w)\\
&\quad + \Gamma_1(f,w)K_0 g + f\Gamma_1(K_0 g,w) + f\Gamma_1(g,\chi) + \Gamma_1(f,\chi)g\\
&= \left((K_0 f)\frac{dT}{dt} - K_0\Gamma_1(f,w) + \Gamma_1(K_0 f,w) + \Gamma_1(f,\chi)\right)g\\
&\quad + f\left((K_0 g)\frac{dT}{dt} - K_0\Gamma_1(g,w) + \Gamma_1(K_0 g,w) + \Gamma_1(g,\chi)\right)\\
&\quad - \Gamma_1(f,g)\frac{dT}{dt} + \Gamma_1(\Gamma_1(f,w),g) + \Gamma_1(f,\Gamma_1(g,w)) - \Gamma_1(\Gamma_1(f,g),w).
\end{aligned} \tag{7.88}$$

An application of either (a) of Theorem 7.5 or (a′) of Theorem 7.6 or of Corollary 7.1 then yields (7.87) in Proposition 7.3. □

Remark 7.11. Let the functions T, w and ψ satisfy (a′) and (b′) of Corollary 7.1. Put $\chi = K_0 w + \psi$. Then the triple (T, w, χ) satisfies:

(a) $K_0 f \dfrac{dT}{dt} = K_0 \Gamma_1(f, w) - \Gamma_1(K_0 f, w) + \Gamma_1(f, \chi)$ (for f in a dense subspace of $L^2(E, m)$);

(b) $\dfrac{\partial^2 w}{\partial t^2} + K_0 \chi = -\Gamma_1(V, w) - \dfrac{\partial (TV)}{\partial t}$;

(c) $\dfrac{\partial (\chi - K_0 w)}{\partial t} = 0$.

In order to find Noether observables the equations (a), (b) and (c) have to be integrated simultaneously. Proposition 7.3 simplifies this somewhat in the sense that one first tries to find w, then χ. The couple (w, χ) also has to satisfy (b). Notice that in case $E = \mathbb{R}^d$ and $K_0 f = -\dfrac{1}{2}\sum_{j,k=1}^{d} \dfrac{\partial}{\partial x_j} a_{j,k} \dfrac{\partial}{\partial x_k} f$, then $\Gamma_1(f, g) = \sum_{j,k=1}^{d} a_{j,k} \dfrac{\partial f}{\partial x_j}\dfrac{\partial g}{\partial x_k}$. Upon choosing linear functions f and g we see that w has to satisfy:

$$\begin{aligned} & a_{j,k}\frac{dT}{dt} + \sum_{\ell,m=1}^{d} a_{\ell,m}\frac{\partial a_{j,k}}{\partial x_m}\frac{\partial w}{\partial x_\ell} \\ &= 2\sum_{\ell,m=1}^{d} a_{j,\ell}a_{k,m}\frac{\partial^2 w}{\partial x_\ell \partial x_m} + \sum_{\ell,m=1}^{d} a_{j,m}\frac{\partial a_{k,\ell}}{\partial x_m}\frac{\partial w}{\partial x_\ell} + \sum_{\ell,m=1}^{d} \frac{\partial a_{j,\ell}}{\partial x_m} a_{k,m}\frac{\partial w}{\partial x_\ell}. \end{aligned}$$

It follows that the matrix with entries $\dfrac{\partial^2 w}{\partial x_l \partial x_m}$ is, up to a first order perturbation, $\dfrac{1}{2}\dfrac{dT}{dt}\times$ the inverse of the matrix $(a_{\ell,m})_{\ell,m=1}^{d}$.

7.4.1 *Classical Noether theorem*

Let Q $(= E)$ be the *configuration manifold* of a *classical dynamical system*. The paths are C^2-maps $q : t \mapsto q(t)$, $t \in I := [t_0, T]$. The Lagrangian is written as $(q, \dot{q}, t) \mapsto L(q, \dot{q}, t)$: $\dot{q} \in TQ$, the tangent bundle of Q. For simplicity we assume here $Q = \mathbb{R}^3$. Then TQ may be identified with Q. We assume an *external force* of the form $F = -\nabla V$, where V is a *scalar potential*. Then $L = \frac{1}{2}|\dot{q}|^2 - V(q, t)$. The *action functional* S, defined on a a domain $D(S) \subseteq C^2\left([t_0,T],Q\right)$, is given by

$$S(q(\cdot);t_0,u) = \int_{t_0}^{u} L(q(s),\dot{q}(s),s)ds.$$

Hamilton's least action principle says that among all regular trajectories between two fixed configurations $q(t_0) = q_0$ and $q(T) = q_1$, the physical motion $\overline{q}$ is a *critical point* of the action S, i.e. its *variational* (= its Gâteaux) derivative in any smooth direction δq cancels: $\delta S(\overline{q})(\delta q) = 0$. Equivalently $\overline{q}$ solves the *Euler-Lagrange* equations in Q:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = \frac{\partial L}{\partial q}.$$

For the *Hamilton-Jacobi theory* one adds an initial or final boundary condition: $S(q_0) = S_0(q_0)$ or $S(T) = S_T(q_1)$. *Noether's theorem* is the second most important theorem of classical Lagrangian mechanics. Let $U_\alpha : Q \times I \to Q \times I$ be given a given one-parameter group ($\alpha \in \mathbb{R}$) local group of *transformations* of the (q,t)-space: $(q,t) \mapsto (Q(q,t;\alpha), \tau(q,t;\alpha))$. The functions Q and τ are supposed to be C^2 in their variables, and $Q(q,t;0) = q$, $\tau(q,t;0) = t$. Therefore

$$\begin{aligned} Q\left(q,t;\alpha\right) &= q + \alpha X(q,t) + o(\alpha);\\ \tau\left(q,t;\alpha\right) &= t + \alpha T(q,t) + o(\alpha). \end{aligned} \tag{7.89}$$

The pair $(X(q,t),T(q,t))$ is called the tangent vector field of the family $\{U_\alpha\}$, and (T,X) its infinitesimal generator. The action S is said to be *divergence invariant* if there exists a C^2-function Φ, such that for all $\alpha > 0$ but small enough, the equality

$$S\left(q(\cdot);t_0',t_1'\right) = S\left(Q(\cdot);\tau_0',\tau_1'\right) - \alpha \int_{t_0'}^{t_1'} \frac{d\Phi}{dt}(q(t),t)dt + o(\alpha), \tag{7.90}$$

for any C^2-trajectory $q(\cdot)$ in $D(S)$ and for any time interval $[t_0',t_1']$ in $[t_0,T]$. Noether's theorem says that for a divergence invariant Lagrangian action the expression

$$\left[\left(\frac{\partial L}{\partial \dot{q}}\right)X + \left(L - \frac{\partial L}{\partial \dot{q}}\dot{q}\right)T - \Phi\right](\overline{q}(t),t)$$

is constant. The first factor $p = \dfrac{\partial L}{\partial \dot{q}}$ defines the *momentum observable*, and the second one *the energy* $-H = L - \dfrac{\partial L}{\partial \dot{q}}\dot{q}$. According to E. Cartan the Noether constant can be considered as the central geometrical object of *classical Hamiltonian mechanics*.

7.4.2 *Some problems*

We want to mention some problems which are related to this and earlier chapters. As we proved in Chapter 3 Theorems 2.9 through 2.13 are true if the space E is a Polish space, and if $C_b(E)$ is the space of all bounded continuous functions on E. Instead of the topology of uniform convergence we consider the *strict topology*. This topology is generated by semi-norms of the form: $f \mapsto \sup_{x\in E} |u(x)f(x)|$, $f \in C_b(E)$. The functions $u \geqslant 0$ have the property that for every $\alpha > 0$ the set $\{u \geqslant \alpha\}$ is compact (or is contained in a compact subset of E). The functions u need not be continuous.

Problem 7.1. *Is there a relationship with work done by Eberle [Eberle (1995, 1996, 1999)]?*

In [Altomare and Attalienti (2002a)] the authors Altomare and Attaliente take a somewhat different point of view. Their state space is still second countable and locally compact. They take a bounded continuous function $w : E \to (0, \infty)$ and the consider the space $C_0^w(E)$ as being the collection of those function $f \in C(E)$ with the property that the function wf belongs to $C_0(E)$. The space $C_0^w(E)$ is supplied with the norm $\|f\|_w = \|wf\|_\infty$, $f \in C_0^w(E)$. They study the semigroup $P^w(t)f := w^{-1}P(t)(wf)$, where $P(t)$, $t \geqslant 0$, is a Feller semigroup. Properties of $P(t)$ are transferred to ones of $P^w(t)$ and vice versa. Using these weighted continuous function spaces the authors prove some new results on the well-posedness of the Black-Scholes equation in a weighted continuous function space; see [Altomare and Attalienti (2002b)]; see Chapter 5 for more on this in the usual case. In [Mininni and Romanelli (2003)] Mininni and Romanelli estimate the trend coefficient in the Black-Scholes equation. The paper is somewhat complementary to what we do in Chapter 5.

Problem 7.2. *Is it possible to rephrase Theorems 2.9 through 2.13 for reciprocal Markov processes and diffusions?*

Martingales should then replaced with differences of forward and backward martingales. A stochastic process $(M(t) : t \geqslant 0)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *backward martingale* if $\mathbb{E}\left[M(t) \,\middle|\, \mathcal{F}^s\right] = M(s)$, $\mathbb{P}$-almost surely, where $t < s$, and $\mathcal{F}^s$ is the σ-field generated by the information from the future: $\mathcal{F}^s = \sigma\left(X(u) : u \geqslant s\right\}$. Of course we assume that $M(t)$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$, $t \geqslant 0$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An E-valued process $(X(t) : 0 \leqslant t \leqslant 1)$ is called *reciprocal* if for any $0 \leqslant s < t \leqslant 1$ and every pair of events $A \in \sigma(X(\tau) : \tau \in (s,t))$, $B \in \sigma(X(\tau) : \tau \in [0,s] \bigcup [t,1])$ the equality

$$\mathbb{P}\left[A \bigcap B \mid X(s), X(t)\right] = \mathbb{P}\left[A \mid X(s), X(t)\right] \mathbb{P}\left[B \mid X(s), X(t)\right] \quad (7.91)$$

is valid. By $\mathcal{D}$ we denote the set

$$\mathcal{D} = \{(s,x,t,B,u,z) : (x,z) \in E \times E,\ 0 \leqslant s < t < u \leqslant 1,\ B \in \mathcal{E}\}. \quad (7.92)$$

A function $P : \mathcal{D} \to [0,\infty)$ is called a *reciprocal probability distribution* or a *Bernstein probability* if the following conditions are satisfied:

(i) the mapping $B \mapsto P(s,x,t,B,u,z)$ is a probability measure on $\mathcal{E}$ for any $(x,z) \in E \times E$ and for any $0 \leqslant s < t < u \leqslant 1$;
(ii) the function $(x,z) \mapsto P(s,x,t,B,u,z)$ is $\mathcal{E} \otimes \mathcal{E}$-measurable for any $0 \leqslant s < t < u \leqslant 1$;
(iii) For every pair $(C,D) \in \mathcal{E} \otimes E$, $(x,y) \in E \times E$, and for all $0 \leqslant s < t < u \leqslant 1$ the following equality is valid:

$$\begin{aligned}&\int_D P(s,x,u,d\xi,v,y)\, P(s,x,t,C,u,\xi) \\ &= \int_C P(s,x,t,d\eta,v,y)\, P(t,\eta,u,D,v,y).\end{aligned}$$

Then the following theorem is valid for $E = \mathbb{R}^\nu$ (see [Jamison (1974)]).

Theorem 7.7. *Let $P(s,x,t,B,u,y)$ be a reciprocal transition probability function and let μ be a probability measure on $\mathcal{E} \otimes \mathcal{E}$. Then there exists a unique probability measure $\mathbb{P}_\mu$ on $\mathcal{F}$ with the following properties:*

(1) With respect to $\mathbb{P}_\mu$ the process $(X(t) : 0 \leqslant t \leqslant 1)$ is reciprocal;
(2) For all $(A,B) \in \mathcal{E} \otimes \mathcal{E}$ the equality $\mathbb{P}_\mu[X_0 \in A,\ X_1 \in B] = \mu(A \times B)$ is valid;
(3) For every $0 \leqslant s < t < u \leqslant 1$ and for every $A \in \mathcal{E}$ the equality

$$\mathbb{P}_\mu\left[X(t) \in A \,\middle|\, X(s), X(u)\right] = P(s, X(s), t, A, u, X(u)) \quad \textit{is valid.}$$

For more details see [Thieullen (1993)] and [Thieullen (1998)]. An example of a reciprocal Markov probability can be constructed as follows; it is kind of a pinned Markov process. Let

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t) : t \geqslant 0), (\vartheta_t : t \geqslant 0), (E, \mathcal{E})\}$$

be a (strong) time-homogeneous Markov process, and suppose that for every $t > 0$ and every $x \in E$, the probability measure $B \mapsto \mathbb{P}[X(t) \in B]$ has a

Radon-Nikodym derivative $p_0(t,x,y)$ with respect to some reference measure dy. Also suppose that $p_0(t,x,y)$ is strictly positive and continuous on $(0,\infty)\times E\times E$. Put

$$p(s,x,u,\xi,v,y) = \frac{p_0(u-s,x,\xi)\,p_0(v-u,\xi,y)}{p_0(v-s,x,y)}, \quad 0 \leqslant s < u < v.$$

Put $P(s,x,u,B,v,y) = \int_B p(s,x,u,\xi,v,y)\,d\xi$. Then P is a reciprocal Markov probability.

7.4.2.1 *Conclusion*

This chapter is a reworked version of [Van Casteren (2003)]. One of the main results is contained in Theorem 7.3. The method of proof is based on martingale methods. For more information on viscosity solutions the reader is referred to [Crandall *et al.* (1992b)]. Another feature of the present chapter is the statement and proof of a generalized Noether theorem (Theorem 7.5) and its complex companion (Theorem 7.6). The proofs are of a computational character; they only depend on the properties of the generator of the diffusion and the corresponding carré du champ operator. They imitate and improve results obtained by Zambrini in [Zambrini (1998a)]. Moreover the results solve problems posed in [Van Casteren (2001)] (Problem 4, Theorem 16, pp. 257-258) and in §2 of [Van Casteren (2000a)]. In particular see Problem 4 and the question prior and related to the suggested Theorem 6 on pp. 48–50 of [Van Casteren (2000a)]. The present chapter is a substantial extension of [Van Casteren (2000b)].

PART 4

Long Time Behavior

This page is intentionally left blank

Chapter 8

On non-stationary Markov processes and Dunford projections

The aim of this chapter is to present some criteria for checking ergodicity of time-continuous finite or infinite Markov chains in the sense that $\dot{\mu}(t) = K(t)\mu(t)$, where every $K(t)$, $t \in \mathbb{R}$, is a weak*-closed linear Kolmogorov operator on the space of complex Borel measures $M(E)$ on a complete metrizable separable Hausdorff space E, and so E is a Polish space. The obtained results are valid in the non-stationary case and can be used as reliable and valuable tools to establish ergodicity. Some theoretical approximation results are given as well. The present chapter was initiated by some results in the Ph.D. thesis of Katilova [Katilova (2004)]: see [Van Casteren (2005a)] as well. What in the present chapter is called $\sigma(M(E), C_b(E))$-convergence, or $\sigma(M(E), C_b(E))$-topology, in the probability literature is often referred to as weak convergence, or weak topology. In functional analytic terms these notions should be called weak*-convergence, or weak*-topology. Here "weak*" refers to the pre-dual space of $M(E)$ which is the space $C_b(E)$ endowed with the strict topology. In order to avoid misunderstandings we sometimes write "$\sigma(M(E), C_b(E))$" instead of "weak" (probabilistic notion) or "weak*" (functional analytic notion). Nevertheless, we will employ the notation "weak*" and "$\sigma(M(E), C_b(E))$" interchangeably; we will write e.g. "weak*-continuous semigroup" where, strictly speaking, we mean "$\sigma(M(E), C_b(E))$-continuous semigroup". For applications of the use of invariant measures for time-dependent problems the reader may want to consult [Geissert *et al.* (2009)] and [Hieber *et al.* (2009)].

8.1 Introduction

Let E be a complete metrizable topological space which is separable with Borel field $\mathcal{E}$: in other words E is a Polish space. By $M(E)$ we denote the vector space of all complex Borel measures on E, supplied with the total variation norm:

$$\operatorname{Var}(\mu) = \sup\left\{\sum_{j=1}^{n} |\mu(B_j)| : B_j \text{ is a partition of } E\right\}. \tag{8.1}$$

In view of inequality (3) in Theorem 8.1 in Section 8.2 (except in Example 8.4) we will not use the total variation norm, but the following equivalent one:

$$\|\mu\| = \sup\{|\mu(B)| : B \in \mathcal{E}\}, \quad \mu \in M(E). \tag{8.2}$$

In fact we have $\|\mu\| \leqslant \operatorname{Var}(\mu) \leqslant 4\|\mu\|$. In the other sections and in Example 8.4 the symbol $\operatorname{Var}(\mu)$, $\mu \in M(E)$, stands for the total variation norm of the measure μ. Let f be a bounded Borel function and μ a measure in $M(E)$. Instead of $\int_E f d\mu$ we often write $\langle f, \mu\rangle$. By hypothesis the family $K(t)$, $t \in \mathbb{R}$, is a family of linear operators with domain and range in $M(E)$ which are $\sigma(M(E), C_b(E))$-closed. This means that if $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in $D(K(t))$, the domain of $K(t)$, for which there exists Borel measures μ and $\nu \in M(E)$ such that, for all $f \in C_b(E)$, $\lim_{n\to\infty}\langle f, \mu_n\rangle = \langle f, \mu\rangle$ and such that $\lim_{n\to\infty}\langle f, K(t)\mu_n\rangle = \langle f, \nu\rangle$, that then μ belongs to $D(K(t))$ and $K(t)\mu = \nu$. Instead of $\sigma(M(E), C_b(E))$-closed we usually write weak*-closed. An important example of a weak*-closed linear operator is the adjoint of an operator with domain and range in $C_b(E)$. We consider a continuous system of the form:

$$\dot{\mu}(t) = K(t)\mu(t), \quad -\infty < t < \infty, \tag{8.3}$$

where each $K(t)$ is a weak*-closed linear operator on $M(E)$.

Definition 8.1. Let K be a weak*-closed linear operator on $M(E)$.

(a) An eigenvalue μ of K is called dominant if $\lim_{t\to\infty}\left\|e^{tK}(I-P)\right\| = 0$. Here P is the Dunford projection on the generalized eigenspace corresponding to μ; i.e. $P = \dfrac{1}{2\pi i}\displaystyle\int_\gamma (\lambda I - K)^{-1}\, d\lambda$, where γ is a (small) positively oriented circle around μ. The disc centered at μ and with circumference γ does not contain other eigenvalues.

(b) An eigenvalue μ of K is called critical if it is dominant and the zero space of $K - \mu I$ is one-dimensional.

We consider the simplex $P(E) \subset M(E)$ consisting of all Borel probability measures on E:

$$P(E) = \{\mu \in M(E) : \mu(E) = 1 \text{ and } \mu(B) \geqslant 0 \text{ for all Borel subsets of } E\}, \tag{8.4}$$

and the subspace $M_0(E)$ of co-dimension one in $M(E)$:

$$M_0(E) = \{\mu \in M(E) : \mu(E) = 0\}. \tag{8.5}$$

8.2 Kolmogorov operators and weak*-continuous semigroups

Under appropriate conditions on the family $K(t)$, $t \geqslant t_0$, a solution to the equation in (8.3), i.e. a solution to

$$\frac{d}{dt}\langle f, \mu(t)\rangle = \langle f, K(t)\mu(t)\rangle, \quad t_0 \leqslant t < \infty, \ f \in C_b(E), \tag{8.6}$$

where $\mu(t_0) \in P(E)$ is given, can be written in the form:

$$\mu(t) = X(t, t_0)\,\mu(t_0), \quad t_0 \leqslant t < \infty; \tag{8.7}$$

the operator-valued function $X(t, t_0)$ satisfies the following differential equation in weak*-sense:

$$\frac{\partial}{\partial t} X(t, t_0) = K(t) X(t, t_0). \tag{8.8}$$

It is an evolution family in the sense that $X(t, t_2)\, X(t_2, t_1) = X(t, t_1)$, $t \geqslant t_2 \geqslant t_1 \geqslant t_0$, $X(t,t) = I$. We also assume that weak^{ast}- $\lim_{t\downarrow s} X(t, s)\,\mu = \mu$, i.e.

$$\lim_{t\downarrow s} \langle f, X(t, s)\,\mu\rangle = \langle f, \mu\rangle \text{ for all } f \in C_b(E) \text{ and } \mu \in M_0(E).$$

Suppose now that for every t the operator $K(t)$ is Kolmogorov or, what is the same, has the Kolomogorov property. This in the meaning that for the operator $K(t)$ the following formulas are valid:

$$\Re K(t)\mu(E) = \Re\langle \mathbf{1}, K(t)\mu\rangle = 0 \text{ for all } \mu \in P(E) \text{ and} \tag{8.9}$$

$$\Re\langle f, K(t)\mu\rangle \geqslant 0 \text{ for all } (f, \mu) \in C_b^+(E) \times P(E) \text{ for which}$$

$$\operatorname{supp}(f) \bigcap \operatorname{supp}(\Re\mu) = \varnothing. \tag{8.10}$$

Here $C_b^+(E)$ is the convex cone of all nonnegative functions in $C_b(E)$. Unfortunately this notion is too weak for our purposes. In fact for our purposes we need a modification of the notion of (sub-)Kolmogorov operator which we label as sectorial sub-Kolmogorov operator. It is somewhat stronger than (8.10).

Definition 8.2. Let K be a linear operator with domain and range in $M(E)$. Suppose that its graph $G(K) := \{(\mu, K\mu) : \mu \in D(K)\}$ is closed in the product space $(M(E), \|\cdot\|) \times (M(E), \sigma(M(E), C_b(E)))$. Here $\sigma(M(E), C_b(E))$ stands for weak*-topology which $M(E)$ gets from its pre-dual space $C_b(E)$. The operator K is called a sub-Kolmogorov operator if for every $\mu \in D(K)$ the equality

$$\begin{aligned}&\sup\{\Re\langle f,\mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E)\}\\ &= \inf_{\varepsilon>0}\sup\{\Re\langle f,\mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\ \Re\langle f, K\mu\rangle \leqslant \varepsilon,\ f \in C_b(E)\}. \qquad (8.11)\end{aligned}$$

holds.

The sub-Kolmogorov operator K is called sectorial if it is a sub-Kolmogorov operator with the property that there exists a finite constant C such that the inequality

$$\begin{aligned}&|\lambda|\sup\{|\langle f,\mu\rangle| : |f| \leqslant \mathbf{1},\ f \in C_b(E)\}\\ &\leqslant C\sup\{|\langle f, \lambda\mu - K\mu\rangle| : |f| \leqslant \mathbf{1},\ f \in C_b(E)\} \qquad (8.12)\end{aligned}$$

holds for all $\mu \in D(K)$ and for all $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$.

The following definition should be compared with the corresponding definition in Definition 4.2: see (4.13). In fact these two notions are equivalent: this is a consequence of assertion (f) in Proposition 4.3. Lemma 8.1 is in fact a rewording of assertion (f) in the latter proposition.

Lemma 8.1. *Let L be a linear operator with domain and range in $C_b(E)$. The following assertions are equivalent:*

(i) For every $\lambda > 0$ and for every $f \in D(L)$ the following inequality holds:

$$\lambda\|f\|_\infty \leqslant \|\lambda f - Lf\|_\infty; \qquad (8.13)$$

(ii) For every $\varepsilon > 0$ the following inequality holds for all $f \in D(L)$:

$$\sup\{|f(x)| : x \in E\} \leqslant \sup\left\{|f(x)| : \Re\left(\overline{f(x)}Lf(x)\right) \leqslant \varepsilon\right\}. \qquad (8.14)$$

Definition 8.3. An operator L with domain and range in $C_b(E)$ is said to be dissipative, if for every $f \in D(L)$ and every $\varepsilon > 0$ the following identity holds:

$$\sup\{|f(x)| : x \in E\} = \sup\left\{|f(x)| : \Re\left(\overline{f(x)}Lf(x)\right) \leqslant \varepsilon,\ x \in E\right\}. \qquad (8.15)$$

An operator L with domain and range in $C_b(E)$ which satisfies the maximum principle is called sectorial if there exists a constant C such that for all $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$ the inequality

$$|\lambda|\|f\|_\infty \leqslant C\|(\lambda I - L)f\|_\infty, \quad \text{holds for all } f \in D(L). \qquad (8.16)$$

Notice that the notion of dissipativeness is equivalent to the following one: for every $f \in D(L)$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ such that

$$\lim_{n\to\infty} |f(x_n)| = \|f\|_\infty \quad \text{and} \quad \lim_{n\to\infty} \Re\left(\overline{f(x_n)} Lf(x_n)\right) \leqslant 0. \tag{8.17}$$

From (8.17) it follows that the present notion of "being dissipative" coincides with the notion in Chapter 4: see §4.2. In particular, the reader is referred to (4.13) in Definition 4.2, and to assertion (f) in Proposition 4.3. Apparently, the conditions in Remarks 8.1 and 8.2 below are not verifiable (or they might not be satisfied in interesting cases, where the operators K respectively L generate analytic semigroups). In §8.6 we give a new characterization of operators which generate an analytic (or holomorphic) semigroup. It also contains a triviality result in the sense that it characterizes an operator L as being the zero operator if only appropriate boundedness conditions are imposed on the absolute values of the expressions $\dfrac{\langle Lx, x^*\rangle}{\langle x, x^*\rangle}$, $x \in X$, $x^* \in X^*$, $\langle x, x^*\rangle \neq 0$, where X is a Banach space with dual X^{ast}: see Proposition 8.9.

Remark 8.1. Suppose that there exists $0 < \gamma < \frac{1}{2}\pi$ such that for every $\mu \in D(K)$ and every $\varepsilon > 0$ there exists a function $f \in C_b(E)$, $0 \leqslant |f| \leqslant \mathbf{1}$, such that $\mathrm{Var}(\mu) \leqslant |\langle f, \mu\rangle| + \varepsilon$ and such that there exists $\vartheta(\mu) \in \mathbb{R}$ satisfying $\pi \geqslant |\vartheta(\mu)| \geqslant \gamma + \frac{1}{2}\pi$ and $\dfrac{\langle f, K\mu\rangle}{\langle f, \mu\rangle} = \dfrac{|\langle f, K\mu\rangle|}{|\langle f, \mu\rangle|} e^{i\vartheta(\mu)}$. Then (8.12) is satisfied with C satisfying $C \sin\gamma = 1$.

Remark 8.2. Similarly, let L be an operator with domain and range in $C_b(E)$. Suppose that there exists $0 < \gamma < \frac{1}{2}\pi$ such that for every $f \in C_b(E)$ and every $\varepsilon > 0$ there exists $x \in E$, $0 \leqslant |f| \leqslant \mathbf{1}$, such that $\|f\|_\infty = |f(x)|$ and such that there exists $\vartheta(x) \in \mathbb{R}$ satisfying $\pi \geqslant |\vartheta(x)| \geqslant \gamma + \frac{1}{2}\pi$ and $\dfrac{Lf(x)}{f(x)} = \dfrac{|Lf(x)|}{|f(x)|} e^{i\vartheta(x)}$. Then the operator L is sectorial in the sense that

$$\|\lambda f - Lf\|_\infty \geqslant \sin\gamma\, |\lambda|\, \|f\|_\infty, \quad f \in D(L).$$

How to check a condition like the one in (8.11) or (8.12)? Therefore we first analyze the right-hand side of (8.11). Let $E = E^+_{\Re\mu} \bigcup E^-_{\Re\mu}$ be the Hahn-decomposition of E corresponding to the Jordan-decomposition of the measure $\Re\mu$. Then $E^+_{\Re\mu} \bigcap E^-_{\Re\mu} = \varnothing$, $(\Re\mu)_+\left(E^-_{\Re\mu}\right) = (\Re\mu)_-\left(E^+_{\Re\mu}\right) = 0$, and if $B \in \mathcal{E}$ is a subset of $E^+_{\Re\mu}$, then $(\Re\mu)(B) \geqslant 0$. In other words the signed measure $\Re\mu$ is positive on $E^+_{\Re\mu}$. Similarly the signed measure $-\Re\mu$ is positive on $E^-_{\Re\mu}$. In addition we have

$$\sup\left\{\Re\langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\, f \in C_b(E)\right\}$$

$$= \Re\mu\left(E^{+}_{\Re\mu}\right) = \sup\left\{\Re\mu(C) : C \subset E^{+}_{\Re\mu},\, C \text{ compact}\right\}. \tag{8.18}$$

Let C_n, $n \in \mathbb{N}$, be a sequence of compact subsets of $E^{+}_{\Re\mu}$ and let O_n, $n \in \mathbb{N}$, be a sequence of open subsets of E such that $C_n \subset E^{+}_{\Re\mu} \subset O_n$, and such that

$$\lim_{n\to\infty} \Re\left\langle \mathbf{1}_{C_n}, K\mu\right\rangle = \lim_{n\to\infty} \Re\left\langle \mathbf{1}_{O_n}, K\mu\right\rangle = \Re\left\langle \mathbf{1}_{E^{+}_{\Re\mu}}, K\mu\right\rangle. \tag{8.19}$$

In addition suppose that

$$\lim_{n\to\infty} \left\langle \mathbf{1}_{O_n} - \mathbf{1}_{C_n}, |\Re\mu|\right\rangle = \lim_{n\to\infty} \left\langle \mathbf{1}_{O_n} - \mathbf{1}_{C_n}, |\Re K\mu|\right\rangle = 0. \tag{8.20}$$

Here the measures $|\Re\mu|$ and $|\Re K\mu|$ stand for the variation measures of $\Re\mu$ and $\Re K\mu$ respectively. Suppose that $\Re\left\langle \mathbf{1}_{E^{+}_{\Re\mu}}, K\mu\right\rangle \leqslant 0$. Let f_n, $n \in \mathbb{N}$, be a sequence of functions in $C_b(E)$ with the property that $\mathbf{1}_{C_n} \leqslant f_n \leqslant \mathbf{1}_{O_n}$. Then from (8.19) and (8.20) it follows that

$$\Re\mu\left(E^{+}_{\Re\mu}\right) = \Re\left\langle \mathbf{1}_{E^{+}_{\Re\mu}}, \mu\right\rangle = \lim_{n\to\infty} \Re\left\langle f_n, \mu\right\rangle \quad \text{and} \tag{8.21}$$

$$0 \geqslant \Re\left\langle \mathbf{1}_{E^{+}_{\Re\mu}}, K\mu\right\rangle = \lim_{n\to\infty} \Re\left\langle f_n, K\mu\right\rangle. \tag{8.22}$$

Suppose that the inequality $\Re\left\langle \mathbf{1}_{E^{+}_{\Re\mu}}, K\mu\right\rangle \leqslant 0$ holds. Then the (in-)equalities in (8.18), (8.21), and (8.22) show that the left-hand side of (8.11) is less than or equal to its right-hand side. The converse inequality being trivial shows that equality (8.11) holds.

In order to establish an equality like the one in (8.12) it suffices to exhibit a Borel measurable function $g : E \to \mathbb{C}$ with the following properties: $|g| = \mathbf{1}$, the expression $\overline{\langle g, \mu\rangle}\, \langle g, K\mu\rangle$ is a negative real number, and

$$\operatorname{Var}(\mu) = \sup\left\{|\langle f, \mu\rangle| : |f| \leqslant 1,\, f \in C_b(E)\right\} = \langle g, \mu\rangle.$$

Next let $K = L^*$, where L is a closed linear operator with domain and range in $C_b(E)$. Suppose that $\Re Lf \leqslant 0$ on C whenever C is a compact subset of E and $f \in D(L)$ is such that $\mathbf{1}_C \leqslant f \leqslant \mathbf{1}$. Then the operator K satisfies (8.11).

Next let $\mu \in D(K)$ be such that $\Re\langle \mathbf{1}, K\mu\rangle \leqslant 0$, and suppose

$$\begin{aligned} &\sup\left\{\Re\langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\, f \in C_b(E)\right\} \\ &= \inf_{\varepsilon>0} \sup\left\{\Re\langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\, \Re\langle \mathbf{1} - f, K\mu\rangle \geqslant -\varepsilon,\, f \in C_b(E)\right\}. \end{aligned} \tag{8.23}$$

Then the inequality in (8.11) follows from (8.23).

Suppose that for every positive measure $\mu \in D(K)$ the following inequalities are satisfied: $K\mu(E) \leqslant 0$ and $\mu(B) = 0$ implies $K\mu(B) \geqslant 0$. Then, for every measure $\mu \in D(K)$ there exists a Borel subset E^+ of E such that $K\mu(E^+) \leqslant 0$, provided that in the Jordan-decomposition $\mu = \mu_+ - \mu_-$ the measure μ_+ belongs to $D(K)$. This fact follows from the next observation. Let $E = E^+ \bigcup E^-$ be the Hahn-decomposition of E corresponding to the Jordan-decomposition of the measure μ. Then $E^+ \bigcap E^- = \varnothing$, $\mu_+(E^-) = \mu_-(E^+) = 0$ and hence, by the new hypotheses,

$$K\mu\left(E^+\right) = K\mu_+(E) - K\mu_-\left(E^+\right) - K\mu_+\left(E^-\right) \leqslant 0.$$

For more details on Hahn-Jordan decompositions see e.g. Chapter 14 in [Zaanen (1997)]. The following theorem is the main motivation to introduce (sub-)Kolmogorov operators K.

Theorem 8.1. *Let K be an sub-Kolmogorov operator as in Definition 8.2. Then, for every $\lambda > 0$ and $\mu \in D(K)$, the following inequalities hold:*

$$\lambda \sup_{B\in\mathcal{E}} \Re\mu(B) \leqslant \sup_{B\in\mathcal{E}} \Re\left(\lambda I - K\right)\mu(B); \tag{8.24}$$

$$\lambda \inf_{B\in\mathcal{E}} \Re\mu(B) \geqslant \inf_{B\in\mathcal{E}} \Re\left(\lambda I - K\right)\mu(B); \tag{8.25}$$

$$\lambda \sup_{B\in\mathcal{E}} \left|\mu(B)\right| \leqslant \sup_{B\in\mathcal{E}} \left|\left(\lambda I - K\right)\mu(B)\right|. \tag{8.26}$$

Proof. [Proof of Theorem 8.1.] First we notice the equality:

$$\sup_{B\in\mathcal{E}} \Re\mu(B) = \sup\left\{\Re\langle f,\mu\rangle : 0 \leqslant f \leqslant 1,\ f \in C_b(E)\right\}.$$

Assertion (8.25) is a consequence of (8.24): apply (8.24) with $-\mu$ replacing μ. Assertion (8.26) also follows from (8.24) by noticing that

$$\left|\langle f,\mu\rangle\right| = \sup_{\vartheta\in[-\pi,\pi]} \Re\left\langle f, e^{i\vartheta}\mu\right\rangle,$$

and then applying (8.24) to the measures $e^{i\vartheta}\mu$ and $\vartheta \in [-\pi,\pi]$. The inequality in (8.24) remains to be shown. Fix $\mu \in M(E)$ and $f \in C_b(E)$, $0 \leqslant f \leqslant \mathbf{1}$. Then we have

$$\lambda\Re\langle f,\mu\rangle = \Re\langle f,(\lambda I - K)\mu\rangle + \Re\langle f,K\mu\rangle. \tag{8.27}$$

From (8.27) we get

$$\begin{aligned}
&\lambda \sup\left\{\Re\langle f,\mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E),\ \Re\langle f,K\mu\rangle \leqslant \varepsilon\right\}\\
&\quad\leqslant \sup\left\{\Re\langle f,(\lambda I - K)\mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E),\ \Re\langle f,K\mu\rangle \leqslant \varepsilon\right\} + \varepsilon\\
&\quad\leqslant \sup\left\{\Re\langle f,(\lambda I - K)\mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E)\right\} + \varepsilon.
\end{aligned} \tag{8.28}$$

Employing equality (8.11) in Definition 8.2 and (8.28) we get

$$\begin{aligned}
&\lambda \sup\left\{\Re\langle f,\mu\rangle : f \in C_b(E), 0 \leqslant f \leqslant \mathbf{1}\right\}\\
&= \inf_{\varepsilon>0} \sup\left\{\Re\langle f,\mu\rangle : 0 \leqslant f \leqslant 1,\ \Re\langle f,K\mu\rangle \leqslant \varepsilon,\ f \in C_b(E)\right\}\\
&\leqslant \sup\left\{\Re\langle f,(\lambda I - K)\mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E)\right\}.
\end{aligned} \tag{8.29}$$

Inequality (8.24) follows from (8.29). This concludes the proof of Theorem 8.1. □

8.3 Kolmogorov operators and analytic semigroups

In the present section we recall some properties of weak*-continuous bounded analytic semigroups acting on $M(E)$. These results have their counterparts for strongly continuous bounded analytic semigroups.

Theorem 8.2. *Suppose, in addition to the fact that K is a sectorial sub-Kolmogorov operator, that there exists $\lambda_0 > 0$ such that $(\lambda_0 I - K)\, D(K) = M(E)$. Then for every real-valued function $f \in C_b(E)$ and every $\mu \in D(K)$ with values in $\mathbb{R}$ the expression $\langle f, K\mu\rangle$ is real. Assume that the graph of the operator K is $\sigma(M(E), C_b(E))$-closed, and that the same is true for all operators $\mu \mapsto \mathbf{1}_C K\mu$, $\mu \in D(K)$, where C is any compact subset of E. Here the measure $\mathbf{1}_C K\mu$ is defined by the equality $\langle f, \mathbf{1}_C K\mu\rangle = \int_C f dK\mu$, $f \in C_b(E)$. Moreover, there exists a finite constant C such that for every $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$ the following assertions hold:*

(1) $(\lambda I - K)\, D(K) = M(E)$.
(2) Let $\mu \in D(K)$ be a real-valued measure on $\mathcal{E}$. Then

$$\begin{aligned}&|\lambda| \sup\{|\langle f, \mu\rangle| : 0 \leqslant f \leqslant \mathbf{1},\, f \in C_b(E)\}\\ &\leqslant \sup\{|\langle f, (\lambda I - K)\, \mu\rangle| : 0 \leqslant f \leqslant \mathbf{1},\; f \in C_b(E)\}.\end{aligned} \tag{8.30}$$

(3) The inequality

$$\begin{aligned}&|\lambda| \sup\{|\langle f, \mu\rangle| : |f| \leqslant \mathbf{1},\, f \in C_b(E)\}\\ &\leqslant C \sup\{|\langle f, (\lambda I - K)\, \mu\rangle| : |f| \leqslant \mathbf{1},\, f \in C_b(E)\}\end{aligned} \tag{8.31}$$

holds for all measures $\mu \in D(K)$.
(4) Suppose that the function $x \mapsto (\lambda I - K)^{-1}\delta_x$, $x \in E$, is Borel measurable. Let μ be a bounded Borel measure on E. Then the following equality holds:

$$\int \lambda\, (\lambda I - K)^{-1}\, \delta_x\, d\mu(x) = \lambda\, (\lambda I - K)^{-1}\, \mu. \tag{8.32}$$

Proof. [Proof of Theorem 8.2.] First we will show the following assertion. If a function $f \in C_b(E)$ and a measure $\mu \in D(K)$ are real-valued, then the expression $\langle f, K\mu\rangle$ belongs to $\mathbb{R}$. For this purpose we choose measures $\nu_\lambda \in M(E)$, $\lambda > 0$, such that $\lambda\mu = (\lambda I - K)\, \nu_\lambda$. Then $\lambda\, (i\mu) = (\lambda I - K)\, (i\nu_\lambda)$. By (8.24) in Theorem 8.1 we have for $B \in \mathcal{E}$

$$-\lambda\Im\nu_\lambda(B) = \lambda\Re(i\nu_\lambda(B)) \geqslant \inf_{B\in\mathcal{E}} \Re(\lambda I - K)(i\nu_\lambda)(B) = \inf_{B\in\mathcal{E}} \lambda\Re(i\mu(B)) = 0. \tag{8.33}$$

From (8.33) it follows that $\Im\nu_\lambda(B) \leqslant 0$ for all $B \in \mathcal{E}$. By the same procedure with $-\mu$ instead of μ we see $\Im\nu_\lambda(B) \geqslant 0$ for all $B \in \mathcal{E}$. Hence we get $\Im\nu_\lambda(B) = 0$ for all $B \in \mathcal{E}$, or, what is the same, the measures $\nu_\lambda = \lambda R(\lambda)\mu$, $\lambda > 0$, take their values in the reals. From (8.26) it follows that $\|\mathbf{1}_C\lambda R(\lambda)K\mu\| \leqslant \|K\mu\|$, $\lambda > 0$, C compact subset of E. Let $(C_k)_{k\in\mathbb{N}}$ be increasing sequence of compact subsets of E such that $\lim_{k\to\infty} |K\mu|(C_k) = |K\mu|(E)$. By the theorem of Banach-Alaoglu, which states the closed dual unit ball in a dual Banach space is weak*-compact, it follows that there exists a double sequence $\{\lambda_{k,n} : k, n \in \mathbb{N}\}$ such that for every fixed k $\lambda_{k,n}$ tends to ∞ as $n \to \infty$, and measures $\nu_k \in M(E)$ such that

$$\lim_{n\to\infty} \langle f, \mathbf{1}_{C_k} K(\lambda_{k,n} R(\lambda_{k,n})\mu)\rangle = \lim_{n\to\infty} \langle f, \mathbf{1}_{C_k}\lambda_{k,n} R(\lambda_{k,n}) K\mu\rangle = \langle f, \nu_k\rangle, \tag{8.34}$$

$f \in C_b(E)$. Since $\lambda_{k,n}R(\lambda_{k,n})\mu - \mu = R(\lambda_{k,n})K\mu$ inequality (8.26) implies

$$\lambda_{k,n}\|\lambda_{k,n}R(\lambda_{k,n})\mu - \mu\| \leqslant \|K\mu\|,$$

we see that

$$\lim_{n\to\infty} \|\lambda_{k,n}R(\lambda_{k,n})\mu - \mu\| = 0. \tag{8.35}$$

From (8.34) and (8.35) it follows that the pair (μ, ν_k) belongs to the closure of $G(\mathbf{1}_{C_k}K)$ in the space $(M(E), \|\cdot\|) \times (M(E), \sigma(M(E), C_b(E)))$. Since by assumption the subspace $G(\mathbf{1}_{C_k}K)$ is closed for this topology we see that $\nu_k = \mathbf{1}_{C_k}K\mu$, and hence $\mathbf{1}_{C_k}K\mu$ being the $\sigma(M(E), C_b(E))$-limit of a sequence of real measures is itself a real-valued measure. Since $\langle f, K\mu\rangle = \lim_{k\to\infty}\langle f, \mathbf{1}_{C_k}K\mu\rangle$ we see that $K\mu$ is a real measure.

(1). As a second step we prove assertion (1), i.e. we show that for every $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$ the equality $(\lambda I - K)D(K) = M(E)$ holds. Therefore we put

$$R(\lambda_0) = (\lambda_0 I - K)^{-1}, \quad \text{and} \quad R(\lambda) = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k R(\lambda_0)^{k+1}.$$

By the inequality (8.26) this series converges for λ in the open disc

$$\{\lambda \in \mathbb{C} : C|\lambda - \lambda_0| < \lambda_0\}.$$

Moreover, for such λ we have $(\lambda I - K)R(\lambda) = I$ (and $R(\lambda)(\lambda I - K)$ is the identity on $D(K)$). Next, consider the subset of $\mathbb{C}$ defined by

$$\{\lambda \in \mathbb{C} : \Re\lambda > 0,\ (\lambda I - K)D(K) = M(E)\}. \tag{8.36}$$

Then the set in (8.36) is open and closed in the half plane $\{\lambda \in \mathbb{C} : \Re\lambda > 0\}$. Hence it coincides with the half-plane $\{\lambda \in \mathbb{C} : \Re\lambda > 0\}$. It follows that there exists a family of bounded linear operators $R(\lambda)$, $\Re\lambda > 0$, such that $R(\lambda) = (\lambda I - K)^{-1}$. Note that in this construction we equipped the space $M(E)$ with the norm $\|\mu\| = \sup\{|\langle f, \mu\rangle| : 0 \leqslant f \leqslant \mathbf{1}\}$. Altogether this proves (1).

(2). We fix $0 \leqslant f \leqslant \mathbf{1}$, $f \in C_b(E)$, and $\mu \in D(K)$, $\mu(B) \in \mathbb{R}$, $B \in \mathcal{E}$. Then $\langle f, K\mu\rangle$ belongs to $\mathbb{R}$, and for an appropriate choice of $\vartheta \in [-\pi/2, \pi/2]$ we have

$$\begin{aligned} |\lambda| \langle f, \mu\rangle &= \Re\left\langle f, \lambda e^{i\vartheta}\mu\right\rangle = \Re\left\langle f, (\lambda I - K)\, e^{i\vartheta}\mu\right\rangle + \Re\left\langle f, K\left(e^{i\vartheta}\mu\right)\right\rangle \\ &= \Re\left\langle f, (\lambda I - K)\, e^{i\vartheta}\mu\right\rangle + \cos\vartheta \langle f, K\mu\rangle . \end{aligned} \tag{8.37}$$

From (8.37) and equality (8.11) in Definition 8.2 we infer:

$$\begin{aligned} &|\lambda| \sup\{\langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\, f \in C_b(E)\} \\ &= |\lambda| \inf_{\varepsilon>0} \sup\{\langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\, \langle f, K\mu\rangle \leqslant \varepsilon,\, f \in C_b(E)\} \\ &\leqslant \inf_{\varepsilon>0} \sup\left\{\Re\left\langle f, (\lambda I - K)\left(e^{i\vartheta}\mu\right)\right\rangle + \cos\vartheta \langle f, K\mu\rangle :\ 0 \leqslant f \leqslant \mathbf{1},\right. \\ &\qquad \left.\langle f, K\mu\rangle \leqslant \varepsilon,\, f \in C_b(E)\right\} \\ &\leqslant \inf_{\varepsilon>0} \left(\sup\{|\langle f, (\lambda I - K)\,\mu\rangle| : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E)\} + \varepsilon\right). \end{aligned} \tag{8.38}$$

From (8.38) we infer

$$\begin{aligned} &|\lambda| \sup\{\langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1},\, f \in C_b(E)\} \\ &\leqslant \sup\{|\langle f, (\lambda I - K)\,\mu\rangle| : 0 \leqslant f \leqslant \mathbf{1},\ f \in C_b(E)\} . \end{aligned} \tag{8.39}$$

The conclusion in (8.30) of item (2) of Theorem 8.2 now follows by applying (8.39) to the real measures μ and $-\mu$.

(3) The inequality in (8.31) is the same as (8.12) in Definition 8.2.

(4) Let μ be a bounded Borel measure on E, and let $\lambda > 0$. Then we want to show the equality in (8.32). Therefore we put $\nu_x = \lambda\,(\lambda I - K)^{-1}\,\delta_x$, $x \in E$. So that $\nu_x \in D(K)$ and $(\lambda I - K)\,\nu_x = \lambda\delta_x$. Then since the operator K is $\sigma\,(C_b(E), M(E))$-closed we see

$$\lambda\mu = \lambda \int \delta_x\, d\mu(x) = \int (\lambda I - K)\,\nu_x\, d\mu(x) = (\lambda I - K) \int \nu_x\, d\mu(x), \tag{8.40}$$

and consequently,

$$\lambda\,(\lambda I - K)^{-1}\,\mu = \int \nu_x\, d\mu(x) = \int \lambda\,(\lambda I - K)^{-1}\,\delta_x\, d\mu(x). \tag{8.41}$$

The equality in (8.41) is the same as the one in (8.32). The final step in (8.40) can be justified as follows. We choose double sequences

$$\{x_{j,n} : n \in \mathbb{N},\, 1 \leqslant j \leqslant N_n\} \subset E \quad \text{and} \quad \{C_{j,n} : n \in \mathbb{N},\, 1 \leqslant j \leqslant N_n\} \subset \mathcal{E}$$

such that

$$\langle f,\mu\rangle = \lim_{n\to\infty}\langle f,\mu_n\rangle \quad\text{and}\quad \left\langle f,\int \nu_x\,d\mu(x)\right\rangle = \lim_{n\to\infty}\langle f,\nu_n\rangle\,,\quad f\in C_b(E), \tag{8.42}$$

where $\langle f,\mu_n\rangle = \sum_{j=1}^{N_n}\mu\left(C_{j,n}\right)f\left(x_{j,n}\right)$, and $\langle f,\nu_n\rangle = \sum_{j=1}^{N_n}\mu\left(C_{j,n}\right)\int f\,d\nu_{x_{j,n}}$. Here we employ the Borel measurability of the function $x\mapsto(\lambda I-K)^{-1}\delta_x$. As a consequence of (8.42) we infer that

$$\sigma\left(M(E),C_b(E)\right)\text{-}\lim_{n\to\infty}\nu_n = \int\nu_x\,d\mu(x)\quad\text{and} \tag{8.43}$$

$$\sigma\left(M(E),C_b(E)\right)\text{-}\lim_{n\to\infty}(\lambda I-K)\,\nu_n = \lambda\sigma\left(M(E),C_b(E)\right)\text{-}\lim_{n\to\infty}\mu_n = \lambda\mu.$$

Since the graph of the operator K is $\sigma\left(M(E),C_b(E)\right)$-closed, the equalities in (8.43) imply that the measure $B\mapsto\int\nu_x(B)\,d\mu(x)$, $B\in\mathcal{E}$, belongs to $D(K)$ and that $\int(\lambda I-K)\,\nu_x\,d\mu(x) = (\lambda I-K)\left(\int\nu_x\,d\mu(x)\right)$, which is the same as (8.40).

This completes the proof of assertion (4), and also of Theorem 8.2 □

Corollary 8.1. *Let the sectorial sub-Kolmogorov operator K in Theorem 8.1 have the additional property that for some $\lambda_0\in\mathbb{C}$, with $\lambda_0>0$, the range of $\lambda_0 I-K$ coincides with $M(E)$. Then for all $\lambda\in\mathbb{C}$ with $\Re\lambda>0$ the operator $(\lambda I-K)^{-1}$ exists as a bounded linear operator which is defined on all of $M(E)$, and which satisfies*

$$|\lambda|\,\mathrm{Var}\left((\lambda I-K)^{-1}\mu\right)\leqslant C\mathrm{Var}\left(\mu\right),\quad \Re\lambda>0,\ \mu\in M(E). \tag{8.44}$$

Here Var(μ) *stands for the total variation norm of the measure μ; it satisfies*

$$\mathrm{Var}\left(\mu\right) = \sup\left\{|\langle f,\mu\rangle| : |f|\leqslant\mathbf{1}\right\}.$$

Proof. [Proof of Corollary 8.1.] From assertion (1) and (2) in Theorem 8.2 it follows that the inverse operators $(\lambda I-K)^{-1}$, $\Re\lambda>0$, exist as continuous linear operators. Then the inequality in (8.31) implies that

$$|\lambda|\,\mathrm{Var}\left(\mu\right)\leqslant C\mathrm{Var}\left((\lambda I-K)\,\mu\right),\quad \Re\lambda>0,\ \mu\in M(E). \tag{8.45}$$

The inequality in (8.44) follows from the inequality in (8.45). The representation of the operator e^{tK} given in (8.47) is explained in (the proof of) Theorem 8.8 (see equality (8.287)). □

Proposition 8.1. *Operators K which have weak*-dense domain and which satisfy (8.44) generate weak*-continuous analytic semigroups*

$$\left\{e^{tK} : |\arg t| \leqslant \alpha\right\}, \quad \textit{for some } 0 < \alpha < \frac{\pi}{2}.$$

The operators $t^{\ell+1}e^{tK}$ and $(-t)^{\ell}K^{\ell}e^{tK}$, $t > 0$, $\ell \in \mathbb{N}$, have the representations

$$\frac{t^{\ell+1}}{(\ell+1)!}e^{tK} = \frac{1}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty}\left(e^{t\lambda}+e^{-t\lambda}-2\right)(\lambda I-K)^{-\ell-2}\,d\lambda, \tag{8.46}$$

and

$$\begin{aligned}&\frac{(-t)^{\ell}}{(\ell+1)!}K^{\ell}e^{tK} &(8.47)\\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\xi}{\xi^2}\left\{I-2i\xi\left(2i\xi I-tK\right)^{-1}\right\}^{\ell}\left(2i\xi\left(2i\xi I-tK\right)^{-1}\right)^2 d\xi,\end{aligned}$$

respectively. Consequently, with $C(0) = \sup\left\{|\lambda|\left\|(\lambda I-K)^{-1}\right\| : \Re\lambda > 0\right\}$ and with $C_1(0) = \sup\left\{\left\|I-\lambda(\lambda I-K)^{-1}\right\| : \Re\lambda > 0\right\}$, the following inequality holds:

$$\frac{\left\|t^{\ell}K^{\ell}e^{tK}\right\|}{\ell!} \leqslant (\ell+1)\,C(0)^2C_1(0)^{\ell}, \quad t \geqslant 0,\ \ell \in \mathbb{N}. \tag{8.48}$$

For $\ell = 0$ formula (8.46) can be rewritten as:

$$e^{tK} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\xi}{\xi^2}\left(2i\xi\left(2i\xi I-tK\right)^{-1}\right)^2 d\xi. \tag{8.49}$$

The formula in (8.49) can be used to define the semigroup e^{tK}, $t \geqslant 0$. For $\ell = 1$ formula (8.47) reduces to

$$-tKe^{tK} = \frac{2}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\xi}{\xi^2}\left\{\left(2i\xi(2i\xi I-tK)^{-1}\right)^2-\left(2i\xi(2i\xi I-tK)^{-1}\right)^3\right\}d\xi. \tag{8.50}$$

Proof. From Cauchy's theorem it follows that the right-hand side of (8.46) multiplied by $\dfrac{(\ell+1)!}{t^{\ell+1}}$ is equal to

$$\frac{(\ell+1)!}{2\pi i t^{\ell+1}}\int_{\omega-i\infty}^{\omega+i\infty}e^{t\lambda}(\lambda I-K)^{-\ell-2}\,d\lambda = \frac{(\ell+1)!}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty}e^{\lambda}(\lambda I-tK)^{-\ell-2}\,d\lambda. \tag{8.51}$$

Integration by parts shows that the right-hand side of (8.51) does not depend $\ell \in \mathbb{N}$, and hence

$$\frac{(\ell+1)!}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty}e^{\lambda}(\lambda I-tK)^{-\ell-2}\,d\lambda = \frac{1}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty}e^{\lambda}(\lambda I-tK)^{-2}\,d\lambda. \tag{8.52}$$

The right-hand side of (8.52) is the inverse Laplace transform at $s = 1$ of the function $s \mapsto se^{stK}$ and thus it is equal to e^{tK}. This shows (8.46). Since $I - \lambda\left(\lambda I - K\right)^{-1} = -K\left(\lambda I - K\right)^{-1}$ the equality in (8.46) entails:

$$\begin{aligned}&\frac{(-t)^\ell}{(\ell+1)!}K^\ell e^{tK}\\&= \frac{1}{2\pi i t}\int_{\omega-i\infty}^{\omega+i\infty}\frac{e^{t\lambda}+e^{-t\lambda}-2}{\lambda^2}\left(I-\lambda\left(\lambda I-K\right)^{-1}\right)^\ell\lambda^2\left(\lambda I-K\right)^{-2}d\lambda\\&= \frac{1}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty}\frac{e^{\lambda}+e^{-\lambda}-2}{\lambda^2}\left(I-\lambda\left(\lambda I-tK\right)^{-1}\right)^\ell\lambda^2\left(\lambda I-tK\right)^{-2}d\lambda,\end{aligned} \tag{8.53}$$

and hence (8.47) follows. The inequality in (8.48) follows immediately from (8.47). The equalities in (8.49) and (8.50) are easy consequences of (8.46) and (8.47) respectively.

Altogether this proves Proposition 8.1. □

Lemma 8.2. *Suppose that for $\Re\lambda > 0$ the operator $\lambda I - K$ has a bounded inverse defined on $M(E)$. Suppose that $C(0)$ defined by*

$$C(0) := \sup\left\{\left\|\lambda\left(\lambda I - K\right)^{-1}\right\| : \Re\lambda > 0\right\} \tag{8.54}$$

is finite. Let $0 < \alpha < \frac{1}{2}\pi$ be such that $2C(0)\sin\left(\frac{1}{2}\alpha\right) < 1$. Then for $\lambda \in \mathbb{C}$ with the property that $|\arg(\lambda)| < \frac{1}{2}\pi + \alpha$ the operator $\lambda I - K$ has a bounded inverse with the property that

$$|\lambda|\left\|\left(\lambda I - K\right)^{-1}\right\| \leqslant C(\alpha), \quad |\arg(\lambda)| \leqslant \frac{1}{2}\pi + \alpha,$$

where

$$C(\alpha) := \sup\left\{\left\|\lambda\left(\lambda I - K\right)^{-1}\right\| : |\arg(\lambda)| \leqslant \frac{1}{2}\pi + \alpha\right\}. \tag{8.55}$$

If $0 \leqslant 2\sin\left(\frac{1}{2}\alpha\right)C(0) < 1$, then $C(\alpha) < \infty$, and

$$C(\alpha) \leqslant \frac{C(0)}{1 - 2\sin\left(\frac{1}{2}\alpha\right)C(0)}. \tag{8.56}$$

In addition, the analytic semigroup e^{sK}, $|\arg(s)| \leqslant \alpha$, can be defined by the same formula as employed in (8.49):

$$e^{sK}\mu = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\xi}{\xi^2}\left(2i\xi\right)^2\left(2i\xi I - sK\right)^{-2}\mu\, d\xi,\ \ \mu \in M(E), \quad |\arg(s)| \leqslant \alpha, \tag{8.57}$$

and hence

$$\left\|e^{sK}\right\| \leqslant C(\alpha)^2 \leqslant \frac{C(0)^2}{\left(1 - 2\left|\sin\frac{1}{2}\alpha\right|C(0)\right)^2}, \quad |\arg(s)| \leqslant \alpha. \tag{8.58}$$

Proof. Fix $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$, and observe the equality

$$\lambda\left(\lambda I - e^{-i\alpha}K\right)^{-1} = e^{i\alpha}\lambda\left(\lambda I - K\right)^{-1}\left(I - \left(1-e^{i\alpha}\right)\lambda\left(\lambda I - K\right)^{-1}\right)^{-1}$$
$$= e^{i\alpha}\lambda\left(\lambda I - K\right)^{-1}\sum_{j=0}^{\infty}\left(1-e^{i\alpha}\right)^{j}\left(\lambda\left(\lambda I - K\right)^{-1}\right)^{j}. \tag{8.59}$$

The inequality in (8.56) then follows from (8.59). The equality in (8.57) follows from (8.49) and the fact that the vector-valued functions in the right-hand side and the left-hand side of (8.57) are holomorphic in s on an open neighborhood of the indicated sector in $\mathbb{C}$.

This proves Lemma 8.2. □

Proposition 8.2. *The powers of the resolvent operators* $(\lambda I - K)^{-j-1}$ *have the representation*

$$\lambda^{j+1}\left(\lambda I - K\right)^{-j-1}\mu = \frac{\left(\lambda e^{-i\alpha}\right)^{j+1}}{j!}\int_0^{\infty} s^j e^{-se^{-i\alpha}\lambda}e^{se^{-i\alpha}K}\mu\,ds, \quad j \in \mathbb{N}, \tag{8.60}$$

where $0 < \alpha < \frac{1}{2}\pi$ *if* $\Im\lambda \geqslant 0$ *and* $\Re\lambda > 0$, *and* $0 > \alpha > -\frac{1}{2}\pi$ *if* $\Im\lambda \leqslant 0$ *and* $\Re\lambda > 0$. *Next choose* $0 \leqslant \alpha' < \alpha < \frac{1}{2}\pi$ *in such a way that* $0 \leqslant 2\sin\left(\frac{1}{2}\alpha\right)C(0) < 1$. *In addition the following estimate holds for all* $j \in \mathbb{N}$ *and for all* $\lambda \in \mathbb{C}$ *with* $|\arg(\lambda)| \leqslant \frac{1}{2}\pi + \alpha' < \frac{1}{2}\pi + \alpha$*:*

$$|\lambda|^{j+1}\left\|\left(\lambda I - K\right)^{-j-1}\right\| \leqslant \frac{1}{\left(\cos\left(|\arg\lambda| - \alpha\right)\right)^{j+1}}\frac{C(0)^2}{\left(1 - 2\sin\left(\frac{1}{2}\alpha\right)C(0)\right)^2}$$
$$\leqslant \frac{1}{\left(\sin\left(\alpha - \alpha'\right)\right)^{j+1}}\frac{C(0)^2}{\left(1 - 2\sin\left(\frac{1}{2}\alpha\right)C(0)\right)^2}. \tag{8.61}$$

Proof. Let $C(\alpha)$ be as in (8.55) and suppose $C(\alpha) < \infty$. Then the measure $e^{se^{-i\alpha}K}\mu$ has the representation:

$$e^{se^{-i\alpha}K}\mu = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\xi}{\xi^2}\left(2e^{i\alpha}i\xi\right)^2\left(2e^{i\alpha}i\xi I - sK\right)^{-2}\mu\,d\xi. \tag{8.62}$$

From (8.62) and (8.57) the following estimate is obtained:

$$\left\|e^{se^{-i\alpha}K}\right\| \leqslant C(\alpha)^2 \leqslant \frac{C(0)^2}{\left(1 - 2C(0)\left|\sin\frac{1}{2}\alpha\right|\right)^2}, \quad s > 0, \tag{8.63}$$

where $C(0)$ is defined in (8.55). From (8.63) and (8.60) we see that the following estimate holds for all $j \in \mathbb{N}$ and for all $\lambda \in \mathbb{C}$ with $|\arg(\lambda)| \leqslant$

$\frac{1}{2}\pi + \alpha' < \frac{1}{2}\pi + \alpha$:

$$|\lambda|^{j+1} \left\| (\lambda I - K)^{-j-1} \right\| \leqslant \frac{1}{(\cos(|\arg\lambda| - \alpha))^{j+1}} \frac{C(0)^2}{\left(1 - 2\sin\left(\frac{1}{2}\alpha\right) C(0)\right)^2}. \tag{8.64}$$

It is clear that (8.64) implies (8.61).

This completes the proof of Proposition 8.2. □

Proposition 8.3. *Let the constants C_0 and C_1 be such that $C_1 \geqslant 1$ and*

$$\left\| t^\ell K^\ell e^{tK} \right\| \leqslant (\ell+1)! C_0^2 C_1^\ell, \quad t \geqslant 0, \ \ell \in \mathbb{N}. \tag{8.65}$$

Then the following inequality is valid:

$$|\lambda| \left\| (\lambda I - K)^{-1} \right\| \leqslant \frac{27}{4} \frac{6}{\sqrt{35}} C_0^2 C_1, \quad \Re\lambda > 0. \tag{8.66}$$

Proof. Suppose that $|\arg(s)| \leqslant \alpha$ where α satisfies $0 \leqslant 2C_1 \sin\left(\frac{1}{2}\alpha\right) < 1$. Then the measure $e^{sK}\mu$ can be written as

$$e^{sK}\mu = e^{(s-|s|)K} e^{|s|K}\mu = \sum_{\ell=0}^{\infty} \frac{\left(\frac{s}{|s|} - 1\right)^\ell}{\ell!} (|s|\, K)^\ell e^{|s|K}\mu, \tag{8.67}$$

and the representation (8.67) together with (8.65) implies the inequality:

$$\left\| e^{sK} \right\| \leqslant \frac{C_0^2}{\left(1 - 2C_1 \sin\left(\frac{1}{2}\alpha\right)\right)^2}, \quad |\arg(s)| \leqslant \alpha, \tag{8.68}$$

provided $0 < 2C_1 \sin\left(\frac{1}{2}\alpha\right) < 1$. Again the representation in (8.60) ia available. The inequality in (8.61) is replaced with

$$\begin{aligned} |\lambda|^{j+1} \left\| (\lambda I - K)^{-j-1} \right\| &\leqslant \frac{1}{(\cos(|\arg\lambda| - \alpha))^{j+1}} \frac{C_0^2}{\left(1 - 2\sin\left(\frac{1}{2}\alpha\right) C_1\right)^2} \\ &\leqslant \frac{1}{(\sin(\alpha - \alpha'))^{j+1}} \frac{C_0^2}{\left(1 - 2\sin\left(\frac{1}{2}\alpha\right) C_1\right)^2}, \end{aligned} \tag{8.69}$$

provided $|\arg(\lambda)| \leqslant \frac{1}{2}\pi + \alpha' < \frac{1}{2}\pi + \alpha$. Since $2(j+3)C_1 > j+1$, the angle α can be chosen in such a way that $2(j+3)C_1 \sin\left(\frac{1}{2}\alpha\right) = j+1$ to obtain the estimate (note that $C_1 \geqslant 1$ and take $\alpha' = 0$):

$$\begin{aligned} &|\lambda|^{j+1} \left\| (\lambda I - K)^{-j-1} \right\| \\ &\leqslant \frac{1}{4} \frac{(j+3)^{j+3}}{(j+1)^{j+1}} \frac{2^{j+1} C_1^{j+1} (j+3)^{j+1}}{\left(4C_1^2 (j+3)^2 - (j+1)^2\right)^{(j+1)/2}} C_0^2 C_1^{j+1} \\ &\leqslant \frac{1}{4} \frac{(j+3)^{j+3}}{(j+1)^{j+1}} \frac{2^{j+1} (j+3)^{j+1}}{\left(4(j+3)^2 - (j+1)^2\right)^{(j+1)/2}} C_0^2 C_1^{j+1} \end{aligned} \tag{8.70}$$

for $\Re\lambda > 0$. If $j = 0$ (8.70) reduces to (8.66). This shows Proposition 8.3. □

Remark 8.3. The assumption that $C_1 \geqslant 1$ in the inequalities in (8.65) is not too surprising. In fact from (8.65) it follows that $C_1 \geqslant 1$ (by the spectral mapping theorem).

Corollary 8.2. *Let the operator L be the generator of a $\mathcal{T}_\beta$-continuous Feller semigroup in $C_b(E)$. Suppose that L is sectorial. Then its adjoint $K = L^*$ is a sectorial sub-Kolmogorov operator like in Definition 8.2. Moreover, the graph of the operator K is weak*-closed and K generates a weak*-continuous bounded analytic semigroup on $M(E)$.*

Proof. As in Theorem 8.1 K has the additional property that for some $\lambda_0 \in \mathbb{C}$, with $\lambda_0 > 0$, the range of $\lambda_0 I - K$ coincides with $M(E)$. Then for all $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$ the operator $(\lambda I - K)^{-1}$ exists as a bounded linear operator which is defined on all of $M(E)$. Since L is sectorial it follows that for the operator L the following inequality holds for all $\lambda \in \mathbb{C}$ with $\Re\lambda \geqslant 0$ and for all $f \in D(L)$:

$$|\lambda| \|f\|_\infty \leqslant C \|(\lambda I - L) f\|_\infty . \tag{8.71}$$

Of course, from (8.71) we see:

$$|\lambda| \left\| (\lambda I - L)^{-1} f \right\|_\infty \leqslant C \|f\|_\infty , \quad f \in C_b(E),\ \Re\lambda > 0. \tag{8.72}$$

From (8.72) we obtain, by duality,

$$|\lambda| \operatorname{Var}\big((\lambda I - K)^{-1} \mu\big) \leqslant C \operatorname{Var}(\mu) , \quad \mu \in M(E),\ \Re\lambda > 0. \tag{8.73}$$

We still have to prove that the operator K is a sub-Kolmogorov operator. This can be achieved as follows. Let $\Re\mu$ be the real part of the measure $\mu \in D(K)$. Then there exists a Borel subset $E^+_{\Re\mu}$ on which $\Re\mu$ is a positive measure and which has the property that

$$\sup \{\Re \langle f, \mu\rangle : 0 \leqslant f \leqslant \mathbf{1}\} = \Re\mu \left(E^+_{\Re\mu}\right). \tag{8.74}$$

Choose compact subsets C_n and open subsets O_n of E such that $C_n \subset E^+_{\Re\mu} \subset O_n$, and such that $\lim_{n\to\infty} |\Re\mu| (O_n \backslash C_n) = 0$. Since $\mu \in D(K)$ we get

$$\begin{aligned}\Re\mu \left(E^+_{\Re\mu}\right) &= \Re \left\langle \mathbf{1}_{E^+_{\Re\mu}}, \mu\right\rangle = \lim_{\lambda\to\infty} \Re \left\langle \mathbf{1}_{E^+_{\Re\mu}}, \lambda (\lambda I - K)^{-1} \mu\right\rangle \\ &= \lim_{\lambda\to\infty} \lim_{n\to\infty} \Re \left\langle f_n, \lambda (\lambda I - K)^{-1} \mu\right\rangle \\ &= \lim_{\lambda\to\infty} \lim_{n\to\infty} \Re \left\langle \lambda (\lambda I - L)^{-1} f_n, \mu\right\rangle ,\end{aligned} \tag{8.75}$$

where $\mathbf{1}_{C_n} \leqslant f_n \leqslant \mathbf{1}_{O_n}$, $f_n \in C_b(E)$. Fix $\lambda > 0$ and consider the function $g_{\lambda,n} := \lambda (\lambda I - L)^{-1} f_n$, which satisfies $0 \leqslant g_{\lambda,n} \leqslant \mathbf{1}$. Moreover, we have

$$\Re \left\langle \lambda (\lambda I - L)^{-1} f_n, K\mu\right\rangle = \Re \left\langle \lambda L (\lambda I - L)^{-1} f_n, \mu\right\rangle$$

$$
\begin{aligned}
&= \Re \left\langle \lambda^2 (\lambda I - L)^{-1} f_n - \lambda f_n, \mu \right\rangle \\
&= \Re \left\langle \mathbf{1}_{E\setminus O_n} \lambda^2 (\lambda I - L)^{-1} f_n, \mu \right\rangle + \Re \left\langle \mathbf{1}_{O_n \setminus C_n} (\lambda^2 (\lambda I - L)^{-1} f_n - \lambda f_n), \mu \right\rangle \\
&\quad + \Re \left\langle \mathbf{1}_{C_n} \left(\lambda^2 (\lambda I - L)^{-1} f_n - \lambda f_n \right), \mu \right\rangle .
\end{aligned} \tag{8.76}
$$

Since the measure $\Re\mu$ is positive on $E\setminus O_n$ and the function $g_{\lambda,n}$ is nonnegative the first term on the right-hand side of (8.76) is less than or equal to zero. The function $g_{\lambda,n}$ satisfies $g_{\lambda,n} \leqslant \mathbf{1}$ and the measure $\Re\mu$ is positive on C_n, and hence the third term in (8.76) is less than or equal to zero as well. Here we also used the fact that $f_n = 1$ on C_n. The middle term in the right-hand side of (8.76) is dominated by

$$
2\lambda \left\langle \mathbf{1}_{O_n \setminus C_n} g_{\lambda,n}, |\Re\mu| \right\rangle \leqslant 2\lambda \left|\Re\mu\right| (O_n \setminus C_n) . \tag{8.77}
$$

Inserting (8.77) in (8.76) and using the fact that the first and the third term of the right-hand side of (8.76) are dominated by 0 shows the inequality:

$$
\Re \left\langle \lambda (\lambda I - L)^{-1} f_n, K\mu \right\rangle \leqslant 2\lambda \left|\Re\mu\right| (O_n \setminus C_n) . \tag{8.78}
$$

Since $\lim_{n\to\infty} \left|\Re\mu\right| (O_n \setminus C_n) = 0$, from (8.74), (8.75), and (8.78) we infer that the operator K is a sub-Kolmogorov operator: see Definition 8.2.

The proof of Corollary 8.2 is now complete. □

Remark 8.4. In fact in Section 8.2 we will need an inequality of the form

$$
|\lambda| \operatorname{Var}(\mu) \leqslant C \operatorname{Var}((\lambda I - K)\mu), \quad \Re\lambda > 0, \ \mu \in M(E). \tag{8.79}
$$

In the presence of (8.79) the operator K generates a bounded analytic semigroup; see Theorem 8.8 below. This is the case if $K = L^*$, where L is an operator with domain and range in $C_b(E)$ with the property that

$$
|\lambda| \left\| f \right\|_\infty \leqslant C \left\| (\lambda I - L) f \right\|_\infty , \quad \Re\lambda > 0, \quad f \in D(L).
$$

The following theorem is related to a similar result for continuous function spaces rather than for measures by Cerrai (see [Cerrai (1994)] and Appendix B in [Cerrai (2001)]). In Kühnemund (see [Kühnemund (2003)]) the reader may find a generalization of such a result in the context of so-called bi-continuous semigroups. The notion of strongly continuous semigroup is replaced with bi-continuity in the sense that the convergence of semigroups is always assumed with respect to the topology $\mathcal{T}_\beta$, whereas the boundedness is always meant in the norm sense. The notion of (infinitesimal) generator is also adapted: for $\mathcal{T}_\beta$-generators convergence is considered in the $\mathcal{T}_\beta$-sense, and boundedness is phrased in terms of the norm. In the present situation the Banach space is the space of all bounded signed measures on E endowed with the variation norm and the topology $\mathcal{T}_\beta$ is the weak*-topology. A related paper is [Dorroh and Neuberger (1993)]. A result which includes Theorem 8.3 below is formulated in [Bratteli and Robinson (1987)] as Theorem 3.1.10 page 171. The strict topology is also called mixed topology, or $\mathcal{K}$-topology: see the references given in Subsection 2.3.1 of Chapter 2.

Theorem 8.3. *Let K be a weak*-closed linear operator with weak*-dense domain in $M(E)$. Suppose that K possesses the sub-Kolmogorov property in the sense of Definition 8.2. Fix $\lambda_0 > 0$ and suppose that for every $x \in E$ there exists a measure $\mu_x^{\lambda_0}$ such that*

$$\lambda_0 \delta_x = (\lambda_0 I - K)\, \mu_x^{\lambda_0}. \tag{8.80}$$

Then there exists a weak-continuous semigroup $S(t) := e^{tK}$, $t \geqslant 0$, such that*

$$\lim_{t \downarrow 0} \frac{\left\langle f, \left(e^{tK} - I\right)\mu\right\rangle}{t} = \langle f, K\mu\rangle\,, \quad \textit{for all } f \in C_b(E) \textit{ and } \mu \in D(K).$$

From Theorem 8.1 it follows that the measures $\mu_{1,x}^{\lambda_0} := \mu_x^{\lambda_0}$, $x \in E$, are sub-probability measures. If $\langle \mathbf{1}, K\mu\rangle = 0$, then these measures are probability measures.

Proof. We will show that our assumptions imply the conditions set forth in Theorem 3.1.10 of [Bratteli and Robinson (1987)]. Assertion (3) of Theorem 8.1 implies

$$\lambda \left\|\mu\right\| \leqslant \left\|(\lambda I - K)\, \mu\right\|, \quad \lambda > 0, \quad \mu \in D(K), \tag{8.81}$$

where $\|\mu\|$ denotes the norm of μ as defined in (8.2). The inequality in (8.81) is the first condition which is required to apply Theorem 3.1.10. Let μ be a measure in $M(E)$. Then by (8.80) we have

$$\lambda_0 \mu = \int_E \lambda_0 \delta_y d\mu(y) = \int_E (\lambda_0 I - K)\, \mu_y^{\lambda_0} d\mu(y) = (\lambda_0 I - K) \int_E \mu_y^{\lambda_0} d\mu(y),$$

and so the range of $\lambda_0 I - K$ coincides with $M(E)$. Hence, the result in Theorem 8.3 follows from Theorem 3.1.10 in [Bratteli and Robinson (1987)]. □

Since the operators $K(t)$, $t \geqslant t_0$, in equation (8.6) are supposed to have the Kolmogorov property, the evolution family $X(t,s)$, $t \geqslant s \geqslant t_0$, consists of Markov operators in the sense that $\langle f, X(t,t_0)\,\mu\rangle \geqslant 0$ whenever $f \in C_b(E)$ is non-negative and μ belongs $P(E)$; in addition, $\langle \mathbf{1}, X(t,t_0)\,\mu\rangle = 1$ for

$\mu \in P(E)$. Since all operators $K(t)$ are Kolmogorov it follows that $X(t, t_0)$ is Markov for all $t \geqslant t_0$. This can be seen by the following approximation argument. Fix $t_0 < T$ and put

$$K_n(t) = K\left(t_0 + (T - t_0)\, 2^{-n} \left\lfloor \frac{t - t_0}{T - t_0} 2^n \right\rfloor\right) = K\left(\varphi_n(t)\right), \tag{8.82}$$

where

$$\varphi_n(t) = t_0 + (T - t_0)\, 2^{-n} \left\lfloor \frac{t - t_0}{T - t_0} 2^n \right\rfloor.$$

Then $K_n(t) = K\left(t_0 + (T - t_0)\, j 2^{-n}\right)$ for

$$t_0 + (T - t_0)\, \frac{j}{2^n} \leqslant t < t_0 + (T - t_0)\, \frac{j+1}{2^n}.$$

Solutions to the system $\dot{\mu}(t) = K(t)\mu(t)$, $t_0 \leqslant t \leqslant T$, are approximated by solutions to the equation:

$$\dot{\mu}_n(t) = K_n(t)\mu_n(t), \quad t_0 \leqslant t \leqslant T. \tag{8.83}$$

A solution to (8.83) can be written in the form $\mu_n(t) = X_n(t, t_0)\, \mu_n(t_0)$, with

$$X_n(t, s) = e^{(t - t_{\ell,n}) K(t_{\ell,n})} \prod_{j=k+1}^{\ell-1} e^{(t_{j+1,n} - t_{j,n}) K(t_{j,n})} e^{(t_{k+1,n} - s) K(t_{k,n})}, \tag{8.84}$$

where $t_0 \leqslant s \leqslant t \leqslant T$, $t_{j,n} = t_0 + (T - t_0)\, \dfrac{j}{2^n}$, $0 \leqslant j \leqslant 2^n$, $t_{k,n} \leqslant s < t_{k+1,n}$, and $t_{\ell,n} \leqslant t < t_{\ell+1,n}$. We also need Duhamel's formula:

$$\left(X_n(t, t_0) - X_m(t, t_0)\right) \mu = \int_{t_0}^{t} X_n(t, s) \left(K_n(s) - K_m(s)\right) X_m(s, t_0)\, \mu\, ds. \tag{8.85}$$

In (8.85) we let $m \to \infty$ and use weak*-convergence to obtain:

$$\left(X_n(t, t_0) - X(t, t_0)\right) \mu = \int_{t_0}^{t} X_n(t, s) \left(K_n(s) - K(s)\right) X(s, t_0)\, \mu\, ds. \tag{8.86}$$

Of course, we assume that the sequences

$$\int_{t_0}^{t} X_n(t, s)\, K_n(s) X_m(s, t_0)\, \mu\, ds \quad \text{and} \quad \int_{t_0}^{t} X_n(t, s)\, K_m(s) X_m(s, t_0)\, \mu\, ds$$

converge in weak*-sense to

$$\int_{t_0}^{t} X_n(t, s)\, K_n(s) X(s, t_0)\, \mu\, ds \quad \text{and} \quad \int_{t_0}^{t} X_n(t, s)\, K(s) X(s, t_0)\, \mu\, ds$$

respectively as $m \to \infty$.

Theorem 8.4. *Let the sequences* $\{K_n(t) : n \in \mathbb{N}\}$ *and* $\{X_n(t, t_0) : n \in \mathbb{N}\}$ *be as in (8.82) and in (8.84). Suppose that for all* $m \in \mathbb{N}$ *and all* $t_0 \leqslant t_1 \leqslant t_2 \leqslant T$ *the measure* $X_m(t_2, t_0)\mu$ *belongs to* $D(K(t_1))$ *for all measures* $\mu \in M(E)$. *Also suppose that for every probability measure* $\mu \in M(E)$ *the family of measures* $\{K_n(t) X_m(t, t_0)\mu : t_0 \leqslant t \leqslant T, 1 \leqslant n \leqslant m\}$ *is* $\mathcal{T}_\beta$*-equi-continuous, i.e. there exists a function* $u \in H(E)$ *such that*

$$\sup_{t_0 \leqslant t \leqslant T} \sup_{n \leqslant m} |\langle f, K_n(t) X_m(t, t_0)\mu\rangle| \leqslant \|uf\|_\infty, \quad f \in C_b(E). \tag{8.87}$$

(a) *Then* $X(t, t_0)\mu := \|\cdot\|\text{-}\lim_{n\to\infty} X_n(t, t_0)\mu$ *exists and* $\mu(t) := X(t, t_0)\mu$ *satisfies:* $\dot{\mu}(t) = K(t)\mu(t)$, *provided that for all* $t_0 < s \leqslant T$

$$\lim_{t\uparrow s} \|(K(t) - K(s))\mu\| = 0 \tag{8.88}$$

for all measures $\mu \in \bigcap_{s-h<t<s} D(K(t))$ *for some* $h > 0$.

(b) *Suppose that for every* $s, t \in [t_0, T]$, $s \leqslant t$, *the sequence* $\{X_n(t, s) : n \in \mathbb{N}\}$ *is uniformly weak*-continuous, and that for all measures*

$$\mu \in \bigcap_{s-h<t<s} D(K(t))$$

the following equality holds

$$\text{weak}^*\text{-}\lim_{t\uparrow s} K(t)\mu = K(s)\mu. \tag{8.89}$$

Then $X(t, t_0)\mu := \text{weak}^*\text{-}\lim_{n\to\infty} X_n(t, t_0)\mu$, $\mu \in M(E)$, *exists and* $\mu(t) := X(t, t_0)\mu$ *satisfies:* $\dot{\mu}(t) = K(t)\mu(t)$.

For more details on $\mathcal{T}_\beta$-equi-continuous families of measures see Theorem 2.3. The sequence $\{X_n(t, s) : n \in \mathbb{N}\}$ is called uniformly weak*-continuous, if for every function $f \in C_b(E)$ and every measure $\mu \in M(E)$ the sequence of continuous functions $(s, t) \mapsto \langle f, X_n(t, s)\mu\rangle$, $t_0 \leqslant s \leqslant T$, $n \in \mathbb{N}$, is uniformly continuous. See Remark 2.4 as well.

Let $u \geqslant 0$ be a function in $H(E)$; i.e. for every $\alpha > 0$ the set $\{u \geqslant \alpha\}$ is contained in a compact subset of E. In the proof we apply the Banach-Alaoglu theorem to the effect that the collection of measures

$$B_u = \bigcap_{f \in C_b(E)} \{\mu \in M(E) : |\langle f, \mu\rangle| \leqslant \|uf\|_\infty\} \tag{8.90}$$

is $\sigma(M(E), C_b(E))$-compact: see Theorem 2.6. As a consequence, we see that every sequence in the collection B_u defined in (8.90) has a $\sigma(M(E), C_b(E))$-convergent subsequence. Here we use the fact that the space $C_b(E)$ endowed with the strict topology is separable; i.e. $C_b(E)$ contains a $\mathcal{T}_\beta$-dense countable subset.

Proof. [Proof of Theorem 8.4.] By hypothesis (8.87) both terms in the right-hand side of in Duhamel's formula (8.85) are $\mathcal{T}_\beta$-equi-continuous. So there exists a function $u \in H(E)$ such that

$$\sup_{t_0 \leqslant t \leqslant T} \sup_{n \leqslant m} \left| \left\langle f, \int_{t_0}^{t} X_n\left(t,s\right) K_n(s) X_m\left(s,t_0\right) \mu \, ds \right\rangle \right| \leqslant \left\| uf \right\|_\infty \quad \text{and} \tag{8.91}$$

$$\sup_{t_0 \leqslant t \leqslant T} \sup_{n \leqslant m} \left| \left\langle f, \int_{t_0}^{t} X_n\left(t,s\right) K_m(s) X_m\left(s,t_0\right) \mu \, ds \right\rangle \right| \leqslant \left\| uf \right\| \tag{8.92}$$

for all $f \in C_b(E)$. By the Banach-Alaoglu theorem we may assume that through a subsequence (m_j) the weak* limit in the right-hand side of (8.85) exists for all $t_0 \leqslant t \leqslant T$, and that therefore the weak* limit of the sequence $X_{m_j}\left(t,t_0\right)\mu$ exists as well for all $t_0 \leqslant t \leqslant T$. Again employing the $\mathcal{T}_\beta$-equi-continuity condition in (8.87) we may assume that, for every $n \in \mathbb{N}$, the weak* limit weak*-$\lim_{j\to\infty} K_n\left(s\right) X_{m_j}\left(s,t_0\right)\mu$ exists. Since, in addition, the operators $X_n\left(t,s\right)$ are continuous for the weak*-topology, we let $m \to \infty$ along an appropriate subsequence and use weak*-convergence to obtain:

$$(X_n(t,t_0) - X(t,t_0))\mu = \int_{t_0}^{t} X_n(t,s)(K_n(s) - K(s))X(s,t_0)\mu \, ds. \tag{8.93}$$

Our extra hypothesis (8.88) then completes the proof of assertion (a) of Theorem 8.4. The assumption that for every s, $t \in [t_0, T]$, $s \leqslant t$, the sequence $\{X_n\left(t,s\right) : n \in \mathbb{N}\}$ is uniformly weak*-continuous together with weak*-$\lim_{t\uparrow s} K(t)\mu = K(s)\mu$ completes the proof of assertion (b) of Theorem 8.4 as well. □

Remark 8.5. Under $\mathcal{T}_\beta$-equi-continuity conditions the sequences

$$X_m\left(s,t_0\right)\mu \quad \text{and} \quad K_m(s) X_m\left(s,t_0\right)\mu$$

possess subsequences which converge in weak*-sense for all $t_0 \leqslant s \leqslant T$. The Kolmogorov property of the operator function $K(t)$ entails that solutions $\mu_n(t)$ of (8.83) are non-negative, i.e. $\langle f, \mu_n(t)\rangle \geqslant 0$ for $f \geqslant 0$, $f \in C_b(E)$, and take their values in the simplex $P(E)$ for each initial condition $\mu(t_0) \in P(E)$. The latter is true because if $\langle \mathbf{1}, \mu_n(t_0)\rangle = 1$, then $\langle \mathbf{1}, \mu_n(t)\rangle = 1$ for all $T \geqslant t \geqslant t_0$. Consequently, the mappings $\mu_n\left(t_0\right) \mapsto \mu_n(t)$, $t \geqslant t_0$, leave the simplex $P(E)$ invariant, provided that $\mu(t)$ is a solution to (8.83). Passing to the limit in (8.83) yields the desired result. This passage can be justified under certain conditions. If the function $\mu_n(t)$ satisfies (8.83), then

$$\mu_n(t) = \mu_n\left(t_0\right) + \int_{t_0}^{t} K_n(s)\mu_n(s) ds.$$

Example 8.1. Let K be a weak*-closed Kolmogorov operator with weak*-dense domain, and let $p(t,x)$ be a Borel measurable strictly positive function defined on $[t_0,T] \times E$ with the property that for every $x \in E$ the function $t \mapsto p(t,x)$ is continuous. Define the families of operators $K_1(t)$ and $K_2(t)$, $t \in [t_0,T]$ by

$$K_1(t)\,\mu(B) = \int_B p(t,x)\,(K\mu)\,(dx) \quad \text{and} \quad K_2(t)\,\mu(B) = \int_B K\left(p(t,\cdot)\mu\right)(dx)$$

Suppose that K has the following property, which is somewhat stronger than the standard Kolmogorov property of Definition 8.2. For every $\mu \geqslant 0$, $\mu \in D(K)$, and every $B \in \mathcal{E}$ for which $\mu(B) = 0$ we have $K\mu(B) \geqslant 0$. Then the operators $K_1(t)$ and $K_2(t)$ share this stronger Kolmogorov property. Fix $\lambda_0 > 0$ and suppose that for every $x \in E$ there exists a measure μ_x^{t,λ_0} such that $\lambda_0\delta_x = (\lambda_0 - p(t,\cdot)\,K)\,\mu_x^{t,\lambda_0}$. Then the operator $K_1(t)$ generates a weak*-continuous semigroup: see Theorem 8.3. If for every $x \in E$ there exists a measure ν_x^{t,λ_0} such that $\lambda_0 p(t,\cdot)\,\delta_x = (\lambda_0 - p(t,\cdot)\,K)\,\nu_x^{t,\lambda_0}$. Then the measure μ_x^{t,λ_0} defined by the equality $\nu_x^{t,\lambda_0} = p(t,\cdot)\,\mu_x^{t,\lambda_0}$ satisfies: $\lambda_0\delta_x = (\lambda_0 - Kp(t,\cdot))\,\mu_x^{t,\lambda_0}$. Hence, by Theorem 8.3 the operator $K_2(t)$ generates a weak*-continuous semigroup in $M(E)$. If the function $(t,x) \mapsto p(t,x)$ is uniformly bounded, then the results of (a) in Theorem 8.4 are applicable for the family $K_1(t)$. If the domains of the operators $K_2(t)$ do not depend on $t \in [t_0,T]$, then the results of (b) in Theorem 8.4 are applicable for the family $K_2(t)$.

Example 8.2. A better example is a family of operators $K(t)$, $t \geqslant 0$, which are adjoint of operators $L(t)$ with domain and range in $C_b(E)$, i.e. $K(t) = L(t)^*$, which generate a time-dependent strong Markov process

$$\left\{\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : t \geqslant \tau\right), \left(E, \mathcal{E}\right)\right\}$$

such that

$$\frac{\partial}{\partial t}\mathbb{E}_{\tau,x}[f(t,X(t))] = \mathbb{E}_{\tau,x}[(D_1 + L(t))f(t,X(t))], \quad f \in D(D_1)\bigcap D(L(t)),$$

where $0 \leqslant \tau < t \leqslant \infty$. The operator D_1 stands for the derivative with respect to time: see Definition 2.8. We put $Y(\tau,t)\,f(x) = \mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right]$, $f \in C_b(E)$, and $X(t,\tau)\,\mu = Y(\tau,t)^*\,\mu$, $\mu \in M(E)$. This means that $\langle Y(\tau,t)\,f, \mu\rangle = \langle f, X(t,\tau)\,\mu\rangle$, $f \in C_b(E)$, $\mu \in M(E)$. Put $P(\tau,x;t,B) = \mathbb{P}_{\tau,x}\left[X(t) \in B\right]$, $0 \leqslant \tau \leqslant t < \infty$, $B \in \mathcal{E}$. Then

$$Y(\tau,t)\,f(x) = \int f(y)P(\tau,x;t,dy)\,, \quad f \in C_b(E),\ 0 \leqslant \tau \leqslant t < \infty. \tag{8.94}$$

Hence,

$$\langle f, X(t,\tau)\mu\rangle = \int f(y)\int P(\tau,x;t,dy)\,d\mu(x), \quad f\in C_b(E),\ 0\leqslant\tau\leqslant t<\infty. \tag{8.95}$$

It is assumed that for every $t\geqslant 0$ the operator $L(t)$ generates a bounded analytic Feller semigroup $e^{sL(t)}$, $|\arg s|\leqslant\alpha(t)$. In addition, assume that the operator $K(t)=L(t)^*$ has a spectral gap of width $2\omega(t)$, and that $|\lambda|\left\|(\lambda I-L(t))^{-1}\right\|\leqslant c(t)$ for $\Re\lambda\geqslant-\omega(t)$, $\lambda\neq 0$. It follows that the operators $L(t)$ generate analytic semigroups $e^{sL(t)}$ where $s\in\mathbb{C}$ belong to a sector with angle opening. Then it follows that there exist a constant $c(t)$ and an angle $\frac{1}{2}\pi<\beta(t)<\pi$ such that

$$|\lambda|\left\|(\lambda I-L(t))^{-1}\right\|\leqslant c(t), \quad \text{for all } \lambda\in\mathbb{C} \text{ with } |\arg(\lambda)|\leqslant\beta(t). \tag{8.96}$$

For a proof see Theorem 8.8 and its corollaries 8.4 and 8.5. Let $e^{sL(t)}$, $s\geqslant 0$, be the (analytic) semigroup generated by the operator $L(t)$. Then the (unbounded) inverse of the operator $-L(t)$ is given by the strong integral $f\mapsto\int_0^\infty e^{sL(t)}f\,ds$. From (8.228) it follows that for $\mu\in M_0(E)$ and $\Re\lambda>0$ the inequality

$$|\lambda|\left|\left\langle g,\left(\lambda I|_{M_0(E)}-L(t)^*|_{M_0(E)}\right)^{-1}\mu\right\rangle\right|\leqslant\|g\|_\infty\operatorname{Var}(\mu), \tag{8.97}$$

holds whenever the function g is of the form $g=\lambda f-L(t)f$, with $f\in D(L(t))$. Here $M_0(E)$ is the space of all complex Borel measures μ on E with the property that $\mu(E)=0$: see (8.5). Suppose that $\operatorname{Var}\left(e^{sL(t)^*}\mu\right)\leqslant c(t)e^{-2\omega(t)s}\operatorname{Var}(\mu)$ for all $\mu\in M_0(E)$ and $s\geqslant 0$. Then for $\Re\lambda\geqslant\omega(t)$, $g\in C_0(E)$ and $\mu\in M_0(E)$ we have

$$\begin{aligned}&(\lambda-2\omega(t))\left\langle g,\left((\lambda-2\omega(t))I|_{M_0(E)}-L(t)^*|_{M_0(E)}\right)^{-1}\mu\right\rangle\\&=(\lambda-2\omega(t))\int_0^\infty\left\langle g,e^{-s\left((\lambda-2\omega(t))I|_{M_0(E)}-L(t)^*|_{M_0(E)}\right)}\mu\right\rangle ds,\end{aligned} \tag{8.98}$$

and hence, if $|\lambda-2\omega(t)|\leqslant 2\omega(t)$ we have

$$\begin{aligned}&|\lambda-2\omega(t)|\left|\left\langle g,\left((\lambda-2\omega(t))I|_{M_0(E)}-L(t)^*|_{M_0(E)}\right)^{-1}\mu\right\rangle\right|\\&\leqslant|\lambda-2\omega(t)|\int_0^\infty\left|\left\langle g,e^{-s\left((\lambda-2\omega(t))I|_{M_0(E)}-L(t)^*|_{M_0(E)}\right)}\mu\right\rangle\right|ds\\&\leqslant|\lambda-2\omega(t)|\int_0^\infty e^{-s(\Re\lambda-2\omega(t))}\operatorname{Var}\left(e^{sL(t)^*|_{M_0(E)}}\mu\right)ds\,\|g\|_\infty\\&\leqslant c(t)\,|\lambda-2\omega(t)|\int_0^\infty e^{-s(\Re\lambda-2\omega(t))}e^{-2s\omega(t)}ds\operatorname{Var}(\mu)\,\|g\|_\infty\end{aligned}$$

$$= c(t) \frac{|\lambda - 2\omega(t)|}{\Re\lambda} \|g\|_\infty \operatorname{Var}(\mu) \leqslant 2c(t) \|g\|_\infty \operatorname{Var}(\mu). \tag{8.99}$$

In view of (8.96), (8.97) and (8.99) it makes sense to consider the largest $\omega(t)$ with the property that for all functions $g \in C_0(E)$, and all Borel measures $\mu \in M_0(E)$ the complex-valued function

$$\lambda \mapsto \lambda \left\langle g, \left(\lambda I|_{M_0(E)} - L(t)^*|_{M_0(E)}\right)^{-1} \mu \right\rangle$$

extends to a bounded holomorphic function on all half-planes of the form

$$\left\{\lambda \in \mathbb{C} : \Re\lambda > -2\omega'(t)\right\}$$

with $\omega'(t) < \omega(t)$. In follows that there exists a constant $c(t)$ such that for all functions $g \in C_b(E)$ and $\mu \in M_0(E)$ the following inequality holds:

$$|\lambda| \left| \left\langle g, \left(\lambda I|_{M_0(E)} - L(t)^*|_{M_0(E)}\right)^{-1} \mu \right\rangle \right| \leqslant c(t) \|g\|_\infty \operatorname{Var}(\mu), \quad \Re\lambda \geqslant -\omega(t).$$

The following definition is to be compared with the definitions 8.5 and 9.14 (in Chapter 9).

Definition 8.4. The number $2\omega(t)$ is called the $M(E)$-spectral gap of the operator $L(t)^*$. It is also called the uniform or L^∞-spectral gap of the operator $L(t)$.

Next let $P(\tau, x; t, B)$ be the transition probability function of the process

$$\left\{(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}), (X(t) : t \geqslant \tau), (E, \mathcal{B})\right\}$$

generated by the operators $L(t)$. Suppose that, for every $\tau \in (0, \infty)$ and every Borel probability measure on E, the following condition is satisfied:

$$\lim_{t\to\infty} \frac{c(t)}{\omega(t)} \int_E \operatorname{Var}\left(\frac{\partial}{\partial t} P(\tau, x; t, \cdot)\right) d\mu(x) = 0.$$

Let μ be any Borel probability measure on E. Put $\mu(t) = Y(\tau, t)^* \mu$, where

$$Y(\tau, t) f(x) = \mathbb{E}_{\tau,x}\left[f(X(t))\right], \quad f \in C_b(E).$$

Then $\dot\mu(t) = L(t)^*\mu(t)$. Moreover,

$$\lim_{t\to\infty} \frac{c(t)}{\omega(t)} \operatorname{Var}(\dot\mu(t)) = 0.$$

We will show this. With the above notation we have:

$$\operatorname{Var}(\dot\mu(t))$$
$$= \sup\left\{\left|\frac{d}{dt}\langle f, \mu(t)\rangle\right| : f \in C_b(E), \|f\|_\infty = 1\right\}$$

$$
\begin{aligned}
&= \sup\left\{\left|\frac{\partial}{\partial t}\left\langle Y(\tau,t)\,f,\mu\right\rangle\right| : f\in C_b(E),\ \|f\|_\infty = 1\right\}\\
&= \sup\left\{\left|\frac{\partial}{\partial t}\int_E\int_E f(y)P(\tau,x;t,dy)\,d\mu(x)\right| : f\in C_b(E),\ \|f\|_\infty = 1\right\}\\
&= \sup\left\{\left|\int_E f(y)\frac{\partial}{\partial t}\int_E P(\tau,x;t,dy)\,d\mu(x)\right| : f\in C_b(E),\ \|f\|_\infty = 1\right\}\\
&= \operatorname{Var}\left(\frac{\partial}{\partial t}\int_E P(\tau,x;t,\cdot)\,d\mu(x)\right) \leqslant \int_E \operatorname{Var}\left(\frac{\partial}{\partial t}P(\tau,x;t,\cdot)\right)\,d\mu(x).
\end{aligned}
\tag{8.100}
$$

If the probability measure $B \mapsto P(\tau,x;t,B)$ has density $p(\tau,x;t,y)$, then the total variation of the measure $B \mapsto \frac{\partial}{\partial t}P(\tau,x;t,B)$ is given by

$$
\operatorname{Var}\left(\frac{\partial}{\partial t}P(\tau,x;t,\cdot)\right) = \int_E \left|\frac{\partial}{\partial t}p(\tau,x;t,y)\right| dy. \tag{8.101}
$$

If there exists a unique $P(E)$-valued function $t \mapsto \pi(t)$ such that $L(t)^*\pi(t) = 0$, then the system $L(t)^*\mu(t) = \dot\mu(t)$ is ergodic. This assertion follows from Theorem 8.5 below.

Observe that versions of the Bismut-Elworthy formula with higher order derivatives can be used to prove that certain Feller type semigroups are analytic: see e.g. [Cerrai (2001)] Chapter 3 and Chapter 6. Section 8.7 is devoted to a discussion on this formula.

8.3.1 *Ornstein-Uhlenbeck process*

The simplest example of this kind of the process is the following one.

Example 8.3. In this example we consider the generator $L := \frac{1}{2}\Delta - x\cdot\nabla$ of the so-called Ornstein-Uhlenbeck process in $C_b\left(\mathbb{R}^d\right)$: see Theorem 1.19 assertion (d), in section E of [Demuth and van Casteren (2000)]. There exists a probability space $(\Omega,\mathcal{F},\mathbb{P})$ together with a $\mathbb{R}^d$-valued Gaussian process $\{X(s) : s \geqslant 0\}$, called *Ornstein-Uhlenbeck process*, such that $\mathbb{E}(X(s)) = 0$ and such that

$$
\mathbb{E}\left(X_j(s_1)X_k(s_2)\right) = \frac{1}{2}\exp\bigl(-(s_1+s_2)\bigr)\left(\exp\bigl(2\min(s_1,s_2)\bigr) - 1\right)\delta_{j,k} \tag{8.102}
$$

$$
= \frac{1}{2}\left(\exp\left(-|s_1-s_2|\right) - \exp\bigl(-(s_1+s_2)\bigr)\right)\delta_{j,k}, \tag{8.103}
$$

for all s_1, $s_2 \geqslant 0$, and for $1 \leqslant j$, $k \leqslant d$. Put $X^x(t) = \exp(-t)x + X(t)$. Then the process $\{X^x(t) : t \geqslant 0\}$ is the *Ornstein-Uhlenbeck process of initial velocity* x. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a bounded Borel measurable function. Then $\mathbb{E}\left[f(X^x(t))\right]$ is given by

$$\mathbb{E}\left[f\left(X^x(t)\right)\right] = \int f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) \frac{\exp\left(-|y|^2\right)}{\left(\sqrt{\pi}\right)^d} dy.$$

Moreover, the Ornstein-Uhlenbeck process is a strong Markov process. This is also true for Brownian motion and for the oscillator process. Its integral kernel $p_0(t,x,y)$ is given by

$$p_0(t,x,y) = \frac{1}{\left(1-e^{-2t}\right)^{d/2}} \exp\left(-\frac{e^{-2t}|x|^2 + e^{-2t}|y|^2 - 2e^{-t}\langle x,y\rangle}{1-e^{-2t}}\right).$$

The semigroup in $C_b\left(\mathbb{R}^d\right)$ is given by

$$\begin{aligned}\left[\exp(tL)f\right](x) &= \int p_0(t,x,y) f(y) \exp\left(-|y|^2\right) \frac{dy}{\left(\sqrt{\pi}\right)^d} \\ &= \frac{1}{\left(\sqrt{\pi}\right)^d} \int f\left(\exp(-t)x + \sqrt{1-\exp(-2t)}y\right) \exp\left(-|y|^2\right) dy.\end{aligned}$$

Its invariant measure is determined by taking the limit:

$$\lim_{t\to\infty} \left[\exp(tL)f\right](x) = \frac{1}{\pi^{d/2}} \int f(y) e^{-|y|^2} dy.$$

For more details the reader is referred to e.g. [Simon (1979)]. The joint distributions of the processes (see Theorem 1.19.(d) of [Demuth and van Casteren (2000)])

$$\{X(t) : t \geqslant 0\} \text{ and } \left\{e^{-t}B\left(\left(e^{2t}-1\right)/2\right) : t \geqslant 0\right\}$$

coincide. The process $\{X(t) : t \geqslant 0\}$ also possesses the same law (i.e. joint distribution) as the process $\left\{\int_0^t \exp\left(-(t-s)\right) dB(s) : t \geqslant 0\right\}$.

The semigroup generated by L is not a bounded analytic one. This can be seen by rewriting the expression for $\lambda R(\lambda) = \lambda\left(\lambda I - L\right)^{-1}$, $\Re\lambda > 0$. For convenience we write:

$$\begin{aligned}q\left(s,x,y\right) &= \frac{1}{\left(1-s^2\right)^{d/2}} \exp\left(-\frac{s^2|x|^2 + |y|^2 - 2s\langle x,y\rangle}{1-s^2}\right) \\ &= \frac{1}{\left(1-s^2\right)^{d/2}} \exp\left(-\frac{|y-sx|^2}{1-s^2}\right). \qquad (8.104)\end{aligned}$$

Then we have $\lim_{t\to\infty} q\left(e^{-t},x,y\right) = \exp\left(-\left|y\right|^2\right)$, and also

$$\begin{aligned}
p_0\left(t,x,y\right)e^{-\left|y\right|^2} &= q\left(e^{-t},x,y\right), \quad t>0,\\
\frac{\partial}{\partial y_j}q\left(s,x,y\right) &= -\frac{2\left(y_j-sx_j\right)}{1-s^2}q\left(s,x,y\right) \quad \text{and}\\
\frac{\partial^2}{\left(\partial y_j\right)^2}q\left(s,x,y\right) &= -\frac{2}{1-s^2}q\left(s,x,y\right)+\frac{4\left(y_j-sx_j\right)^2}{\left(1-s^2\right)^2}q\left(s,x,y\right). \qquad (8.105)
\end{aligned}$$

From the equalities in (8.105) we get:

$$\begin{aligned}
&\frac{1}{2}\Delta_y q\left(s,x,y\right)+dq\left(s,x,y\right)+\left\langle y,\nabla_y q\left(s,x,y\right)\right\rangle\\
&= -d\frac{s^2}{1-s^2}q(s,x,y)+\frac{2s}{(1-s^2)^2}\left\{s\left|x\right|^2+s\left|y\right|^2-\left(1+s^2\right)\left\langle y,x\right\rangle\right\}q(s,x,y)\\
&= -s\frac{\partial}{\partial s}q\left(s,x,y\right) = \frac{\partial}{\partial t}q\left(e^{-t},x,y\right)\Big|_{e^{-t}=s}. \qquad (8.106)
\end{aligned}$$

Let $f\in D(L)$, and let μ_0 be the Borel measure on $\mathbb{R}^d$ which has density $\pi^{-d/2}\exp\left(-\left|y\right|^2\right)$ with respect to the Lebesgue measure. Then integration by parts yields:

$$\begin{aligned}
&\lambda\int_0^\infty e^{-\lambda t}e^{tL}f(x)dt\\
&= \left(1-e^{-\lambda t}\right)e^{tL}f(x)\Big|_{t=0}^{t=\infty}-\int_0^\infty\left(1-e^{-\lambda t}\right)e^{tL}Lf(x)dt\\
&= \left\langle f,\mu_0\right\rangle-\int_0^\infty\left(1-e^{-\lambda t}\right)\int_{\mathbb{R}^d}q\left(e^{-t},x,y\right)Lf(y)\frac{dy}{\pi^{d/2}}\,dt\\
&= \left\langle f,\mu_0\right\rangle-\int_0^\infty\left(1-e^{-\lambda t}\right)\int_{\mathbb{R}^d}q\left(e^{-t},x,y\right)\left(\frac{1}{2}\Delta f(y)-\left\langle y,\nabla f(y)\right\rangle\right)\frac{dy}{\pi^{d/2}}\,dt
\end{aligned}$$

(apply again integration by parts)

$$\begin{aligned}
&= \left\langle f,\mu_0\right\rangle-\int_0^\infty\left(1-e^{-\lambda t}\right)\\
&\quad\int_{\mathbb{R}^d}\left(\frac{1}{2}\Delta_y q\left(e^{-t},x,y\right)+dq\left(e^{-t},x,y\right)+\left\langle y,\nabla_y q\left(e^{-t},x,y\right)\right\rangle\right)f(y)\frac{dy}{\pi^{d/2}}\,dt\\
&= \left\langle f,\mu_0\right\rangle-\int_0^\infty(1-e^{-\lambda t})\\
&\quad\int_{\mathbb{R}^d}\left(\frac{-d}{e^{2t}-1}+\frac{2e^{2t}\{|x|^2+|y|^2-(e^t+e^{-t})\left\langle y,x\right\rangle\}}{(e^{2t}-1)^2}\right)q(e^{-t},x,y)f(y)\frac{dy}{\pi^{d/2}}\,dt
\end{aligned}$$

$$= \langle f, \mu_0\rangle - \int_{\mathbb{R}^d}\int_0^\infty \left(1-e^{-\lambda t}\right)$$
$$\left(\frac{-d}{e^{2t}-1} + \frac{2e^{2t}\left\{|x|^2+|y|^2-\left(e^t+e^{-t}\right)\langle y,x\rangle\right\}}{\left(e^{2t}-1\right)^2}\right) q\left(e^{-t},x,y\right)\, dt f(y)\frac{dy}{\pi^{d/2}}$$

(make the substitution $y = e^{-t}x + \sqrt{1-e^{-2t}}y'$)

$$= \langle f, \mu_0\rangle - \int_{\mathbb{R}^d}\int_0^\infty \frac{1-e^{-\lambda t}}{e^{2t}-1}\left(-d + 2\left\{|y'|^2 - \sqrt{e^{2t}-1}\,\langle y',x\rangle\right\}\right)$$
$$\times \exp\left(-|y'|^2\right) f\left(e^{-t}x + \sqrt{1-e^{-2t}}y'\right)\, dt\frac{dy'}{\pi^{d/2}}. \tag{8.107}$$

By the same token we get

$$\lambda\int_0^\infty e^{-\lambda t}e^{tL}f(x)dt$$
$$= f(x) + \lim_{\eta\downarrow 0}\int_{\mathbb{R}^d}\int_\eta^\infty \frac{e^{-\lambda t}}{e^{2t}-1}\left(-d + 2\left\{|y'|^2 - \sqrt{e^{2t}-1}\,\langle y',x\rangle\right\}\right)$$
$$\times \exp\left(-|y'|^2\right) f\left(e^{-t}x + \sqrt{1-e^{-2t}}y'\right)\, dt\frac{dy'}{\pi^{d/2}}. \tag{8.108}$$

From (8.107) we infer

$$\left|\lambda\int_0^\infty e^{-\lambda t}e^{tL}f(x)dt - \langle f,\mu_0\rangle\right|$$
$$\leqslant \int_{\mathbb{R}^d}\left|\int_0^\infty \left(1-e^{-\lambda t}\right)\left(\frac{-d}{e^{2t}-1} + \frac{2e^{2t}\left\{|x|^2+|y|^2-\left(e^t+e^{-t}\right)\langle y,x\rangle\right\}}{\left(e^{2t}-1\right)^2}\right)\right.$$
$$\left. q\left(e^{-t},x,y\right)\, dt\right| \frac{dy}{\pi^{d/2}}\,\|f\|_\infty$$
$$\leqslant \int_{\mathbb{R}^d}\int_0^\infty \left|1-e^{-\lambda t}\right|\left|\frac{-d}{e^{2t}-1} + \frac{2e^{2t}\left\{|x|^2+|y|^2-\left(e^t+e^{-t}\right)\langle y,x\rangle\right\}}{\left(e^{2t}-1\right)^2}\right|$$
$$q\left(e^{-t},x,y\right)\, dt\frac{dy}{\pi^{d/2}}\,\|f\|_\infty$$

(make the substitution $y = e^{-t}x + \sqrt{1-e^{-2t}}y'$)

$$= \int_{\mathbb{R}^d}\int_0^\infty \left|1-e^{-\lambda t}\right|\left|\frac{-d}{e^{2t}-1} + \frac{2\left\{|y'|^2 - \sqrt{e^{2t}-1}\,\langle y',x\rangle\right\}}{e^{2t}-1}\right|$$

$$\exp\left(-\left|y'\right|^2\right) dt\frac{dy'}{\pi^{d/2}} \|f\|_\infty$$

$$= \int_{\mathbb{R}^d}\int_0^\infty \frac{\left|1-e^{-\lambda t}\right|}{e^{2t}-1} \left|-d + 2\left\{\left|y'\right|^2 - \sqrt{e^{2t}-1}\left\langle y', x\right\rangle\right\}\right|$$

$$\exp\left(-\left|y'\right|^2\right) dt\frac{dy'}{\pi^{d/2}} \|f\|_\infty$$

$$= \int_{\mathbb{R}^d}\int_0^\infty \frac{\left|1-e^{-\lambda t}\right|}{e^{2t}-1} \left|-d + \left\{|y|^2 - \sqrt{2\left(e^{2t}-1\right)}\left\langle y, x\right\rangle\right\}\right|$$

$$\exp\left(-\frac{1}{2}|y|^2\right) dt\frac{dy}{(2\pi)^{d/2}} \|f\|_\infty$$

$$\leqslant \int_{\mathbb{R}^d}\int_0^\infty \frac{\left|1-e^{-\lambda t}\right|}{e^{2t}-1} \left|-d + |y|^2\right| \exp\left(-\frac{1}{2}|y|^2\right) dt\frac{dy}{(2\pi)^{d/2}} \|f\|_\infty$$

$$+ \int_{\mathbb{R}^d}\int_0^\infty \frac{\sqrt{2}\left|1-e^{-\lambda t}\right|}{\sqrt{e^{2t}-1}} \left|\left\langle y, x\right\rangle\right| \exp\left(-\frac{1}{2}|y|^2\right) dt\frac{dy}{(2\pi)^{d/2}} \|f\|_\infty$$

$$\leqslant 2d\int_0^\infty \frac{\left|1-e^{-\lambda t}\right|}{e^{2t}-1} dt \|f\|_\infty + \frac{2}{\sqrt{\pi}}\int_0^\infty \frac{\left|1-e^{-\lambda t}\right|}{\sqrt{e^{2t}-1}} dt\ |x| \|f\|_\infty . \tag{8.109}$$

We will also estimate the absolute value of the quantity:

$$\int q(0,x,y) f(y)\frac{dy}{\pi^{d/2}} - \int q\left(e^{-t},x,y\right) f(y)\frac{dy}{\pi^{d/2}}$$

$$= \int\int_t^\infty \left(-d + \frac{2e^{2s}}{e^{2s}-1}\left\{|x|^2 + |y|^2 - \left(e^s + e^{-s}\right)\left\langle y, x\right\rangle\right\}\right)$$

$$\times \frac{1}{e^{2s}-1} q\left(e^{-s},x,y\right) ds\, f(y)\frac{dy}{\pi^{d/2}}$$

$$= \int\int_t^\infty \left\{-d + 2\left\{\left|y'\right|^2 - \sqrt{e^{2s}-1}\left\langle y', x\right\rangle\right\}\right\} \frac{1}{e^{2s}-1}\exp\left(-\left|y'\right|^2\right)$$

$$f\left(e^{-s}x + \sqrt{1-e^{-2s}}y'\right) ds\frac{dy'}{\pi^{d/2}}$$

$$= \int\int_t^\infty \left\{-d + \left\{|y|^2 - \sqrt{\frac{e^{2s}-1}{2}}\left\langle y, x\right\rangle\right\}\right\} \frac{1}{e^{2s}-1}\exp\left(-\frac{1}{2}|y|^2\right)$$

$$f\left(e^{-s}x + \sqrt{\frac{1-e^{-2s}}{2}}y\right) ds\frac{dy}{(2\pi)^{d/2}}. \tag{8.110}$$

Here we used the equality in (8.106):

$$\left(-d + \frac{2e^{2s}}{e^{2s}-1}\left\{|x|^2 + |y|^2 - \left(e^s + e^{-s}\right)\left\langle y, x\right\rangle\right\}\right) \frac{1}{e^{2s}-1} q\left(e^{-s},x,y\right)$$

$$= \frac{\partial}{\partial s} q\left(e^{-s},x,y\right). \tag{8.111}$$

From (8.110) we obtain the following estimate in the same manner as we got the inequality in (8.109):

$$\left|\int q\left(0,x,y\right)f(y)\frac{dy}{\pi^{d/2}}-\int q\left(e^{-t},x,y\right)f(y)\frac{dy}{\pi^{d/2}}\right|$$
$$\leqslant 2d\int_t^\infty \frac{1}{e^{2s}-1}ds\left\|f\right\|_\infty+\frac{1}{\sqrt{\pi}}\int_t^\infty \frac{1}{\sqrt{e^{2s}-1}}ds\,\left|x\right|\left\|f\right\|_\infty. \tag{8.112}$$

Suppose $y \neq x$. In addition, the substitution $s = e^{-t}$ shows the equality:

$$\lambda\int_0^\infty e^{-\lambda t}p_0\left(t,x,y\right)e^{-\left|y\right|^2}dt=\lambda\int_0^1 s^{\lambda-1}q\left(s,x,y\right)ds$$
$$=d\int_0^1\frac{\left(1-s^\lambda\right)s}{1-s^2}q\left(s,x,y\right)\,ds$$
$$-\int_0^1\frac{2\left(1-s^\lambda\right)}{1-s^2}q\left(s,x,y\right)\,\frac{s\left|x\right|^2+s\left|y\right|^2-\left(1+s^2\right)\langle x,y\rangle}{1-s^2}ds. \tag{8.113}$$

From (8.113) we infer

$$\lambda\int_0^\infty\int f(y)e^{-\lambda t}p_0\left(t,x,y\right)e^{-\left|y\right|^2}\frac{dy}{\pi^{d/2}}dt$$
$$=d\int_0^1\frac{\left(1-s^\lambda\right)s}{1-s^2}\int f(y)q\left(s,x,y\right)\,\frac{dy}{\pi^{d/2}}\,ds$$
$$-\int_0^1\frac{2\left(1-s^\lambda\right)}{1-s^2}\int f(y)q\left(s,x,y\right)\,\frac{s\left|x\right|^2+s\left|y\right|^2-\left(1+s^2\right)\langle x,y\rangle}{1-s^2}\,\frac{dy}{\pi^{d/2}}\,ds$$

(make the substitution $y = sx + \sqrt{1-s^2}y'$)

$$=d\int_0^1\frac{\left(1-s^\lambda\right)s}{1-s^2}\int f\left(sx+\sqrt{1-s^2}y'\right)e^{-\left|y'\right|^2}\frac{dy'}{\pi^{d/2}}\,ds$$
$$-\int_0^1\frac{2\left(1-s^\lambda\right)}{1-s^2}\int f\left(sx+\sqrt{1-s^2}y'\right)\left(s\left|y'\right|^2-\sqrt{1-s^2}\,\langle x,y'\rangle\right)$$
$$\times\, e^{-\left|y'\right|^2}\frac{dy'}{\pi^{d/2}}\,ds$$
$$=\int_0^1\frac{s\left(1-s^\lambda\right)}{1-s^2}\int f\left(sx+\sqrt{1-s^2}y'\right)e^{-\left|y'\right|^2}\left(d-2\left|y'\right|^2\right)\frac{dy'}{\pi^{d/2}}\,ds$$
$$+\int_0^1\frac{2\left(1-s^\lambda\right)}{\sqrt{1-s^2}}\int f\left(sx+\sqrt{1-s^2}y'\right)\langle x,y'\rangle\,e^{-\left|y'\right|^2}\frac{dy'}{\pi^{d/2}}\,ds. \tag{8.114}$$

Let $C(t,s)$, $t \geqslant s$, t, $S \in \mathbb{R}$, be a family of $d \times d$ matrices with real entries, with the following properties:

(a) $C(t,t) = I$, $t \in \mathbb{R}$, (I stands for the identity matrix).
(b) The following identity holds: $C(t,s)C(s,\tau) = C(t,\tau)$ holds for all real numbers t, s, τ for which $t \geqslant s \geqslant \tau$.
(c) The matrix valued function $(t,s,x) \mapsto C(t,s)x$ is continuous as a function from the set $\left\{(t,s) \in \mathbb{R}^d \times \mathbb{R}^d : t \geqslant s\right\} \times \mathbb{R}^d$ to $\mathbb{R}^d$.

Define the backward propagator Y_C on $C_b\left(\mathbb{R}^d\right)$ by $Y_C(s,t)f(x) = f\left(C(t,s)x\right)$, $x \in \mathbb{R}^d$, and $f \in C_b\left(\mathbb{R}^d\right)$. Then Y_C is a backward propagator on the space $C_b\left(\mathbb{R}^d\right)$, which is $\sigma\left(C_b\left(\mathbb{R}^d\right), M\left(\mathbb{R}^d\right)\right)$-continuous. Here the symbols $M\left(\mathbb{R}^d\right)$ stand for the vector space of all signed measures on $\mathbb{R}^d$.

Let $W(t)$ be standard m-dimensional Brownian motion on $(\Omega, \mathcal{F}_t, \mathbb{P})$ and let $\sigma(\rho)$ be a deterministic continuous function which takes its values in the space of $d \times m$-matrices. Put $Q(\rho) = \sigma(\rho)\sigma(\rho)^*$. Another interesting example is the following:

$$\begin{aligned} &Y_{C,Q}(s,t)\, f(x) \\ &= \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f\left(C(t,s)x + \left(\int_s^t C(t,\rho)Q(\rho)C(t,\rho)^* d\rho\right)^{1/2} y\right) dy \\ &= \mathbb{E}\left[f\left(C(t,s)x + \int_s^t C(t,\rho)\sigma(\rho)dW(\rho)\right)\right], \end{aligned} \tag{8.115}$$

where A is an arbitrary $d \times d$ matrix, and where $Q(\rho) = \sigma(\rho)\sigma(\rho)^*$ is a positive-definite $d \times d$ matrix. Then the propagators $Y_{C,Q}$ and $Y_{C,S}$ are backward propagators on $C_b\left(\mathbb{R}^d\right)$. We will prove this.

Next suppose that the forward propagator C on $\mathbb{R}^d$ consists of contractive operators, i.e. $C(t,s)C(t,s)^* \leqslant I$ (this inequality is to be taken in matrix sense). Choose a family $S(t,s)$ of square $d \times d$-matrices such that $C(t,s)C(t,s)^* + S(t,s)\,S(t,s)^* = I$, and put

$$Y_{C,S}(s,t)\, f(x) = \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f\left(C(t,s)x + S(t,s)y\right) dy. \tag{8.116}$$

In fact the example in (8.116) is a special case of the example in (8.115) provided $Q(\rho)$ is given by the following limit:

$$Q(\rho) = \lim_{h \downarrow 0} \frac{I - C(\rho - h)\, C(\rho - h)^*}{h}. \tag{8.117}$$

If $Q(\rho)$ is as in (8.117), then

$$S(t,s)\, S(t,s)^* = I - C(t,s)\, C(t,s)^* = \int_s^t C(t,\rho)\, Q(\rho) C(t,\rho)^*\, d\rho.$$

The following auxiliary lemma will be useful. It is the finite-dimensional analog of Proposition 1.5 in Chapter 1. Condition (8.118) is satisfied if the three pairs (C_1, S_1), (C_2, S_2), and (C_3, S_3) satisfy: $C_1C_1^* + S_1S_1^* = C_2C_2^* + S_2S_2^* = C_3C_3^* + S_3S_3^* = I$. It also holds if $C_2 = C(t_2, t_1)$, and

$$S_jS_j^* = \int_{t_{j-1}}^{t_j} C(t_j, \rho)\,\sigma(\rho)\sigma(\rho)^* C(t_j, \rho)^*\,d\rho, \quad j = 1,\ 2, \quad \text{and}$$

$$S_3S_3^* = \int_{t_0}^{t_2} C(t_2, \rho)\,\sigma(\rho)\sigma(\rho)^* C(t_2, \rho)^*\,d\rho.$$

Lemma 8.3. *Let C_1, S_1, C_2, S_2, and C_3, S_3 be $d \times d$-matrices with the following properties:*

$$C_3 = C_2C_1, \quad \text{and} \quad C_2S_1S_1^*C_2^* + S_2S_2^* = S_3S_3^*. \tag{8.118}$$

Let $f \in C_b\left(\mathbb{R}^d\right)$, and put

$$Y_{1,2}f(x) = \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f(C_1x + S_1y)\,dy; \tag{8.119}$$

$$Y_{2,3}f(x) = \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f(C_2x + S_2y)\,dy; \tag{8.120}$$

$$Y_{1,3}f(x) = \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f(C_3x + S_3y)\,dy. \tag{8.121}$$

Then $Y_{1,2}Y_{2,3} = Y_{1,3}$.

Proof. Let the matrices C_j and S_j, $1 \leqslant j \leqslant 3$, be as in (8.118). Let $f \in C_b\left(\mathbb{R}^d\right)$. First we assume that the matrices S_1 and C_2 are invertible, and we put $A_3 = S_1^{-1}C_2^{-1}S_3$, and $A_2 = S_1^{-1}C_2^{-1}S_2$. Then, using the equalities in (8.118) we see $A_3A_3^* = I + A_2A_2^*$. We choose a $d \times d$-matrix A such that $A^*A = I + A_2^*A_2$, and we put $D = \left(A^{-1}\right)^* A_2^*A_3$. Then we have $A_3^*A_3 = I + D^*D$. Let $f \in C_b\left(\mathbb{R}^d\right)$. Let the vectors $(y_1, y_2) \in \mathbb{R}^d \times \mathbb{R}^d$ and $(y, z) \in \mathbb{R}^d \times \mathbb{R}^d$ be such that

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} A_3 & -A_2A^{-1} \\ 0 & A^{-1} \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}. \tag{8.122}$$

Since

$$A_2A_2^*\left(I + A_2A_2^*\right)^{-1} = A_2\left(I + A_2^*A_2\right)^{-1} A_2^*,$$

we obtain

$$\det\left(I + A_2A_2^*\right) = \det\left(I + A_2^*A_2\right).$$

Hence, the absolute value of the determinant of the matrix in the right-hand side of (8.122) can be rewritten as:

$$\left|\det \begin{pmatrix} A_3 & -A_2A^{-1} \\ 0 & A^{-1} \end{pmatrix}\right|^2 = \left|\det A_3 \left(\det A\right)^{-1}\right|^2$$

$$= \frac{\det(A_3 A_3^*)}{\det(A^* A)} = \frac{\det(I + A_2 A_2^*)}{\det(I + A_2^* A_2)} = 1. \tag{8.123}$$

From (8.122) and (8.123) it follows that the corresponding volume elements satisfy: $dy_1\, dy_2 = dy\, dz$. We also have

$$|y_1|^2 + |y_2|^2 = |y|^2 + |z - Dy|^2 . \tag{8.124}$$

Employing the substitution (8.122) together with the equalities $dy_1\, dy_2 = dy\, dz$ and (8.124) and applying Fubini's theorem we obtain:

$$\begin{aligned} Y_{1,2} Y_{2,3} f(x) &= \frac{1}{(2\pi)^d} \iint e^{-\frac{1}{2}\left(|y_1|^2 + |y_2|^2\right)} f\left(C_2 C_1 x + C_2 S_1 y_1 + S_2 y_2\right) dy_1 dy_2 \\ &= \frac{1}{(2\pi)^d} \iint e^{-\frac{1}{2}\left(|y|^2 + |z - Dy|^2\right)} f\left(C_3 x + S_3 y\right) dy\, dz \\ &= \frac{1}{(2\pi)^d} \int e^{-\frac{1}{2}|y|^2} f\left(C_3 x + S_3 y\right) dy = Y_{1,3} f(x) \end{aligned} \tag{8.125}$$

for all $f \in C_b\left(\mathbb{R}^d\right)$. If the matrices S_1 and C_2 are not invertible, then we replace the C_1 with $C_{1,\varepsilon} = e^{-\varepsilon} C_1$ and $S_{1,\varepsilon}$ satisfying $C_{1,\varepsilon} C_{1,\epsilon}^* + S_{1,\varepsilon} S_{1,\varepsilon}^* = I$, and $\lim_{\varepsilon \downarrow 0} S_{1,\varepsilon} = S_1$. We take $S_{2,\varepsilon} = e^{-\varepsilon} S_2$ instead of S_2. In addition, we choose the matrices $C_{2,\varepsilon}$, $\varepsilon > 0$, in such a way that $C_{2,\varepsilon} C_{2,\epsilon}^* + S_{2,\varepsilon} S_{2,\varepsilon}^* = I$, and $\lim_{\varepsilon \downarrow 0} C_{2,\varepsilon} = C_2$.

This completes the proof of Lemma 8.3. □

Proposition 8.4. *Put $X^{\tau,x}(t) = C(t,\tau)\, x + \int_\tau^t C(t,\rho)\, \sigma(\rho) dW(\rho)$. Then the process $X^{\tau,x}(t)$ is Gaussian. Its expectation is given by $\mathbb{E}\left[X^{\tau,x}(t)\right] = C(t,\tau)\, x$, and its covariance matrix has entries*

$$\mathbb{P}\text{-cov}\left(X_j^{\tau,x}(s), X_k^{\tau,x}(t)\right) = \left(\int_s^t C(t,\rho)\, Q(\rho) C(t,\rho)^*\, d\rho\right)_{j,k} . \tag{8.126}$$

Let $\left\{(\Omega, \mathcal{F}, \mathbb{P}_{\tau,x}), (X(t), t \geqslant 0), \left(\mathbb{R}^d, \mathcal{B}^d\right)\right\}$ be the corresponding time-inhomogeneous Markov process. Then this process is generated by the family operators $L(t)$, $t \geqslant 0$, where

$$L(t) f(x) = \frac{1}{2} \sum_{j,k=1}^{d} Q_{j,k}(t) D_j D_k f(x) + \langle \nabla f(x), A(t) x \rangle . \tag{8.127}$$

Here the matrix-valued function $A(t)$ is given by $A(t) = \lim_{h \downarrow 0} \dfrac{C(t+h,t) - I}{h}$. The semigroup $e^{sL(t)}$, $s \geqslant 0$, is given by

$e^{sL(t)} f(x)$

$$
\begin{aligned}
&= \mathbb{E}\left[f\left(e^{sA(t)}x + \int_0^s e^{(s-\rho)A(t)}\sigma(t)dW(\rho) \right) \right] \\
&= \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f\left(e^{sA(t)}x + \left(\int_0^s e^{\rho A(t)} Q(t) e^{\rho A(t)^*} d\rho \right)^{1/2} y \right) dy \\
&= \int p\left(s,x,y;t\right) f(y) dy
\end{aligned} \tag{8.128}
$$

where, with $Q_{A(t)}(s) = \int_0^s e^{\rho A(t)} Q(t) e^{\rho A(t)^*} d\rho$, *the integral kernel* $p(s,x,y;t)$ *is given by*

$$
\begin{aligned}
&p\left(s,x,y;t\right) \\
&= \frac{1}{(2\pi)^{d/2}\sqrt{\det Q_{A(t)}(s)}} e^{\left(-\frac{1}{2}\left\langle \left(Q_{A(t)}(s)\right)^{-1}\left(y-e^{sA(t)}x\right), y-e^{sA(t)}x \right\rangle\right)}.
\end{aligned}
$$

If all eigenvalues of the matrix $A(t)$ *have strictly negative real part, then the measure*

$$
B \mapsto \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} \mathbf{1}_B \left(\int_0^\infty e^{\rho A(t)} Q(t) e^{\rho A(t)^*} d\rho y \right) dy
$$

defines an invariant measure for the semigroup $e^{sL(t)}$, $s \geqslant 0$.

From Remark 8.7 below it follows that our theory is not directly applicable to the Ornstein-Uhlenbeck process as exhibited in Proposition 8.4. Therefore we will modify the example in the next proposition.

Proposition 8.5. *Let the* $\mathbb{R}^d$*-valued process* $X(t)$ *be a solution to the following stochastic differential equation:*

$$
X(t) = C(t,\tau)\, X(\tau) + \int_\tau^t C(t,\rho)\, F(\rho, X(\rho))\, d\rho + \int_\tau^t C(t,\rho)\, \sigma(\rho, X(\rho))\, dW(\rho). \tag{8.129}
$$

Under appropriate conditions on the functions F *and* σ *the equation in (8.129) has a unique weak solution. More precisely, it is assumed that* $x \mapsto \sigma(t,x)$ *is Lipschitz continuous with a constant which depends continuously on* t, *and that for some strictly positive continuous functions* $k_1(t)$, $k_2(t)$ *and* $k_3(t)$, *and strictly positive finite constants* $\varepsilon > 0$, $\alpha > 0$, *the following inequality holds for all* y, $z \in \mathbb{R}^d$*:*

$$
\left\langle F(t, y+z), \frac{y}{|y|} \right\rangle \leqslant -k_1(t)\,|y|^{1+\varepsilon} + k_2(t)\,|z|^\alpha + k_3(t). \tag{8.130}
$$

It is also assumed that the functions $y \mapsto F(t,y)$ and $y \mapsto \sigma(t,y)$ are locally Lipschitz, i.e., for every compact subset K of $\mathbb{R}^d$ there exists a continuous function $t \mapsto C_K(t)$ such that for all y_1 and $y_2 \in K$ the inequalities

$$\begin{aligned} &|F(t,y_2) - F(t,y_1)| \leqslant C_K(t)\,|y_2 - y_1|\,, \quad \textit{and} \\ &|\sigma(t,y_2) - \sigma(t,y_1)| \leqslant C_K(t)\,|y_2 - y_1|\,, \end{aligned} \tag{8.131}$$

hold. The corresponding Markov process

$$\left\{\left(\Omega, \mathcal{F}_\tau^t, \mathbb{P}_{\tau,x}\right), \left(X(t), t \geqslant 0\right), \left(\mathbb{R}^d, \mathcal{B}^d\right)\right\}$$

is generated by the time-dependent linear differential operators $L(t)$, given by

$$L(t)f(x) = \langle \nabla f(x), A(t)x + F(t,x)\rangle + \frac{1}{2}\sum_{j,k=1}^{d} D_j D_k f(x) a_{j,k}(t,x), \tag{8.132}$$

where

$$\begin{aligned} A(t) &= \lim_{h\downarrow 0}\frac{C(t+h,t) - C(t,t)}{h}, \quad \textit{and} \\ a_{j,k}(t,x) &= \sum_{\ell=1}^{d} \sigma_{j,\ell}(t,x)\,\sigma_{k,\ell}(t,x)\,. \end{aligned}$$

It is assumed that the operator $A(t)$ satisfies $\langle A(t)y, y\rangle \leqslant 0$, $y \in \mathbb{R}^d$. Moreover, let $X^{\tau,x}(t)$, $t \geqslant \tau$, be a solution to (8.129) with $X(\tau) = x$. Then

$$\mathbb{E}_{\tau,x}\left[\prod_{j=1}^{n} f_j\left(X\left(t_j\right)\right)\right] = \mathbb{E}\left[\prod_{j=1}^{n} f_j\left(X^{\tau,x}\left(t_j\right)\right)\right],$$

where $\mathbb{E}$ is the expectation with respect to the distribution of Brownian motion. In addition,

$$\frac{\partial}{\partial t}\mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right] = \mathbb{E}_{\tau,x}\left[L(t)f\left(X(t)\right)\right].$$

Proof. Fix a C^1-function $\varphi : \mathbb{R}^d \to [0,\infty)$ such that $\int_{\mathbb{R}^d}\varphi(y)dy = 1$, and $\operatorname{supp}(\varphi) \subset \{|y| \leqslant 1\}$. Moreover, assume that $\varphi(y)$ is symmetric in the sense that $\varphi(y) = \varphi(-y)$, $y \in \mathbb{R}^d$. This property implies e.g. $\int_{\mathbb{R}^d} y\varphi(y)dy = 0$. In addition, let ε_n, $n \in \mathbb{N}$, be a sequence of positive real numbers such that $0 < \varepsilon_{n+1} \leqslant \varepsilon_n \leqslant 1$, $n \in \mathbb{N}$, and such that $\lim_{n\to\infty}\varepsilon_n = 0$. Let the process $t \mapsto Y(t)$ be such that $\mathbb{E}\left[\sup_{\tau\leqslant t\leqslant T}|Y(t)|\right] < \infty$, and define the processes $s \mapsto F_n(s, Y(s))$, $n \in \mathbb{N}$, by

$$F_n(s, Y(s)) = \int_{\mathbb{R}^d} F(s, Y(s) - \varepsilon_n y)\,\varphi(y)dy$$

$$= \int_{\mathbb{R}^d} F(s,y)\,\varphi\left(\frac{Y(s)-y}{\varepsilon_n}\right)\frac{dy}{\varepsilon_n^k}.$$

Then the functions F_n have properties similar to F, and for each fixed n, the functional $Y(\cdot) \mapsto F_n(t, Y(t))$ is globally Lipschitz continuous. Instead of looking at the equation in (8.129) we consider the sequence of equations ($n \in \mathbb{N}$):

$$\begin{aligned} X_n(t) &= C(t,\tau)\,X_n(\tau) + \int_\tau^t C(t,\rho)\,F_n(\rho, X_n(\rho))\,d\rho \\ &\quad + \int_\tau^t C(t,\rho)\,\sigma(\rho, X_n(\rho))\,dW(\rho). \end{aligned} \tag{8.133}$$

Assuming that the equation in (8.133) has a solution $X_n(t)$, then we write

$$Z_n(t) = \int_\tau^t C(t,\rho)\,\sigma(\rho, X_n(\rho))\,dW(\rho) \quad \text{and} \quad Y_n(t) = X_n(t) - Z_n(t). \tag{8.134}$$

In terms of $Y_n(t)$ and $Z_n(t)$ the equation in (8.133) reads as follows (notice that $Z_n(\tau) = 0$):

$$Y_n(t) = C(t,\tau)\,Y_n(\tau) + \int_\tau^t C(t,\rho)\,F_n(\rho, Y_n(\rho) + Z_n(\rho))\,d\rho. \tag{8.135}$$

Moreover, from (8.130) it follows that

$$\begin{aligned} &\left\langle F_n(t, Y_n(t) + Z_n(t)), \frac{Y_n(t)}{|Y_n(t)|}\right\rangle \\ &\leqslant -k_1(t)\,|Y_n(t)|^{1+\varepsilon} + k_2(t)\,|Z_n(t)|^\alpha + k_3(t) + \varepsilon_n. \end{aligned} \tag{8.136}$$

From our hypotheses it follows that

$$\begin{aligned} \frac{d}{dt}|Y_n(t)| &= \left\langle \frac{d}{dt}Y_n(t), \frac{Y_n(t)}{|Y_n(t)|}\right\rangle \\ &= \left\langle A(t)Y_n(t), \frac{Y_n(t)}{|Y_n(t)|}\right\rangle + \left\langle F_n(t, Y_n(t) + Z_n(t)), \frac{Y_n(t)}{|Y_n(t)|}\right\rangle \\ &\leqslant -k_1(t)\,|Y_n(t)|^{1+\varepsilon} + k_2(t)\,|Z_n(t)|^\alpha + k_3(t) + \varepsilon_n. \end{aligned} \tag{8.137}$$

From (8.134) we also see:

$$Z_n(t) = \int_\tau^t C(t,\rho)\,\sigma(\rho, Y_n(\rho) + Z_n(\rho))\,dW(\rho). \tag{8.138}$$

Applying Hölder's inequality to (8.137) shows

$$\frac{d}{dt}\mathbb{E}_{\tau,x}\left[|Y_n(t)|\right] \leqslant -k_1(t)\left(\mathbb{E}_{\tau,x}\left[|Y_n(t)|\right]\right)^{1+\varepsilon} + k_2(t)\mathbb{E}_{\tau,x}\left[|Z_n(t)|^\alpha\right] + k_3(t) + \varepsilon_n. \tag{8.139}$$

Next put $y_{1,n}(t) = \mathbb{E}_{\tau,x}\left[|Y_n(t)|\right]$, and let $y_{2,n}(t)$ be any continuously differentiable positive function with the following properties: $y_{2,n}(\tau) \geqslant y_{1,n}(\tau) = |x|$, and

$$\dot{y}_{2,n}(t) \geqslant -k_1(t)y_{2,n}(t)^{1+\varepsilon} + k_2(t)\mathbb{E}_{\tau,x}\left[|Z_n(t)|^{\alpha}\right] + k_3(t) + \varepsilon_n. \tag{8.140}$$

Then from (8.137), (8.139), and Lemma 8.4 below we obtain $y_{2,n}(t) \geqslant y_{1,n}(t)$, $t \geqslant \tau$.

A martingale solution to equation (8.129) can be found as follows. First find a (weak) solution $X_0(t)$, $t \geqslant \tau \geqslant 0$, to the equation

$$X_0(t) = C\left(t,\tau\right)X_0(\tau) + \int_\tau^t C\left(t,\rho\right)\sigma\left(\rho, X_0(\rho)\right)dW(\rho). \tag{8.141}$$

Then choose $F_0\left(t,y\right)$ in such a way that $F\left(t,y\right) = \sigma\left(t,y\right)F_0\left(t,y\right)$. After that we define the finite-dimensional distributions of the process $X_F(t)$ as follows. First we introduce the process $\zeta\left(t,\tau\right)$, $t \geqslant \tau$:

$$\zeta\left(t,\tau\right) = \int_\tau^t F_0\left(\rho, X_0(\rho)\right)dW(\rho) - \frac{1}{2}\int_\tau^t \left|F_0\left(\rho, X_0(\rho)\right)\right|^2 d\rho. \tag{8.142}$$

Then the process $t \mapsto e^{\zeta(t,\tau)}$, $t \geqslant \tau$, is a local martingale with respect to the filtration $\mathcal{F}_t^{W,\tau} := \sigma\left(W(\rho) : \tau \leqslant \rho \leqslant t\right)$, $t \geqslant \tau$, generated by Brownian motion, and which is such that $X_0(\tau) = x$, $\mathbb{P}$-almost surely. This means that if $\mathbb{E}\left[e^{\zeta(t,\tau)} \,\middle|\, X_0(\tau) = x\right] = 1$, then the process $t \mapsto e^{\zeta(t,\tau)}$, $t \geqslant \tau$, is a martingale with respect to the measure $A \mapsto \mathbb{P}\left[A \,\middle|\, X(\tau) = x\right]$, $A \in \mathcal{F}_t^{W,\tau}$. The finite-dimensional distributions of the process $X_F(t)$, $t \geqslant \tau$, are given by the Girsanov formula:

$$\begin{aligned}&\mathbb{E}_{\tau,x}\left[f\left(X_F\left(t_1\right),\dots,X_F\left(t_n\right)\right)\right]\\ &= \mathbb{E}\left[e^{\zeta(t,\tau)}f\left(X_0\left(t_1\right),\dots,X_0\left(t_n\right)\right) \,\middle|\, X_0(\tau) = x\right].\end{aligned} \tag{8.143}$$

Here we assume that the function $f : \underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{n \text{ times}} \to \mathbb{R}$ is a bounded Borel function, and $\tau \leqslant t_1 < \cdots < t_n \leqslant t$. In order to prove that equality (8.143) determines the distribution of the process $X_F(t)$, $t \geqslant \tau$, we have to show that the martingale problem for the family of operators $L(t)$, $t \geqslant \tau$, in (8.127) is well-posed. Therefore, we apply Itô's formula to obtain:

$$\begin{aligned}&e^{\zeta(t,\tau)}f\left(X_0(t)\right) - e^{\zeta(\tau,\tau)}f\left(X_0(\tau)\right)\\ &= \int_\tau^t e^{\zeta(\rho,\tau)}f\left(X_0(\rho)\right)\left\langle F_0\left(\rho, X(\rho)\right), dW(\rho)\right\rangle\\ &\quad + \int_\tau^t e^{\zeta(\rho,\tau)}\left\langle \nabla f\left(X_0(\rho)\right), \sigma\left(\rho, X(\rho)\right)dW(\rho)\right\rangle\end{aligned}$$

$$
\begin{aligned}
&+\int_\tau^t e^{\zeta(\rho,\tau)}\left\langle \nabla f\left(X_0(\rho)\right), A(\rho)X_0(\rho)\right\rangle d\rho \\
&+\int_\tau^t e^{\zeta(\rho,\tau)}\left\langle \nabla f\left(X_0(\rho)\right), \sigma\left(\rho, X_0(\rho)\right) F_0\left(\rho, X_0(\rho)\right)\right\rangle d\rho \\
&+\frac{1}{2}\sum_{j,k=1}^{d}\int_\tau^t e^{\zeta(\rho,\tau)}\left(\sigma\left(\rho, X_0(\rho)\right)\sigma\left(\rho, X_0(\rho)\right)^*\right)_{j,k} D_jD_k f\left(X_0(\rho)\right) d\rho \\
=&\int_\tau^t e^{\zeta(\rho,\tau)} f\left(X_0(\rho)\right)\left\langle F_0\left(\rho, X(\rho)\right), dW(\rho)\right\rangle \\
&+\int_\tau^t e^{\zeta(\rho,\tau)}\left\langle \nabla f\left(X_0(\rho)\right), \sigma\left(\rho, X(\rho)\right) dW(\rho)\right\rangle \\
&+\int_\tau^t e^{\zeta(\rho,\tau)} L(\rho) f\left(X_0(\rho)\right) d\rho. \qquad (8.144)
\end{aligned}
$$

It follows that that the process

$$
t \mapsto e^{\zeta(t,\tau)} f\left(X_0(t)\right) - f\left(X_0(\tau)\right) - \int_\tau^t e^{\zeta(\rho,\tau)} L(\rho) f\left(X_0(\rho)\right) d\rho
$$

is a martingale with respect to the measure $A \mapsto \mathbb{P}\left[A \mid X(\tau) = x\right]$, $A \in \mathcal{F}_t^{W,\tau}$, provided that $\mathbb{E}\left[e^{\zeta(t,\tau)} \mid X(\tau) = x\right] = 1$. Hence under the latter condition it follows that the process $t \mapsto X_F(t)$ is a $\mathbb{P}_{\tau,x}$-martingale. Essentially speaking this proves that the martingale problem for the operators $L(t)$, $t \geqslant \tau$, possesses solutions. In order to establish the Markov property we need the uniqueness of solutions. The uniqueness of solutions can be achieved as follows. Let $t \mapsto X^1(t)$ and $X^2(t)$, $t \geqslant \tau$, be solutions to equation (8.129). Put $Z^j(t) = \int_\tau^t \sigma\left(\rho, X^j(\rho)\right) dW(\rho)$, and $Y^j(t) = X^j(t) - Z^j(t)$. Then $Z^j(t) = \int_\tau^t \sigma\left(\rho, Y^j(\rho) + Z^j(\rho)\right) dW(\rho)$, and

$$
Y^j(t) = C\left(t,\tau\right) Y^j(\tau) + \int_\tau^t F\left(\rho, Y^j(\rho) + Z^j(\rho)\right) d\rho. \qquad (8.145)
$$

Let K be a compact subset of $[\tau,\infty) \times \mathbb{R}^d$, and define the stopping times τ_K^j, $j = 1, 2$, and τ_K by

$$
\tau_K^j = \inf\left\{t > \tau : \left(t, X^j(t)\right) \in \left([\tau,\infty) \times \mathbb{R}^d\right) \backslash K\right\} \quad \text{and} \quad \tau_K = \min\left(\tau_K^1, \tau_K^2\right).
$$

Then on the event $\{\tau_K > t\}$ by the local Lipschitz property of the function F and σ we have (see (8.131))

$$
\begin{aligned}
&\frac{d}{dt}\left|Y^2(t) - Y^1(t)\right| = \left\langle \frac{d}{dt}\left(Y^2(t) - Y^1(t)\right), \frac{Y^2(t) - Y^1(t)}{\left|Y^2(t) - Y^1(t)\right|}\right\rangle \\
&= \left\langle A(t)\left(Y^2(t) - Y^1(t)\right), \frac{Y^2(t) - Y^1(t)}{\left|Y^2(t) - Y^1(t)\right|}\right\rangle
\end{aligned}
$$

$$
\begin{aligned}
&+\left\langle F\left(t, Y^2(t)+Z^2(t)\right)-F\left(t, Y^1(t)+Z^1(t)\right), \frac{Y^2(t)-Y^1(t)}{\left|Y^2(t)-Y^1(t)\right|}\right\rangle \\
&\leqslant C_K(t)\left(\left|Y^2(t)-Y^1(t)\right|+\left|Z^2(t)-Z^1(t)\right|\right) \\
&= C_K(t)\left(\left|Y^2(t)-Y^1(t)\right|+\left|\int_\tau^t\left(\sigma\left(\rho, X^2(\rho)\right)-\sigma\left(\rho, X^1(\rho)\right)\right) W(\rho)\right|\right).
\end{aligned}
\tag{8.146}
$$

From the Gronwall inequality (8.189) in Lemma 8.6 and (8.146) we infer:

$$
\begin{aligned}
\left|Y^2(t)-Y^1(t)\right| \leqslant &\left|Y^2(\tau)-Y^1(\tau)\right| e^{\int_\tau^t C_K(\rho) d\rho} \\
&+\int_\tau^t e^{\int_\rho^t C_K\left(\rho'\right) d\rho'} C_K(\rho)\left|Z_2(\rho)-Z_1(\rho)\right| d\rho .
\end{aligned}
\tag{8.147}
$$

Inequality (8.147) on the event $\{\tau_K > t\}$ entails:

$$
\begin{aligned}
&\left|Y^2(t)-Y^1(t)\right| \mathbf{1}_{\{\tau_K>t\}} \\
&\leqslant\left|Y^2(\tau)-Y^1(\tau)\right| \mathbf{1}_{\{\tau_K>t\}} e^{\int_\tau^t C_K(\rho) d\rho} \\
&\quad+\int_\tau^t e^{\int_\rho^t C_K\left(\rho'\right) d\rho'} C_K(\rho)\left|Z_2(\rho)-Z_1(\rho)\right| \mathbf{1}_{\{\tau_K>t\}} d\rho \\
&=\left|Y^2(\tau)-Y^1(\tau)\right| \mathbf{1}_{\{\tau_K>t\}} e^{\int_\tau^t C_K(\rho) d\rho} \\
&\quad+\int_\tau^t e^{\int_\rho^t C_K\left(\rho'\right) d\rho'} C_K(\rho)\left|\int_\tau^\rho\left(\sigma\left(\rho', X^2\left(\rho'\right)\right)-\sigma\left(\rho', X^1\left(\rho'\right)\right)\right) dW\left(\rho'\right)\right| \\
&\quad\quad\times \mathbf{1}_{\{\tau_K>t\}} d\rho \\
&=\left|Y^2(\tau)-Y^1(\tau)\right| \mathbf{1}_{\{\tau_K>t\}} e^{\int_\tau^t C_K(\rho) d\rho} \\
&\quad+\int_\tau^t e^{\int_\rho^t C_K\left(\rho'\right) d\rho'} C_K(\rho) \\
&\quad\times\left|\int_\tau^{\rho\wedge\tau_K}\left(\sigma\left(\rho', X^2\left(\rho'\right)\right)-\sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right| \mathbf{1}_{\{\tau_K>t\}} d\rho \\
&\leqslant\left|Y^2(\tau)-Y^1(\tau)\right| \mathbf{1}_{\{\tau_K>\tau\}} e^{\int_\tau^t C_K(\rho) d\rho} \\
&\quad+\int_\tau^t e^{\int_\rho^t C_K\left(\rho'\right) d\rho'} C_K(\rho) \\
&\quad\times\left|\int_\tau^{\rho\wedge\tau_K}\left(\sigma\left(\rho', X^2\left(\rho'\right)\right)-\sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right| \mathbf{1}_{\{\tau_K>\tau\}} d\rho .
\end{aligned}
\tag{8.148}
$$

It follows that

$$
\sup_{\tau\leqslant s\leqslant t}\left|Y^2(s)-Y^1(s)\right| \mathbf{1}_{\{\tau_K>s\}}
$$

$$
\begin{aligned}
&\leqslant \left|Y^2(\tau)-Y^1(\tau)\right| \mathbf{1}_{\{\tau_K>\tau\}} e^{\int_\tau^t C_K(\rho)d\rho} \\
&\quad + \int_\tau^t e^{\int_\rho^t C_K(\rho')d\rho'} C_K(\rho) \\
&\quad \times \left| \int_\tau^{\rho\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right| \\
&\qquad \times \mathbf{1}_{\{\tau_K>\tau\}} d\rho,
\end{aligned}
\tag{8.149}
$$

and hence by the elementary inequalities $2ab \leqslant a^2+b^2$ and $(a+b)^2 \leqslant 2a^2+2b^2$, $a,\, b \in \mathbb{R}$,

$$
\begin{aligned}
&\sup_{\tau\leqslant s\leqslant t} \left|Y^2(s)-Y^1(s)\right|^2 \mathbf{1}_{\{\tau_K>s\}} \\
&\leqslant 2\left|Y^2(\tau)-Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}} e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\quad + 2\int_\tau^t\int_\tau^t e^{\int_{\rho_1}^t C_K(\rho')d\rho'} C_K(\rho_1) e^{\int_{\rho_2}^t C_K(\rho')d\rho'} C_K(\rho_2) \\
&\quad \times \left| \int_\tau^{\rho_1\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right| \\
&\quad \times \left| \int_\tau^{\rho_2\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right| \\
&\quad \times \mathbf{1}_{\{\tau_K>\tau\}} d\rho_1\, d\rho_2 \\
&\leqslant 2\left|Y^2(\tau)-Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}} e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\quad + \int_\tau^t\int_\tau^t e^{\int_{\rho_1}^t C_K(\rho')d\rho'} C_K(\rho_1) e^{\int_{\rho_2}^t C_K(\rho')d\rho'} C_K(\rho_2) \\
&\quad \times \left(\left| \int_\tau^{\rho_1\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right|^2 \right. \\
&\quad \left. + \left| \int_\tau^{\rho_2\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right|^2 \right) \\
&\quad \times \mathbf{1}_{\{\tau_K>\tau\}} d\rho_1\, d\rho_2 \\
&\leqslant 2\left|Y^2(\tau)-Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}} e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\quad + 2\int_\tau^t e^{\int_\rho^t C_K(\rho')d\rho'} C_K(\rho) \left(e^{\int_\tau^t C_K(\rho')d\rho'} - 1 \right) \\
&\quad \times \left| \int_\tau^{\rho\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}} d\rho,
\end{aligned}
\tag{8.150}
$$

The fact that the process

$$\rho \mapsto \int_{\tau}^{\rho\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)$$

is a martingale with respect to Brownian motion entails the equality

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{\tau}^{\rho\wedge\tau_K} \left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right) \mathbf{1}_{\{\tau_K>\rho'\}} dW\left(\rho'\right)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] \\
&= \mathbb{E}\left[\left|\int_{\tau}^{\rho\wedge\tau_K} \left|\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right|^2 \mathbf{1}_{\{\tau_K>\rho'\}} d\rho'\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right].
\end{aligned} \tag{8.151}$$

By taking expectations in (8.150) and using (8.151) we get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\tau\leqslant s\leqslant t}\left|Y^2(s) - Y^1(s)\right|^2 \mathbf{1}_{\{\tau_K>s\}}\right] \\
&\leqslant 2\mathbb{E}\left[\left|Y^2(\tau) - Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\quad + 2\int_\tau^t e^{\int_\rho^t C_K(\rho')d\rho'} C_K(\rho)\left(e^{\int_\tau^t C_K(\rho')d\rho'} - 1\right) \\
&\times \mathbb{E}\left[\left|\int_{\tau}^{\rho\wedge\tau_K} (\sigma(\rho', X^2(\rho')) - \sigma(\rho', X^1(\rho'))) \mathbf{1}_{\{\tau_K>\rho'\}} dW(\rho')\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] d\rho \\
&\leqslant 2\mathbb{E}\left[\left|Y^2(\tau) - Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\quad + 2\int_\tau^t e^{\int_\rho^t C_K(\rho')d\rho'} C_K(\rho)\left(e^{\int_\tau^t C_K(\rho')d\rho'} - 1\right) \\
&\qquad \times \mathbb{E}\left[\int_{\tau}^{\rho\wedge\tau_K} \left|\left(\sigma\left(\rho', X^2\left(\rho'\right)\right) - \sigma\left(\rho', X^1\left(\rho'\right)\right)\right)\right|^2 \mathbf{1}_{\{\tau_K>\rho'\}} d\rho'\right] d\rho
\end{aligned}$$

(employ the local Lipschitz property of the function $x \mapsto \sigma(\rho, x)$ with Lipschitz constant $\widetilde{C}_K(\rho)$)

$$\begin{aligned}
&\leqslant 2\mathbb{E}\left[\left|Y^2(\tau) - Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\quad + 2\int_\tau^t e^{\int_\rho^t C_K(\rho')d\rho'} C_K(\rho)\left(e^{\int_\tau^t C_K(\rho')d\rho'} - 1\right) \\
&\qquad \times \mathbb{E}\left[\int_{\tau}^{\rho\wedge\tau_K} \widetilde{C}_K^2\left(\rho'\right)\left|X^2\left(\rho'\right) - X^1\left(\rho'\right)\right|^2 \mathbf{1}_{\{\tau_K>\rho'\}} d\rho'\right] d\rho \\
&\leqslant 2\mathbb{E}\left[\left|Y^2(\tau) - Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho}
\end{aligned}$$

$$+ 2\left(e^{\int_\tau^t C_K(\rho)d\rho} - 1\right)^2 \mathbb{E}\left[\int_\tau^{t\wedge\tau_K} \widetilde{C}_K^2(\rho)\left|X^2(\rho) - X^1(\rho)\right|^2 \mathbf{1}_{\{\tau_K>\rho\}}d\rho\right]. \tag{8.152}$$

From the Burkholder-Davis-Gundy inequality for $p = 2$ we obtain;

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\tau\leqslant s\leqslant t}\left|Z^2(s) - Z^1(s)\right|^2 \mathbf{1}_{\{\tau_K>s\}}\right] \\
&= \mathbb{E}\left[\sup_{\tau\leqslant s\leqslant t}\left|\int_\tau^{s\wedge\tau_K}\left(\sigma\left(\rho, X^2(\rho)\right) - \sigma\left(\rho, X^1(\rho)\right)\right)\mathbf{1}_{\tau_K>\rho}dW(\rho)\right|^2 \mathbf{1}_{\{\tau_K>s\}}\right] \\
&\leqslant 4\mathbb{E}\left[\int_\tau^{t\wedge\tau_K}\left|\left(\sigma\left(\rho, X^2(\rho)\right) - \sigma\left(\rho, X^1(\rho)\right)\right)\mathbf{1}_{\tau_K>\rho}\right|^2 d\rho \mathbf{1}_{\{\tau_K>\tau\}}\right] \\
&\leqslant 4\mathbb{E}\left[\int_\tau^{t\wedge\tau_K}\widetilde{C}_K^2(\rho)\left|X^2(\rho) - X^1(\rho)\right|^2 \mathbf{1}_{\{\tau_K>\rho\}}d\rho \mathbf{1}_{\{\tau_K>\tau\}}\right].
\end{aligned} \tag{8.153}$$

Next we estimate the expectation of $\max_{\tau\leqslant s\leqslant t}\left|\overline{X}_K(s)\right|^2$ where

$$\overline{X}_K(t) = \left(Y^2(t) - Y^1(t) + Z^2(t) - Z^1(t)\right)\mathbf{1}_{\{\tau_K>t\}} = \overline{Y}_K(t) + \overline{Z}_K(t). \tag{8.154}$$

Here the notations $\overline{Y}_K(t)$ and $\overline{Z}_K(t)$ are self-explanatory. Put

$$u_K(s) = \mathbb{E}\left[\sup_{\tau\leqslant\rho\leqslant s}\left|X^2(\rho) - X^1(\rho)\right|^2 \mathbf{1}_{\{\tau_K>\rho\}}\right].$$

From (8.152) and (8.153) we then obtain:

$$\begin{aligned}
u_K(t) &\leqslant 4\mathbb{E}\left[\left|Y^2(\tau) - Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho} \\
&\qquad + 2\left(2\left(e^{\int_\tau^t C_K(\rho)d\rho} - 1\right)^2 + 1\right)\int_\tau^t \widetilde{C}_K^2(\rho)\, u_K(\rho)d\rho \\
&= \psi(t) + \chi(t)\int_\tau^t c_1(\rho)u_K(\rho)d\rho
\end{aligned} \tag{8.155}$$

where

$$\begin{aligned}
\psi(t) &= 4\mathbb{E}\left[\left|Y^2(\tau) - Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho}, \\
\chi(t) &= 2\left(2\left(e^{\int_\tau^t C_K(\rho)d\rho} - 1\right)^2 + 1\right), \quad \text{and} \quad c_1(t) = \widetilde{C}_K^2(t).
\end{aligned}$$

From the Gronwall inequality (8.188) in Lemma 8.6 below and (8.155) we then obtain:

$$u_K(t) \leqslant \psi(t) + \chi(t)\int_\tau^t e^{\int_\rho^t \chi(\rho')c_1(\rho')d\rho'} c_1(\rho)\psi(\rho)d\rho. \tag{8.156}$$

Since the functions $t \mapsto c_1(t)$ and $t \mapsto \psi(t)$ are increasing from (8.156) we infer

$$\begin{aligned} u_K(t) &\leqslant \psi(t) e^{\chi(t)\int_\tau^t c_1(\rho)d\rho} \\ &= 4\mathbb{E}\left[\left|Y^2(\tau)-Y^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho} e^{\chi(t)\int_\tau^t c_1(\rho)d\rho} \\ &= 4\mathbb{E}\left[\left|X^2(\tau)-X^1(\tau)\right|^2 \mathbf{1}_{\{\tau_K>\tau\}}\right] e^{2\int_\tau^t C_K(\rho)d\rho+\chi(t)\int_\tau^t c_1(\rho)d\rho}. \end{aligned} \tag{8.157}$$

If $X^2(\tau) = X^1(\tau)$ $\mathbb{P}$-almost surely, then (8.157) implies $X^2(t) = X^1(t)$ on the event $\{\tau_K > t\}$. Since K is an arbitrary compact subset of $[\tau, \infty) \times \mathbb{R}^d$ the latter proves that the stochastic differential equation (8.129) in Proposition 8.5 is uniquely solvable in the space $\mathcal{S}^2_{\mathrm{loc}} = \mathcal{S}^2_{\mathrm{loc}}\left(\mathbb{R}^d\right)$ consisting of continuous semi-martingales X withe property that $\mathbb{E}\left[\sup_{\tau\leqslant s\leqslant t}|X(s)|^2\right]$ is finite. Of course all this is true provided that solutions to the stochastic differential equation (8.129) belong to the space $\mathcal{S}^2_{\mathrm{loc}}$. But this follows from general arguments: see e.g. [Ikeda and Watanabe (1998)] or [Øksendal and Reikvam (1998)].

This completes the proof of Proposition 8.5. □

The following lemma and also Lemma 8.6 were employed in the proof of Proposition 8.5. Lemma 8.5 is included because it shows how, in case $g(t, y) = k(t)y(t)^{1+\varepsilon}$, solutions to equations of the form $\dot{y}(t) = -g(t, y(t)) + C(t)$ behave themselves for large t.

Lemma 8.4. *Fix $\tau \leqslant T$, and let $g : [\tau, T] \times \mathbb{R} \to \mathbb{R}$ be a continuous function, which is continuously differentiable in the second variable. In addition, let $C : [\tau, T] \to \mathbb{R}$ be a measurable function. Let the $\mathbb{R}$-valued continuous functions $t \mapsto y_2(t)$ and $t \mapsto y_1(t)$, $\tau \leqslant t \leqslant T$, satisfy the following differential inequalities:*

$$\dot{y}_1(t) \leqslant -g(t, y_1(t)) + C(t), \quad \tau \leqslant t \leqslant T, \quad \textit{and} \tag{8.158}$$

$$\dot{y}_2(t) \geqslant -g(t, y_2(t)) + C(t), \quad \tau \leqslant t \leqslant T. \tag{8.159}$$

If $y_2(\tau) \geqslant y_1(\tau)$, then $y_2(t) \geqslant y_1(t)$, $\tau \leqslant t \leqslant T$.

Proof. Put $\Phi(t) = y_2(t) - y_1(t)$, and

$$\Psi(t) = \exp\left(\int_\tau^t \int_0^1 D_2 g(\rho, (1-s)y_1(\rho) + s y_2(\rho))\, ds\, d\rho\right).$$

Then $\Psi(t) > 0$, and

$$\frac{d}{dt}\left(\Phi(t)\Psi(t)\right)$$

$$= \frac{\partial}{\partial t}\left(\Phi(t)\exp\left(\int_\tau^t \int_0^1 D_2 g\left(\rho,(1-s)y_1(\rho)+sy_2(\rho)\right)ds\,d\rho\right)\right)$$
$$= \left(\dot\Phi(t)+\Phi(t)\int_0^1 D_2 g\left(t,(1-s)y_1(t)+sy_2(t)\right)ds\right)\Psi(t)$$
$$= \left(\dot y_2(t)-\dot y_1(t)+(y_2(t)-y_1(t))\int_0^1 D_2 g\left(t,(1-s)y_1(t)+sy_2(t)\right)ds\right)\Psi(t)$$
$$= \left(\dot y_2(t)+g\left(t,y_2(t)\right)-\dot y_1(t)-g\left(t,y_1(t)\right)\right)\Psi(t)\geqslant 0, \tag{8.160}$$

where in inequality (8.160) we used (8.158) and (8.159). Hence we get

$$\Phi(t)\Psi(t)\geqslant\Phi(\tau)\Psi(\tau)=\Phi(\tau)=y_2(\tau)-y_1(\tau)\geqslant 0, \tag{8.161}$$

and thus $y_2(t)-y_1(t)=\Phi(t)\geqslant 0$, which completes the proof of Lemma 8.4. □

Lemma 8.5. *Let $y:[\tau,\infty)\to[0,\infty)$ be a solution to the following ordinary differential equation*

$$\dot y(t)=C(t)-k(t)y(t)^{1+\varepsilon},\quad t\geqslant\tau. \tag{8.162}$$

It is assumed that the functions $C(t)$ and $k(t)$ are strictly positive and continuous, that $\varepsilon>0$, and that the quotient $\gamma:=\gamma(t)=\dfrac{C(t)}{k(t)}$ does not depend on t. Then $\sup_{t\geqslant\tau}y(t)<\infty$. In addition, the following inequality holds for $t>\tau$:

$$\left|y(t)-\gamma^{1/(1+\varepsilon)}\right|\leqslant\left(\varepsilon\int_\tau^t k(\rho)d\rho\right)^{-1/\varepsilon},\quad t>\tau. \tag{8.163}$$

Moreover, with $\eta=\dfrac{1}{1+\varepsilon}$ the following assertions are true:

(1) If $\int_\tau^\infty k(\rho)d\rho=\infty$, then $\lim_{t\to\infty}y(t)=\gamma^{1/(1+\varepsilon)}$.
(2) If $y(\tau)=\gamma^\eta$, then $y(t)=\gamma^\eta$.
(3) If $\int_\tau^\infty k(\rho)\,d\rho<\infty$ and $y(\tau)>\gamma^\eta$, then the limit $\lim_{t\to\infty}y(t)>\gamma^\eta$.
(4) If $\int_\tau^\infty k(\rho)\,d\rho<\infty$ and $y(\tau)<\gamma^\eta$, then the limit $\lim_{t\to\infty}y(t)<\gamma^\eta$.

The importance of inequality (8.163) lies in the fact that in this inequality there is no reference to the initial value $y(\tau)$ of the solution $t\mapsto y(t)$. It seems that the inequality in (8.163) is somewhat nicer and stronger than the inequality in (2.15) of [Goldys and Maslowski (2001)].

Remark 8.6. As in the proof of Lemma 8.5 put $\eta=\dfrac{1}{1+\varepsilon}$. From (8.169) in the proof of Lemma 8.5 we see that $y(\tau)>\gamma^\eta$ implies $y(\tau)>y(t)>\gamma^\eta$, and that $y(t)$ decreases to its limit. We also see that $y(\tau) < \gamma^\eta$ entails $y(\tau) < y(t) < \gamma^\eta$, and that $y(t)$ increases to its limit. If the integral

$$\int_0^1 \int_\tau^t k(\rho)^\eta \left((1-s)C(\rho)^\eta + sk(\rho)^\eta y(\rho)\right)^\varepsilon d\rho\, ds \tag{8.164}$$

increases to ∞ with t, then $\lim_{t\to\infty} y(t) = \gamma^\eta$. Notice that the integrals in (8.164) tend to ∞ whenever the function $k(t)$ and $C(t)$ are constant. In order that the limit $\lim\limits_{t\to\infty} y(t) = \gamma^{1/(1+\varepsilon)}$ one needs the fact that the integral

$$\int_\tau^\infty k(\rho)^{\frac{1}{1+\varepsilon}} C(\rho)^{\frac{\varepsilon}{1+\varepsilon}} d\rho = \gamma^{\frac{\varepsilon}{1+\varepsilon}} \int_\tau^\infty k(\rho) d\rho$$

diverges. If it converges, then the limit $\lim_{t\to\infty} y(t)$ still exists, but it is not equal to γ^η. Moreover, the limit depends on the initial value. If $y(\tau) < \gamma^\eta$, then equality (8.169) implies $y(\tau) < y(t) < \gamma^\eta$ for all $t \geqslant \tau$. Moreover, $y(t)$ increases to γ^η. If $y(\tau) = \gamma^\eta$, then $y(t) = \gamma^\eta$, $t \geqslant \tau$.

Proof. [Proof of Lemma 8.5.] For brevity we write $\eta = \dfrac{1}{1+\varepsilon}$. We introduce the function $\varphi(t)$, $t \geqslant \tau$, defined by

$$\varphi(t) = \left(\gamma^\eta(t) - y(t)\right) e^{(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)^\eta ((1-s)C(\rho)^\eta + sk(\rho)^\eta y(\rho))^\varepsilon d\rho\, ds}. \tag{8.165}$$

We differentiate the function in (8.165) to obtain

$$\begin{aligned} \dot\varphi(t) &= \frac{\frac{d}{dt}\left(\gamma(t)^\eta - y(t)\right)}{\gamma(t)^\eta - y(t)}\varphi(t) \\ &\quad + \varphi(t)(1+\varepsilon)k(t)^\eta \int_0^1 \left((1-s)C(t)^\eta + sk(t)^\eta y(t)\right)^\varepsilon ds, \end{aligned} \tag{8.166}$$

and hence

$$\begin{aligned} &\left(\gamma(t)^\eta - y(t)\right)\dot\varphi(t) \\ &= \left(\frac{d}{dt}\left(\gamma(t)^\eta - y(t)\right)\right)\varphi(t) \\ &\quad + \varphi(t)(1+\varepsilon)\left(C(t)^\eta - k(t)^\eta y(t)\right)\int_0^1 \left((1-s)C(t)^\eta + sk(t)^\eta y(t)\right)^\varepsilon ds \\ &= \left(\frac{d}{dt}\left(\gamma(t)^\eta - y(t)\right)\right)\varphi(t) \\ &\quad - \varphi(t)\int_0^1 \frac{\partial}{\partial s}\left((1-s)C(t)^\eta + sk(t)^\eta y(t)\right)^{1+\varepsilon} ds \\ &= \left(\frac{d}{dt}\left(\gamma(t)^\eta - y(t)\right)\right)\varphi(t) - \varphi(t)\left(k(t)y(t)^{1+\varepsilon} - C(t)\right) \end{aligned} \tag{8.167}$$

($y(t)$ satisfies equation (8.162))

$$
\begin{aligned}
&= \left(\frac{d}{dt}\left(\gamma(t)^{\eta} - y(t)\right)\right)\varphi(t) + \varphi(t)\dot{y}(t) \\
&= \left(\frac{d}{dt}\left(\gamma(t)^{\eta}\right)\right)\varphi(t) = 0 \qquad (8.168)
\end{aligned}
$$

where we used the fact that $\gamma(t)$ does not depend on $t \geqslant \tau$. Consequently, from (8.168) it follows that the function $\varphi(t)$ does not depend on $t \geqslant \tau$. From the definition of φ (see (8.165)) we see that

$$
\begin{aligned}
&y(t) - \gamma^{\eta} \qquad (8.169) \\
&= (y(\tau) - \gamma^{\eta})\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)^{\eta}((1-s)C(\rho)^{\eta} + sk(\rho)^{\eta}y(\rho))^{\varepsilon}d\rho\, ds\right) \\
&= (y(\tau) - \gamma^{\eta})\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)((1-s)\gamma^{\eta} + s\,y(\rho))^{\varepsilon}\,d\rho\, ds\right).
\end{aligned}
$$

Suppose $\tau < t$. From (8.169) we see that $y(\tau) > \gamma^{\eta}$ implies $y(\tau) > y(t) > \gamma^{\eta}$, and that $y(t)$ decreases to its limit. We also see that $y(\tau) < \gamma^{\eta}$ entails $y(\tau) < y(t) < \gamma^{\eta}$, and that $y(t)$ increases to its limit. If $y(\tau) \leqslant \gamma^{\eta}$, then equality (8.169) implies $y(t) \leqslant \gamma^{\eta}$ for all $t \geqslant \tau$. Even more is true:

$$
\begin{aligned}
&0 \leqslant \gamma^{\eta} - y(t) \qquad (8.170) \\
&= (\gamma^{\eta} - y(\tau))\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)^{\eta}((1-s)C(\rho)^{\eta} + sk(\rho)^{\eta}y(\rho))^{\varepsilon}d\rho\, ds\right) \\
&\leqslant \gamma^{\eta}\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)\left((1-s)\gamma^{\eta} + sy(\rho)\right)^{\varepsilon}d\rho\, ds\right) \\
&\leqslant \gamma^{\eta}\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)(1-s)^{\varepsilon}\gamma^{\varepsilon\eta}\,d\rho\, ds\right) \\
&= \gamma^{\eta}\exp\left(-\gamma^{\varepsilon\eta}\int_\tau^t k(\rho)d\rho\right).
\end{aligned}
$$

Next we put

$$
\Phi_{\varepsilon}(\tau,\rho) = (1+\varepsilon)\int_0^1\int_\tau^{\rho} k\left(\rho'\right)\left((1-s)\gamma^{\eta} + sy\left(\rho'\right)\right)^{\varepsilon}d\rho' ds. \qquad (8.171)
$$

If the function $y(t)$ solves equation (8.162), then (8.169) implies

$$
\left(y(\rho) - \gamma^{\eta}\right)e^{\Phi_{\varepsilon}(\tau,\rho)} = y(\tau) - \gamma^{\eta}, \qquad (8.172)
$$

and consequently we get

$$
e^{\varepsilon\Phi_{\varepsilon}(\tau,t)} - 1 = \int_\tau^t \frac{\partial}{\partial\rho}e^{\varepsilon\Phi_{\varepsilon}(\tau,\rho)}d\rho
$$

$$
\begin{aligned}
&= \varepsilon\,(1+\varepsilon)\int_\tau^t\int_0^1 k(\rho)\left((1-s)\gamma^\eta + sy(\rho)\right)^\varepsilon e^{\varepsilon\Phi_\varepsilon(\tau,\rho)}ds\,d\rho \\
&= \varepsilon\,(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)\left(\gamma^\eta e^{\Phi_\varepsilon(\tau,\rho)} + s\,(y(\tau)-\gamma^\eta)\right)^\varepsilon d\rho\,ds. \qquad (8.173)
\end{aligned}
$$

If $y(\tau) > \gamma^\eta$, then (8.173) implies

$$
\begin{aligned}
e^{\varepsilon\Phi_\varepsilon(\tau,t)} &= 1 + \varepsilon\,(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)\left(\gamma^\eta e^{\Phi_\varepsilon(\tau,\rho)} + s\,(y(\tau)-\gamma^\eta)\right)^\varepsilon d\rho\,ds \\
&\geqslant 1 + \varepsilon\int_\tau^t k(\rho)d\rho\,(y(\tau)-\gamma^\eta)^\varepsilon. \qquad (8.174)
\end{aligned}
$$

From (8.174) we infer

$$
\begin{aligned}
e^{\Phi_\varepsilon(\tau,t)} &\geqslant \left(1 + \varepsilon\int_\tau^t k(\rho)d\rho\,(y(\tau)-\gamma^\eta)^\varepsilon\right)^{1/\varepsilon} \\
&\geqslant \left(\varepsilon\int_\tau^t k(\rho)d\rho\right)^{1/\varepsilon}(y(\tau)-\gamma^\eta). \qquad (8.175)
\end{aligned}
$$

From (8.172) with $\rho = t$ together with (8.175) we see that

$$
0 \leqslant y(t) - \gamma^\eta \leqslant \left(\varepsilon\int_\tau^t k(\rho)d\rho\right)^{-1/\varepsilon}. \qquad (8.176)
$$

From (8.176) we see that (8.163) holds provided that $y(\tau) > \gamma^\eta$.

If $0 \leqslant y(\tau) < \gamma^\eta$ we proceed as follows. Again we use (8.173) to obtain

$$
\begin{aligned}
e^{\varepsilon\Phi_\varepsilon(\tau,t)} &= 1 + \varepsilon\,(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)\left(\gamma^\eta e^{\Phi_\varepsilon(\tau,\rho)} - s\,(\gamma^\eta - y(\tau))\right)^\varepsilon d\rho\,ds \\
&\geqslant 1 + \varepsilon\,(1+\varepsilon)\int_\tau^t k(\rho)\int_0^1 \left((1-s)\,(\gamma^\eta - y(\tau))\right)^\varepsilon ds\,d\rho \\
&= 1 + \varepsilon\int_\tau^t k(\rho)d\rho\,(\gamma^\eta - y(\tau))^\varepsilon. \qquad (8.177)
\end{aligned}
$$

Hence we see

$$
\begin{aligned}
e^{\Phi_\varepsilon(\tau,t)} &\geqslant \left(1 + \varepsilon\int_\tau^t k(\rho)d\rho\,(\gamma^\eta - y(\tau))^\varepsilon\right)^{1/\varepsilon} \\
&\geqslant \left(\varepsilon\int_\tau^t k(\rho)d\rho\right)^{1/\varepsilon}(\gamma^\eta - y(\tau)). \qquad (8.178)
\end{aligned}
$$

From (8.169) with $\gamma^\eta > y(\tau)$ together with (8.178) we then get

$$
\gamma^\eta - y(t) = (\gamma^\eta - y(\tau))\,e^{-\Phi_\varepsilon(\tau,t)} \leqslant \left(\varepsilon\int_\tau^t k(\rho)d\rho\right)^{-1/\varepsilon}. \qquad (8.179)
$$

Inequality (8.163) in Lemma 8.5 now follows from (8.176) and (8.179). The monotonicity properties of the function $t \mapsto y(t)$ follow from the equality in (8.169).

Proof of assertion (1). If $\int_\tau^\infty k(\rho)\,d\rho = \infty$, then from inequality (8.163) we get $\lim_{t\to\infty} y(t) = \gamma^\eta$.

Proof of (2). This assertion is a direct consequence of (8.169).

Proof of assertion (3). The following arguments show that $\lim_{t\to\infty} y(t) < \gamma^\eta$. Here the function $t \mapsto y(t)$ is a solution to the equation in equation (8.162).

The estimates in (8.182) and (8.186) below are particularly useful $\int_\tau^\infty k(\rho)\,d\rho < \infty$.

If $y(\tau) > \gamma^\eta$, then (8.169) with ρ instead of t implies $y(\rho) < y(\tau)$, $\rho > \tau$, and

$$
\begin{aligned}
&y(t) - \gamma^\eta && (8.180)\\
&= (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)^\eta\left((1-s)C(\rho)^\eta + sk(\rho)^\eta y(\rho)\right)^\varepsilon d\rho\,ds\right)\\
&= (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)\left((1-s)\gamma^\eta + sy(\rho)\right)^\varepsilon d\rho\,ds\right)\\
&\geqslant (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)\left((1-s)y(\rho) + sy(\rho)\right)^\varepsilon d\rho\,ds\right)\\
&= (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\int_\tau^t k(\rho)y(\rho)^\varepsilon\,d\rho\right) && (8.181)\\
&\geqslant (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\,y(\eta)^\varepsilon\int_\tau^t k(\rho)\,d\rho\right). && (8.182)
\end{aligned}
$$

Hence from (8.180) we obtain, with ρ instead of t,

$$
y(\rho) \geqslant \gamma^\eta + (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\int_\tau^\rho k\left(\rho'\right)y\left(\rho'\right)^\varepsilon\,d\rho'\right), \tag{8.183}
$$

and therefore

$$
(1-s)\gamma^\eta + sy(\rho) \geqslant \gamma^\eta + s\left(y(\tau) - \gamma^\eta\right)\exp\left(-(1+\varepsilon)\int_\tau^\rho k\left(\rho'\right)y\left(\rho'\right)^\varepsilon\,d\rho'\right), \tag{8.184}
$$

Again using (8.180) and (8.184) we then obtain:

$$
\begin{aligned}
&0 \leqslant y(t) - \gamma^\eta && (8.185)\\
&= (y(\tau) - \gamma^\eta)\exp\left(-(1+\varepsilon)\int_0^1\int_\tau^t k(\rho)^\eta((1-s)C(\rho)^\eta + sk(\rho)^\eta y(\rho))^\varepsilon d\rho\,ds\right)
\end{aligned}
$$

$$= (y(\tau) - \gamma^\eta) \exp\left(-(1+\varepsilon)\int_0^1 \int_\tau^t k(\rho)\left((1-s)\gamma^\eta + sy(\rho)\right)^\varepsilon d\rho\, ds\right)$$

$$\leqslant (y(\tau) - \gamma^\eta)$$

$$\exp\left\{-(1+\varepsilon)\int_0^1 \int_\tau^t k(\rho)\right.$$

$$\left.\left(\gamma^\eta + s\left(y(\tau)-\gamma^\eta\right)\, \exp\left(-\left(1+\varepsilon\right)\int_\tau^\rho k\left(\rho'\right) y\left(\rho'\right)^\varepsilon\, d\rho'\right)\right)^\varepsilon d\rho\, ds\right\}$$

(use the elementary inequality $(\gamma^\eta + a)^\varepsilon \geqslant 2^{(\varepsilon-1)\wedge 0}\left(\gamma^{\eta\varepsilon} + a^\varepsilon\right)$, $a > 0$, $\varepsilon > 0$)

$$\leqslant (y(\tau) - \gamma^\eta) \exp\left(-(1+\varepsilon)2^{(\varepsilon-1)\wedge 0}\gamma^{\eta\varepsilon}\int_0^1 \int_\tau^t k(\rho) d\rho\, ds\right)$$

$$\exp\left(-(1+\varepsilon)2^{(\varepsilon-1)\wedge 0}\int_0^1 \int_\tau^t k(\rho) s^\varepsilon \left(y(\tau)-\gamma^\eta\right)^\varepsilon\right.$$

$$\left.\exp\left(-\varepsilon\left(1+\varepsilon\right)\int_\tau^\rho k\left(\rho'\right) y\left(\rho'\right)^\varepsilon\, d\rho'\right) d\rho\, ds\right)$$

$$= (y(\tau) - \gamma^\eta) \exp\left(-(1+\varepsilon)2^{(\varepsilon-1)\wedge 0}\gamma^{\eta\varepsilon}\int_\tau^t k(\rho) d\rho\right)$$

$$\exp\left\{-\left(y(\tau)-\gamma^\eta\right)^\varepsilon 2^{(\varepsilon-1)\wedge 0}\right.$$

$$\left.\int_\tau^t k(\rho) \exp\left(-\varepsilon\left(1+\varepsilon\right)\int_\tau^\rho k\left(\rho'\right) y\left(\rho'\right)^\varepsilon\, d\rho'\right) d\rho\right\}$$

$$\leqslant (y(\tau) - \gamma^\eta) \exp\left(-(1+\varepsilon)2^{(\varepsilon-1)\wedge 0}\gamma^{\eta\varepsilon}\int_\tau^t k(\rho) d\rho\right)$$

$$\exp\left(-2^{(\varepsilon-1)\wedge 0}(y(\tau)-\gamma^\eta)^\varepsilon \int_\tau^t k(\rho) \exp\left(-\varepsilon(1+\varepsilon) y(\tau)^\varepsilon \int_\tau^\rho k(\rho') d\rho'\right) d\rho\right)$$

$$= (y(\tau) - \gamma^\eta) \exp\left(-(1+\varepsilon)2^{(\varepsilon-1)\wedge 0}\gamma^{\eta\varepsilon}\int_\tau^t k(\rho) d\rho\right)$$

$$\exp\left\{-\frac{2^{(\varepsilon-1)\wedge 0}}{\varepsilon(1+\varepsilon)}\left(\frac{y(\tau)-\gamma^\eta}{y(\tau)}\right)^\varepsilon \left(1-\exp\left(-\varepsilon(1+\varepsilon) y(\tau)^\varepsilon \int_\tau^t k(\rho') d\rho'\right)\right)\right\}$$

(use the elementary equality $1 - e^{-a} = \int_0^1 a e^{-sa} ds$, $a \geqslant 0$)

$$= (y(\tau) - \gamma^\eta) \exp\left(-(1+\varepsilon)2^{(\varepsilon-1)\wedge 0}\gamma^{\eta\varepsilon}\int_\tau^t k(\rho) d\rho\right)$$

$$\exp\left\{-2^{(\varepsilon-1)\wedge 0}\left(y(\tau)-\gamma^\eta\right)^\varepsilon \int_\tau^t k\left(\rho'\right)\, d\rho'\right.$$

$$\times \int_0^1 \exp\left(-\varepsilon\left(1+\varepsilon\right) sy(\tau)^{\varepsilon} \int_\tau^t k\left(\rho'\right) d\rho'\right) ds \Biggr\}. \tag{8.186}$$

From the inequalities in (8.181) and (8.186) the assertion in (3) follows. The proof of assertion (4) is similar, and therefore omitted. This completes the proof Lemma 8.5. □

The following lemma contains versions of the Gronwall inequality. It was used in the proof of Proposition 8.5.

Lemma 8.6. *Let $\varphi(t)$, $c_1(t)$, $\chi(t)$ and $\psi(t)$ be nonnegative continuous functions on the interval $[\tau, \infty)$ such that $\varphi(t) \leqslant \psi(t) + \chi(t) \int_\tau^t c_1(\rho)\varphi(\rho)d\rho$, $t \geqslant \tau$. Then*

$$\begin{aligned}\varphi(t) \leqslant \psi(t) + \chi(t) \int_\tau^t \sum_{j=0}^{n-1} \frac{\left(\int_\rho^t \chi\left(\rho'\right) c_1\left(\rho'\right) d\rho'\right)^j}{j!} c_1(\rho)\psi(\rho)d\rho \\ + \chi(t) \int_\tau^t \frac{\left(\int_\rho^t \chi\left(\rho'\right) c_1\left(\rho'\right) d\rho'\right)^n}{n!} c_1(\rho)\varphi(\rho)d\rho.\end{aligned} \tag{8.187}$$

From (8.187) it follows that:

$$\begin{aligned}\varphi(t) &\leqslant \psi(t) + \chi(t) \int_\tau^t \sum_{j=0}^{\infty} \frac{\left(\int_\rho^t \chi\left(\rho'\right) c_1\left(\rho'\right) d\rho'\right)^j}{j!} c_1(\rho)\psi(\rho)d\rho \\ &= \psi(t) + \chi(t) \int_\tau^t e^{\int_\rho^t \chi(\rho')c_1(\rho')d\rho'} c_1(\rho)\psi(\rho)d\rho.\end{aligned} \tag{8.188}$$

If $\psi(t) = \chi(t)\varphi(\tau) + \chi(t)\int_\tau^t z(\rho)d\rho$, then (8.188) implies:

$$\begin{aligned}\varphi(t) &\leqslant \chi(t)\varphi(\tau) + \chi(t) \int_\tau^t z(\rho)d\rho \\ &\quad + \chi(t) \int_\tau^t e^{\int_\rho^t \chi(\rho')c_1(\rho')d\rho'} \chi(\rho)c_1(\rho) \left(\varphi(\tau) + \int_\tau^\rho z\left(\rho''\right) d\rho''\right) d\rho \\ &= \chi(t)\varphi(\tau)e^{\int_\tau^t \chi(\rho)c_1(\rho)d\rho} + \chi(t) \int_\tau^t e^{\int_\rho^t \chi(\rho')c_1(\rho')d\rho'} z(\rho)d\rho.\end{aligned} \tag{8.189}$$

Remark 8.7. If the matrix $A(t) \neq 0$, then the operator $L(t)$ does not generate an analytic semigroup. This means that our theory is not directly applicable to the example in Proposition 8.4. We could make $C\left(t,\tau\right)$ state-dependent and $A(t)$ also. One way of doing this is by taking a unique solution to a stochastic differential equation:

$$dX(t) = b\left(t, X(t)\right) dt + \sigma\left(t, X(t)\right) dW(t), \quad t \geqslant \tau \geqslant 0 \tag{8.190}$$

and then defining the family $C(t,\tau,X(\tau))$ by $X(t) = C(t,\tau,X(\tau))$. The functional $C(t,\tau,X(\tau))$ then depends on the σ-field generated by $X(\tau)$ and $\mathcal{F}_t^\tau = \sigma(W(\rho) : \tau \leqslant \rho \leqslant t)$. The evolution $Y(\tau,t)$ is then defined by

$$Y(\tau,t) f(x) = \mathbb{E}\left[f(X(t)) \bigm| X(\tau) = x\right] = \mathbb{E}_{\tau,x}\left[f(C(t,\tau,X(\tau)))\right], \ t \geqslant \tau.$$

In this general setup we do not have explicit formulas anymore. Moreover, this choice of $C(t,\tau,x)$ does not give any $A(t)$, because the function $t \mapsto X(t) = C(t,\tau,X(\tau))$ is not differentiable in a classical sense. Of course it satisfies (8.190).

8.3.2 *Some stochastic differential equations*

We want to study the processes $t \mapsto X(t)$, $s \mapsto X^{t,A(t)}(s)$ and $t \mapsto X_0^{\tau,A(\tau)}(t)$, which are solutions to the following stochastic integral equations, and their inter-relationships:

$$\begin{aligned} X(t) &= C(t,\tau)X(\tau) + \int_\tau^t C(t,\rho)\,\sigma(\rho, X(\rho))\,dW(\rho), \\ X^{t,A(t)}(s) &= e^{sA(t)}X^{t,A(t)}(0) + \int_0^s e^{(s-\rho)A(t)}\sigma\left(t, X^{t,A(t)}(\rho)\right)dW(\rho), \quad \text{and} \\ X_0^{\tau,A(\tau)}(t) &= e^{(t-\tau)A(\tau)}X_0^{\tau,A(\tau)}(\tau) + \int_\tau^t e^{(\rho-\tau)A(\tau)}\sigma\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right)dW(\rho). \end{aligned} \tag{8.191}$$

For $t \geqslant s \geqslant \tau$ the matrix family $C(t,\tau)$ satisfies $C(t,\tau) = C(t,s)\,C(s,\tau)$, and the matrix family $A(\tau)$ is defined by $A(\tau) = \lim_{h\downarrow 0} \dfrac{C(\tau+h,\tau) - I}{h}$. In differential form the stochastic integral equations in (8.191) read as follows:

$$dX(t) = A(t)X(t)dt + \sigma(t, X(t))\,dW(t); \tag{8.192}$$

$$dX^{t,A(t)}(s) = A(t)X^{t,A(t)}(s)ds + \sigma\left(t, X^{t,A(t)}(s)\right)dW(s), \quad \text{and} \tag{8.193}$$

$$\begin{aligned} dX_0^{\tau,A(\tau)}(t) &= A(\tau)\left(X_0^{\tau,A(\tau)}(t) - \int_\tau^t e^{(\rho-\tau)A(\tau)}\sigma\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right)dW(\rho)\right)dt \\ &\quad + e^{(t-\tau)A(\tau)}\sigma\left(\tau, X_0^{\tau,A(\tau)}(t)\right)dW(t). \end{aligned} \tag{8.194}$$

We will consider the following exponential martingale

$$\mathcal{E}^\tau(t) = \exp\left(-\int_\tau^t b(\rho, X(\rho))\,dW(\rho) - \frac{1}{2}\int_\tau^t |b(\rho, X(\rho))|^2\,d\rho\right), \tag{8.195}$$

and its companions

$$\mathcal{E}^{t,A(t)}(s) \tag{8.196}$$

$$= \exp\left(-\int_0^s b\left(t, X^{t,A(t)}(\rho)\right) dW(\rho) - \frac{1}{2}\int_0^s \left|b\left(t, X^{t,A(t)}(\rho)\right)\right|^2 d\rho\right),$$

and

$$\begin{aligned}&\mathcal{E}_0^{\tau,A(\tau)}(t) \\ &= \exp\left(-\int_\tau^t b\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right) dW(\rho) - \frac{1}{2}\int_\tau^t \left|b\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right)\right|^2 d\rho\right).\end{aligned} \tag{8.197}$$

Instead of $\mathcal{E}^0(t)$ we write $\mathcal{E}(t)$. Put

$$M^\tau(t) = \int_\tau^t b(\rho, X(\rho))dW(\rho), \quad M^{t,A(t)}(s) = \int_0^s b\Big(t, X^{t,A(t)}(\rho)\Big)dW(\rho),$$

and

$$M_0^{\tau,A(\tau)}(t) = \int_\tau^t b\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right) dW(\rho). \tag{8.198}$$

Then by Itô calculus we have:

$$\begin{aligned} d\mathcal{E}^\tau(t) &= -\mathcal{E}^\tau(t)dM^\tau(t), \quad d\mathcal{E}^{t,A(t)}(s) = -\mathcal{E}^{t,A(t)}(s)dM^{t,A(t)}(s), \\ d\mathcal{E}_0^{\tau,A(\tau)}(t) &= -\mathcal{E}_0^{\tau,A(\tau)}(t)dM_0^{\tau,A(\tau)}(t). \end{aligned} \tag{8.199}$$

Let $f : \mathbb{R}^d \to \mathbb{C}$ be a C^2-function. Again employing Itô calculus shows:

$$\begin{aligned} &df\left(X(t)\right) \\ &= \sum_{k=1}^d D_k f\left(X(t)\right) dX_k(t) + \frac{1}{2}\sum_{j,k=1}^d Q_{j,k}\left(t, X(t)\right) D_j D_k f\left(X(t)\right) dt \\ &= \left(\left\langle \nabla f\left(X(t)\right), A(t)X(t)\right\rangle + \frac{1}{2}\sum_{j,k=1}^d Q_{j,k}\left(t, X(t)\right) D_j D_k f\left(X(t)\right)\right) dt \\ &\quad + \left\langle \nabla f\left(X(t)\right), \sigma\left(t, X(t)\right) dW(t)\right\rangle. \end{aligned} \tag{8.200}$$

By the same token we get

$$\begin{aligned} df\left(X^{t,A(t)}(s)\right) &= \sum_{k=1}^d D_k f\left(X^{t,A(t)}(s)\right) dX_k^{t,A(t)}(s) \\ &\quad + \frac{1}{2}\sum_{j,k=1}^d Q_{j,k}\left(t, X^{t,A(t)}(s)\right) D_j D_k f\left(X^{t,A(t)}(s)\right) ds \\ &= \left(\left\langle \nabla f\left(X^{t,A(t)}(s)\right), A(t)X^{t,A(t)}(s)\right\rangle\right. \end{aligned}$$

$$+\frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}\left(t, X^{t,A(t)}(s)\right) D_j D_k f\left(X^{t,A(t)}(s)\right)\Big) ds$$
$$+\left\langle \nabla f\left(X^{t,A(t)}(s)\right), \sigma\left(t, X^{t,A(t)}(s)\right) dW(s)\right\rangle. \quad (8.201)$$

In addition we have, again by Itô calculus,

$$\begin{aligned}
& df\left(X_0^{\tau,A(\tau)}(t)\right) \\
&= \sum_{k=1}^{d} D_k f\left(X_0^{\tau,A(\tau)}(t)\right) dX_{0,k}^{\tau,A(\tau)}(t) \\
&\quad + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}\left(\tau, X_0^{\tau,A(\tau)}(\tau) D_j D_k f\left(X_0^{\tau,A(\tau)}(t)\right)\right) dt \\
&= \Bigg(\left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), A(\tau) X_0^{\tau,A(\tau)}(t)\right\rangle \\
&\quad + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}\left(\tau, X_0^{\tau,A(\tau)}(t)\right) D_j D_k f\left(X_0^{\tau,A(\tau)}(t)\right)\Bigg) dt \\
&\quad + \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \sigma\left(\tau, X_0^{\tau,A(\tau)}(t)\right) dW(t)\right\rangle \\
&\quad - \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), A(\tau)\int_\tau^t e^{(\rho-\tau)A(\tau)}\sigma\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right) dW(\rho)\right\rangle dt.
\end{aligned} \quad (8.202)$$

We also need the covariation processes:

$$\begin{aligned}
&\langle \mathcal{E}(\cdot), f(X(\cdot))\rangle (t), \quad \left\langle \mathcal{E}^{t,A(t)}(\cdot), f\left(X^{t,A(t)}(\cdot)\right)\right\rangle (s), \quad \text{and} \\
&\left\langle \mathcal{E}_0^{\tau,A(\tau)}(\cdot), f\left(X_0^{\tau,A(\tau)}(\cdot)\right)\right\rangle (t).
\end{aligned} \quad (8.203)$$

The covariation process $\langle \mathcal{E}(\cdot), f(X(\cdot))\rangle (t)$ is determined by

$$d\langle \mathcal{E}(\cdot), f(X(\cdot))\rangle (t) = -\mathcal{E}(t)\langle \nabla f(X(t)), \sigma(t, X(t))\, b(t, X(t))\rangle\, dt. \quad (8.204)$$

The covariation process $\left\langle \mathcal{E}^{t,A(t)}(\cdot), f\left(X^{t,A(t)}(\cdot)\right)\right\rangle (s)$ is determined by

$$\begin{aligned}
& d\left\langle \mathcal{E}^{t,A(t)}(\cdot), f\left(X^{t,A(t)}(\cdot)\right)\right\rangle (s) \quad (8.205) \\
&= -\mathcal{E}^{t,A(t)}(s)\left\langle \nabla f\left(X^{t,A(t)}(s)\right), \sigma\left(t, X^{t,A(t)}(s)\right) b\left(t, X^{t,A(t)}(s)\right)\right\rangle ds.
\end{aligned}$$

Likewise the covariation process $\left\langle \mathcal{E}_0^{\tau,A(\tau)}(\cdot), f\left(X_0^{\tau,A(\tau)}(\cdot)\right)\right\rangle (t)$ is determined by

$$d\left\langle \mathcal{E}_0^{\tau,A(\tau)}(\cdot), f\left(X_0^{\tau,A(\tau)}(\cdot)\right)\right\rangle (t) \quad (8.206)$$

$$= -\mathcal{E}_0^{\tau,A(\tau)}(t)\left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \sigma\left(\tau, X_0^{\tau,A(\tau)}(t)\right) b\left(\tau, X_0^{\tau,A(\tau)}(t)\right)\right\rangle dt.$$

Next we calculate the stochastic differential of the processes

$$\mathcal{E}(t) f\left(X(t)\right), \quad \mathcal{E}^{t,A(t)}(s) f\left(X^{t,A(t)}(s)\right) \quad \text{and} \quad \mathcal{E}_0^{\tau,A(\tau)}(t) f\left(X_0^{\tau,A(\tau)}(t)\right).$$

Using Itô calculus, the equality in (8.195), and the first equality in (8.198) and in (8.199), in conjunction with (8.200) and (8.204) shows

$$\begin{aligned}
&d\left(\mathcal{E}(t) f\left(X(t)\right)\right) \\
&= \left(d\mathcal{E}(t)\right) f\left(X(t)\right) + \mathcal{E}(t) df\left(X(t)\right) + d\left\langle \mathcal{E}(\cdot), f\left(X(\cdot)\right)\right\rangle(t) \\
&= -\mathcal{E}(t) f\left(X(t)\right) b\left(t, X(t)\right) dW(t) \\
&\quad + \mathcal{E}(t)\left(\left\langle \nabla f\left(X(t)\right), A(t)X(t)\right\rangle + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}\left(t, X(t)\right) D_j D_k f\left(X(t)\right)\right) dt \\
&\quad + \mathcal{E}(t)\left\langle \nabla f\left(X(t)\right), \sigma\left(t, X(t)\right) dW(t)\right\rangle \\
&\quad - \mathcal{E}(t)\left\langle \nabla f\left(X(t)\right), \sigma\left(t, X(t)\right) b\left(t, X(t)\right)\right\rangle dt \\
&= -\mathcal{E}(t) f\left(X(t)\right) b\left(t, X(t)\right) dW(t) + \mathcal{E}(t)\left\langle \nabla f\left(X(t)\right), \sigma\left(t, X(t)\right) dW(t)\right\rangle \\
&\quad + \mathcal{E}(t) L_b(t) f\left(X(t)\right) dt
\end{aligned} \tag{8.207}$$

where with $Q(t,x) = \sigma(t,x)\sigma(t,x)^*$ we wrote

$$L_b(t) f(x) = \left\langle \nabla f(x), A(t)x - \sigma(t,x) b(t,x)\right\rangle + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}(t,x) D_j D_k f(x). \tag{8.208}$$

Put

$$Q_{t,A(t)}(s) = \int_0^s e^{\rho A(t)}\sigma(t,x)\sigma(t,x)^* e^{\rho A(t)^*} d\rho = \int_0^s e^{\rho A(t)} Q(t,x) e^{\rho A(t)^*} d\rho,$$

and

$$Y(\tau,t) f(x) = \mathbb{E}\left[\mathcal{E}^{\tau}(t) f\left(X(t)\right) \,\middle|\, X(\tau) = x\right]. \tag{8.209}$$

Then

$$Y(\tau,s) Y(s,t) f(x) = Y(\tau,t) f(x), \quad f \in C_b(E),\ x \in E,\ \tau \leqslant s \leqslant t. \tag{8.210}$$

Next we will calculate the stochastic derivative of the process

$$s \mapsto \mathcal{E}^{t,A(t)}(s) f\left(X^{t,A(t)}(s)\right).$$

More precisely, upon employing Itô calculus, the martingale in (8.196), and the second martingale in (8.198) and in (8.199), in conjunction with (8.201) and (8.205) we obtain

$$d\left(\mathcal{E}^{t,A(t)}(s) f\left(X^{t,A(t)}(s)\right)\right)$$

$$
\begin{aligned}
&= \left(d\mathcal{E}^{t,A(t)}(s)\right) f\left(X^{t,A(t)}(s)\right) + \mathcal{E}^{t,A(t)}(s) df\left(X^{t,A(t)}(s)\right) \\
&\quad + d\left\langle \mathcal{E}^{t,A(t)}(\cdot), f\left(X^{t,A(t)}(\cdot)\right)\right\rangle(s) \\
&= -\mathcal{E}^{t,A(t)}(s) f\left(X^{t,A(t)}(s)\right) b\left(\tau, X^{t,A(t)}(s)\right) dW(s) \\
&\quad + \mathcal{E}^{t,A(t)}(s)\left(\left\langle \nabla f\left(X^{t,A(t)}(s)\right), A(\tau)X^{t,A(t)}(s)\right\rangle\right. \\
&\quad \left. + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}\left(\tau, X^{t,A(t)}(s)\right) D_j D_k f\left(X^{t,A(t)}(s)\right)\right) ds \\
&\quad + \mathcal{E}^{t,A(t)}(s)\left\langle \nabla f\left(X^{t,A(t)}(s)\right), \sigma\left(\tau, X^{t,A(t)}(s)\right) dW(s)\right\rangle \\
&\quad - \mathcal{E}^{t,A(t)}(s)\left\langle \nabla f\left(X^{t,A(t)}(s)\right), \sigma\left(\tau, X^{t,A(t)}(s)\right) b\left(\tau, X^{t,A(t)}(s)\right)\right\rangle ds \\
&= -\mathcal{E}^{t,A(t)}(s) f\left(X^{t,A(t)}(s)\right) b\left(\tau, X^{t,A(t)}(s)\right) dW(s) \\
&\quad + \mathcal{E}^{t,A(t)}(s)\left\langle \nabla f\left(X^{t,A(t)}(s)\right), \sigma\left(\tau, X^{t,A(t)}(s)\right) dW(s)\right\rangle \\
&\quad + \mathcal{E}^{t,A(t)}(s) L_b(t) f\left(X^{t,A(t)}(s)\right) ds \qquad (8.211)
\end{aligned}
$$

where $L_b(t)$ is as in (8.208).

In a quite similar manner we obtain the stochastic differential of the process

$$
t \mapsto \mathcal{E}_0^{\tau,A(\tau)}(t) f\left(X_0^{\tau,A(\tau)}(t)\right).
$$

Upon employing Itô calculus, the equality in (8.197), and the third martingale in (8.198) and in (8.199), in conjunction with (8.202) and (8.206) we get

$$
\begin{aligned}
& d\left(\mathcal{E}_0^{\tau,A(\tau)}(t) f\left(X_0^{\tau,A(\tau)}(t)\right)\right) \\
&= \left(d\mathcal{E}_0^{\tau,A(\tau)}(t)\right) f\left(X_0^{\tau,A(\tau)}(t)\right) + \mathcal{E}_0^{\tau,A(\tau)}(t) df\left(X_0^{\tau,A(\tau)}(t)\right) \\
&\quad + d\left\langle \mathcal{E}_0^{\tau,A(\tau)}(\cdot), f\left(X_0^{\tau,A(\tau)}(\cdot)\right)\right\rangle(t) \\
&= -\mathcal{E}_0^{\tau,A(\tau)}(t) f\left(X_0^{\tau,A(\tau)}(t)\right) b\left(\tau, X_0^{\tau,A(\tau)}(t)\right) dW(t) \\
&\quad + \mathcal{E}_0^{\tau,A(\tau)}(t)\left(\left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), A(\tau)X_0^{\tau,A(\tau)}(t)\right\rangle\right. \\
&\quad \left. + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}\left(\tau, X_0^{\tau,A(\tau)}(t)\right) D_j D_k f\left(X_0^{\tau,A(\tau)}(t)\right)\right) dt
\end{aligned}
$$

$$
\begin{aligned}
&+ \mathcal{E}_0^{\tau,A(\tau)}(t) \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \sigma\left(\tau, X_0^{\tau,A(\tau)}(t)\right) dW(t) \right\rangle \\
&- \mathcal{E}_0^{\tau,A(\tau)}(t) \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \right. \\
&\qquad\qquad \left. A(\tau) \int_\tau^t e^{(\rho-\tau)A(\tau)} \sigma\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right) dW(\rho) \right\rangle dt \\
&- \mathcal{E}_0^{\tau,A(\tau)}(t) \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \sigma\left(\tau, X_0^{\tau,A(\tau)}(t)\right) b\left(\tau, X_0^{\tau,A(\tau)}(t)\right) \right\rangle dt \\
= &-\mathcal{E}_0^{\tau,A(\tau)}(t) f\left(X_0^{\tau,A(\tau)}(t)\right) b\left(\tau, X_0^{\tau,A(\tau)}(t)\right) dW(t) \\
&+ \mathcal{E}_0^{\tau,A(\tau)}(t) \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \sigma\left(\tau, X_0^{\tau,A(\tau)}(t)\right) dW(t) \right\rangle \\
&- \mathcal{E}_0^{\tau,A(\tau)}(t) \left\langle \nabla f\left(X_0^{\tau,A(\tau)}(t)\right), \right. \\
&\qquad\qquad \left. A(\tau) \int_\tau^t e^{(\rho-\tau)A(\tau)} \sigma\left(\tau, X_0^{\tau,A(\tau)}(\rho)\right) dW(\rho) \right\rangle dt \\
&+ \mathcal{E}_0^{\tau,A(\tau)}(t) L_b(\tau) f\left(X_0^{\tau,A(\tau)}(t)\right) dt
\end{aligned} \tag{8.212}
$$

where $L_b(\tau)$ is as in (8.208):

$$
\begin{aligned}
L_b(\tau)f(x) = &\left\langle \nabla f(x), A(\tau)x - \sigma(\tau,x)\, b(\tau,x) \right\rangle \\
&+ \frac{1}{2} \sum_{j,k=1}^{d} Q_{j,k}(\tau,x)\, D_j D_k f(x).
\end{aligned} \tag{8.213}
$$

Next let $s \mapsto X^{t,A(t)}(s)$ be the solution to the stochastic integral equation:

$$
X^{t,A(t)}(s) = e^{sA(t)} X^{t,A(t)}(0) + \int_0^s e^{(s-\rho)A(t)} \sigma\left(t, X^{t,A(t)}(\rho)\right) dW(\rho), \tag{8.214}
$$

which is equivalent to

$$
dX^{t,A(t)}(s) = A(t) X^{t,A(t)}(s) ds + \sigma\left(t, X^{t,A(t)}(s)\right) dW(s), \tag{8.215}
$$

which is the same as the second in (8.191) and which in differential form is given in (8.193). In terms of the exponential martingale $s \mapsto \mathcal{E}^{t,A(t)}(s)$ defined in (8.196) the semigroup $e^{sL_b(t)}$, $s \geqslant 0$, is given by:

$$
e^{sL_b(t)} f(x) = \mathbb{E}\left[\mathcal{E}^{t,A(t)}(s) f\left(X^{t,A(t)}(s)\right) \mid X^{t,A(t)}(0) = x \right]. \tag{8.216}
$$

We also want give conditions in order that for every $\mu \in M\left(\mathbb{R}^d\right)$ the limit

$$
\lim_{t\to\infty} \operatorname{Var}\left(L_b(t)^* Y(\tau,t)^* \mu \right) = 0. \tag{8.217}
$$

We suppose that the coefficients $b(t,x) = b(t)$ and $\sigma(t,x) = \sigma(t)$ only depend on time. Then the (formal) adjoint of the operator $L_b(t)$ can be written as follows:

$$
\begin{aligned}
&L_b(t)^* f(x) \qquad (8.218)\\
&= -\langle \nabla f(x), A(t)x - \sigma(t)b(t)\rangle - \operatorname{tr}(A(t))\, f(x) + \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}(t) D_j D_k f(x).
\end{aligned}
$$

Put $Q_C(\tau,t) = \int_\tau^t C(t,\rho)\sigma(\rho)\sigma(\rho)^* C(t,\rho)^* d\rho$. Then for the evolution family $Y(\tau,t)$ we have:

$$
\begin{aligned}
&Y(\tau,t) f(x)\\
&= \frac{1}{(2\pi)^{d/2}} \int e^{-\frac{1}{2}|y|^2} f\left(C(t,\tau)x - \int_\tau^t C(t,\rho)\sigma(\rho)b(\rho)d\rho - \left(Q_C(\tau,t)\right)^{1/2} y\right) dy\\
&= \frac{1}{(2\pi)^{d/2} \det\left(Q_C(\tau,t)\right)^{1/2}} \qquad (8.219)\\
&\int \exp\left(-\frac{1}{2} \left| Q_C(\tau,t)^{-1/2} \left(C(t,\tau)x - \int_\tau^t C(t,\rho)\sigma(\rho)b(\rho)d\rho - y\right)\right|^2 \right) f(y)dy.
\end{aligned}
$$

Next suppose that the coefficients $b(t) = b(t,x)$ and $\sigma(t) = \sigma(t,x)$ only depend on the time t, and put

$$
g_s(x) = \frac{1}{(2\pi)^{d/2} \det\left(Q_{t,A(t)}(s)\right)^{1/2}} \exp\left(-\frac{1}{2} \left\langle Q_{t,A(t)}(s)^{-1} x, x\right\rangle \right).
$$

Then

$$
\widehat{g}_s(\xi) = \exp\left(-\frac{1}{2} \left\langle Q_{t,A(t)}(s)\xi, \xi\right\rangle \right),
$$

and hence by the Fourier inverse formula

$$
\begin{aligned}
&e^{sL_b(t)} f(x)\\
&= \int \mathbb{E}\left[\exp\left(-\langle b(t), W(s)\rangle - \frac{1}{2}|b(t)|^2 s\right) \right.\\
&\qquad \left. \times f\left(e^{sA(t)}x + \int_0^s e^{(s-\rho)A(t)}\sigma(t) dW(\rho)\right)\right]\\
&= \frac{1}{(2\pi)^d} \int \mathbb{E}\left[\exp\left(-\langle b(t), W(s)\rangle - \frac{1}{2}|b(t)|^2 s\right)\right.\\
&\qquad \left. \exp\left(i \left\langle \xi, e^{sA(t)}x + \int_0^s e^{(s-\rho)A(t)}\sigma(t)dW(\rho)\right\rangle\right)\right] \widehat{f}(\xi) d\xi\\
&= \frac{1}{(2\pi)^d} \int \exp\left(i\left\langle \xi, e^{sA(t)}x - \int_0^s e^{(s-\rho)A(t)} d\rho\, \sigma(t)b(t)\right\rangle\right)
\end{aligned}
$$

$$\exp\left(-\frac{1}{2}\left\langle\int_0^s e^{(s-\rho)A(t)}\sigma(t)\sigma(t)^* e^{(s-\rho)A(t)^*}\,d\rho\,\xi,\xi\right\rangle\right)\widehat{f}(\xi)d\xi$$

$$= \frac{1}{(2\pi)^d}\int \exp\left(i\left\langle \xi, e^{sA(t)}x - \int_0^s e^{\rho A(t)}d\rho\,\sigma(t)b(t)\right\rangle\right)$$

$$\exp\left(-\frac{1}{2}\left\langle\int_0^s e^{\rho A(t)}\sigma(t)\sigma(t)^* e^{\rho A(t)^*}\,d\rho\,\xi,\xi\right\rangle\right)\widehat{f}(\xi)d\xi$$

$$= \frac{1}{(2\pi)^d}\int \exp\left(i\left\langle \xi, e^{sA(t)}x - \int_0^s e^{\rho A(t)}d\rho\,\sigma(t)b(t)\right\rangle\right)\widehat{g_s * f}(\xi)d\xi$$

$$= g_s * f\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}d\rho\right)$$

$$= \frac{1}{(2\pi)^{d/2}\left(\det Q_{t,A(t)}(s)\right)^{1/2}}$$

$$\int \exp\left(-\frac{1}{2}\left|Q_{t,A(t)}(s)^{-1/2}\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right)\right|^2\right) f(y)\,dy$$

$$= \frac{1}{(2\pi)^{d/2}}\int e^{-\frac{1}{2}|y|^2} f\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - \left(Q_{t,A(t)}(s)\right)^{1/2} y\right)dy. \tag{8.220}$$

From the representation in (8.220) it follows that the operator $L_b(t)$ does not generate a bounded analytic semigroup. In fact if $f \in C_b\left(\mathbb{R}^d\right)$ is such that its first and second derivative is also continuous and bounded, then we have

$$L_b(t)e^{sL_b(t)}f(x)$$

$$= e^{sL_b(t)}L_b(t)f(x) = L_b(t)^* g_s * f\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}d\rho\right)$$

$$= \int L_b(t)^* g_s\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - \cdot\right)(y) f(y)\,dy$$

$$= -\sum_{j=1,k=1}^{d}\int \frac{\partial g_s}{\partial y_j}\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right)$$

$$\times A_{j,k}(t)\left(y_k - \sum_{\ell=1}^{d}\sigma_{k,\ell}(t)b_\ell(t)\right) f(y)\,dy$$

$$- \operatorname{tr}\left(A(t)\right)\int g_s\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right) f(y)\,dy$$

$$+ \frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}(t)\int \frac{\partial^2 g_s}{\partial y_j \partial y_k}\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right) f(y)\,dy$$

$$= -\sum_{j=1,k=1}^{d}\int \frac{\partial g_s(y)}{\partial y_j}\left(A(t)\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y - \sigma(t)b(t)\right)\right)_j$$
$$\times f\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right)dy$$
$$-\operatorname{tr}(A(t))\int g_s(y)\, f\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right)dy$$
$$+\frac{1}{2}\sum_{j,k=1}^{d} Q_{j,k}(t)\int \frac{\partial^2 g_s(y)}{\partial y_j \partial y_k} f\left(e^{sA(t)}x - \int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - y\right)dy. \tag{8.221}$$

All terms in (8.221) are uniformly bounded in x except the very first one, which grows like a constant times $|x|$. In order that the operator $L_b(t)$ generates a bounded analytic semigroup it is necessary and sufficient that $\sup_{s>0}\left\|sL_b(t)e^{sL_b(t)}\right\| < \infty$ and $\sup_{s>0}\left\|e^{sL_b(t)}\right\| < \infty$: see e.g. [Engel and Nagel (2000)].

Suppose that the real parts of the eigenvalues of the matrix $A(t)$ are strictly negative. From (8.220) it follows that the measure

$$B \mapsto$$
$$\frac{1}{(2\pi)^{d/2}}\int e^{-\frac{1}{2}|y|^2}\mathbf{1}_B\left(-\lim_{s\to\infty}\int_0^s e^{\rho A(t)}\sigma(t)b(t)d\rho - \lim_{s\to\infty}\left(Q_{t,A(t)}(s)\right)^{1/2}y\right)dy$$

serves as an invariant measure for the semigroup $e^{sL_b(t)}$, $s \geqslant 0$. Using the processes $X(t)$, $X^{\tau,A(\tau)}(t)$, and $X_0^{\tau,A(\tau)}(t)$, $t \geqslant \tau$, we introduce the filtered probability spaces $(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x})$ and $\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}^{(0)}\right)$. Here the σ-field $\mathcal{F}_t^\tau$, $\tau \leqslant t$, is generated by the variables $W(\rho)$, $\tau \leqslant \rho \leqslant t$. Let the variable F be $\mathcal{F}_t^\tau$-measurable. Then we put

$$\mathbb{E}_{\tau,x}[F] = \mathbb{E}\left[\mathcal{E}^\tau(t)F \,\middle|\, X(\tau) = x\right]. \tag{8.222}$$

On the other hand the definition of $\mathbb{P}_{\tau,x}^{(0)}$ is more of a challenge. First we take F which is measurable with respect to $\mathcal{F}_t^\tau$ of the form $F = \prod_{j=1}^n f_j\left(X^{\tau,A(\tau)}(t_j)\right)$. Then we put

$$\mathbb{E}_{\tau,x}^{(0)}[F] = \mathbb{E}\left[\mathcal{E}^{\tau,A(\tau)}(t)\prod_{j=1}^n f_j\left(X^{\tau,A(\tau)}(t_j)\right) \mid X(\tau) = x\right]$$
$$= \mathbb{E}\left[\mathcal{E}_0^{\tau,A(\tau)}(t)\prod_{j=1}^n f_j\left(X_0^{\tau,A(\tau)}(t_j)\right) \mid X(\tau) = x\right]. \tag{8.223}$$

Example 8.4. Another not too artificial example is the adjoint of the form

$$L(t) = \frac{1}{2}\sum_{j,k=1}^{d} a_{j,k}\left(t,x\right)\frac{\partial^2}{\partial x_j x_k} + \sum_{j=1}^{d} b_j(t,x)\frac{\partial}{\partial x_j},$$

defined on a dense subspace of the space $C_0\left(\mathbb{R}^d\right)$, i.e. the space of all bounded continuous functions with zero boundary conditions. The least that is required for the square matrix $\left(a_{j,k}\left(t,x\right)\right)_{j,k=1}^{d}$ is that it is invertible, symmetric and positive-definite. We also observe that, for such a choice of the coefficients $a_{j,k}(t,x)$ the operator $L(t)$ satisfies the following maximum principle. For any function $f \in C_0\left(\mathbb{R}^d\right)$ belonging to the domain $D\left(L(t)\right)$ there exists a point $(x_0, y_0) \in \mathbb{R}^d \times \mathbb{R}^d$ such that $\sup\left\{|f(x) - f(y)|\,;\,(x,y) \in \mathbb{R}^d \times \mathbb{R}^d\right\} = \left|f\left(x_0\right) - f\left(y_0\right)\right|$, such that the next inequality holds:

$$\Re\left\{\left(\overline{f\left(x_0\right)} - \overline{f\left(y_0\right)}\right)\left(L(t)f\left(x_0\right) - L(t)f\left(y_0\right)\right)\right\} \leqslant 0. \tag{8.224}$$

Since the function $(x,y) \mapsto |f(x) - f(y)|$ attains its maximum at (x_0, y_0) it follows that $\nabla f\left(x_0\right) = \nabla f\left(y_0\right) = 0$. It also follows that the function $x \mapsto \Re\left(\left(\overline{f\left(x_0\right)} - \overline{f\left(y_0\right)}\right)\left(f(x) - f\left(y_0\right)\right)\right)$ attains it maximum at x_0. Hence, by inequality (8.232) below we see that

$$\begin{aligned}
&\Re\left(\overline{f\left(x_0\right)} - \overline{f\left(y_0\right)}\right)L(t)f\left(x_0\right) \\
&= \Re\left(L(t)\left(\overline{f\left(x_0\right)} - \overline{f\left(y_0\right)}\right)\left(f(\cdot) - f\left(y_0\right)\right)\right)\left(x_0\right) \leqslant 0,
\end{aligned} \tag{8.225}$$

where the functions by the same token we also have:

$$-\Re\left(\overline{f\left(x_0\right)} - \overline{f\left(y_0\right)}\right)L(t)f\left(y_0\right) \leqslant 0. \tag{8.226}$$

From (8.224) we infer for $\alpha \in \mathbb{C}$, $\lambda \geqslant 0$, and $f \in D\left(L(t)\right)$ the string of inequalities:

$$\begin{aligned}
&4\left\|\lambda\left(f - \alpha\mathbf{1}\right) - L(t)f\right\|_\infty^2 \\
&\geqslant \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}\left|\lambda\left(f(x) - f(y)\right) - L(t)f(x) + L(t)f(y)\right|^2 \\
&= \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}\left\{|\lambda|^2\left|f(x) - f(y)\right|^2\right. \\
&\qquad \left. - 2\lambda\Re\left\{\left(\overline{f(x)} - \overline{f(y)}\right)\left(L(t)f(x) - L(t)f(y)\right)\right\} + \left|L(t)f(x) - L(t)f(y)\right|^2\right\} \\
&\geqslant |\lambda|^2\left|f(x_0) - f(y_0)\right|^2 - 2\lambda\Re\left\{\left(\overline{f(x_0)} - \overline{f(y_0)}\right)\left(L(t)f(x_0) - L(t)f(y_0)\right)\right\} \\
&\qquad + \left|L(t)f\left(x_0\right) - L(t)f\left(y_0\right)\right|^2
\end{aligned}$$

$$\geqslant |\lambda|^2 \left|f(x_0) - f(y_0)\right|^2 = |\lambda|^2 \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d} \left|f(x) - f(y)\right|^2$$
$$\geqslant |\lambda|^2 \inf_{\alpha\in\mathbb{C}} \|f - \alpha\mathbf{1}\|_\infty^2 . \tag{8.227}$$

From the inequalities in (8.227) we obtain for $\Re\lambda \geqslant 0$ and $f \in D(L(t))$:

$$2 \inf_{\alpha\in\mathbb{C}} \|\lambda f - L(t)f - \alpha\mathbf{1}\|_\infty \geqslant \lambda \inf_{\alpha\in\mathbb{C}} \|f - \alpha\mathbf{1}\|_\infty . \tag{8.228}$$

A similar argument shows that for $\lambda \geqslant 0$ and $f \in D(L(t))$ we also have:

$$\|\lambda f - L(t)f\|_\infty \geqslant \lambda \|f\|_\infty , \tag{8.229}$$

provided that for every function $f \in C_0\left(\mathbb{R}^d\right)$ belonging to the domain $D(L(t))$ there exists a point $x_0 \in \mathbb{R}^d$ such that $\sup\left\{|f(x)|\,;\, x \in \mathbb{R}^d\right\} = |f(x_0)|$, and such that $\Re\left\{\overline{f(x_0)}L(t)f(x_0)\right\} \leqslant 0$. In fact the operators $L(t)$ satisfy the maximum principle in the sense that $\Re\left(L(t)f(x_0)\right) \leqslant 0$ whenever $f \in D(L(t))$ and $x_0 \in \mathbb{R}^d$ is such that $\Re f(x_0) = \sup_{x\in\mathbb{R}^d} \Re f(x)$. One way of seeing this directly runs as follows. Let $f \in D(L(t))$. If $x_0 \in \mathbb{R}^d$ is such that $\Re f(x_0) = \sup_{x\in\mathbb{R}^d} \Re f(x)$. Then $\Re\langle x - x_0, \nabla f(x_0)\rangle = 0$, and thus, for all $x \in \mathbb{R}^d$,

$$\begin{aligned}\Re f(x) &= \Re f(x_0) + \langle x - x_0, \nabla\Re f(x_0)\rangle \\ &\qquad + \int_0^1 (1-s) \sum_{j,k=1}^d (x_j - x_{0,j})(x_k - x_{0,k}) \frac{\partial^2\Re f}{\partial x_j\partial x_k}((1-s)x_0 + sx)\,ds \\ &= \Re f(x_0) + \int_0^1 (1-s) \sum_{j,k=1}^d (x_j - x_{0,j})(x_k - x_{0,k}) \\ &\qquad\qquad \times \frac{\partial^2\Re f}{\partial x_j\partial x_k}((1-s)x_0 + sx)\,ds.\end{aligned} \tag{8.230}$$

From (8.230) and the fact that the function $\Re f$ attains its maximum at x_0 we see that

$$\Re\int_0^1 (1-s) \sum_{j,k=1}^d (x_j - x_{0,j})(x_k - x_{0,k}) \frac{\partial^2}{\partial x_j\partial x_k} f((1-s)x_0 + sx)\,ds \leqslant 0. \tag{8.231}$$

From the inequality in (8.231) it easily follows that the Hessian $D^2\Re f(x_0)$ which is the matrix with entries $\dfrac{\partial^2}{\partial x_j\partial x_k}\Re f(x_0)$ is negative-definite: i.e. it is symmetric and its eigenvalues are less than or equal to 0. Since the matrix $a(t, x_0) := \left(a_{j,k}(t, x_0)\right)_{j,k=1}^d$ is positive-definite (i.e. its eigenvalues are nonnegative and the matrix is symmetric) and the functions $b_j(t,x)$, $1 \leqslant j \leqslant d$, are real-valued, we infer that

$$\begin{aligned}\Re L(t) f\left(x_0\right) &= \sum_{j,k=1}^{d} a_{j,k}\left(t, x_0\right) \frac{\partial^2}{\partial x_j \partial x_k} \Re f\left(x_0\right) + \sum_{j=1}^{d} b_j\left(t, x_0\right) \frac{\partial}{\partial x_j} \Re f\left(x_0\right) \\ &= \operatorname{Tr}\left(a\left(t, x_0\right) D^2 \Re f\left(x_0\right)\right) \\ &= \operatorname{Tr}\left(\sqrt{a\left(t, x_0\right)} D^2 \Re f\left(x_0\right) \sqrt{a\left(t, x_0\right)}\right) \leqslant 0. \end{aligned} \tag{8.232}$$

(This notation was also used in formula (1.143) in Chapter 1.) The matrix $\sqrt{a\left(t, x_0\right)}$ is a positive-definite matrix with its square equal to $a\left(t, x_0\right)$. In addition, we used the fact that in (8.232) the identity

$$\sum_{j,k=1}^{d} a_{j,k}\left(t, x_0\right) \frac{\partial^2}{\partial x_j \partial x_k} \Re f\left(x_0\right)$$

can be interpreted as $\operatorname{Tr}\left(a\left(t, x_0\right) D^2 \Re f\left(x_0\right)\right)$. It follows that the operators $L(t)$ generate analytic semigroups $e^{sL(t)}$ where $s \in \mathbb{C}$ belongs to a sector with angle opening, which may be chosen independently of t provided that

$$\sup_{t>0} \sup_{s>0} \sup_{x \in \mathbb{R}^d} s \left| \frac{\partial}{\partial s} P_{L(t)}\left(s, x, \cdot\right) \right| \left(\mathbb{R}^d\right) < \infty.$$

Here the Markov transition function $P_{L(t)}\left(s, x, B\right)$, $(s,x) \in [0,\infty) \times \mathbb{R}^d$, $B \in \mathcal{B}_{\mathbb{R}^d}$, $t \geqslant 0$, is determined by the equality

$$e^{sL(t)} f(x) = \int_{\mathbb{R}^d} f(y) P_{L(t)}\left(s, x, dy\right), \quad f \in C_b\left(\mathbb{R}^d\right).$$

For the reason why, see the inequality in (8.100) and the equality in (8.101). Then it follows that there exist a constant C and an angle $\frac{1}{2}\pi < \beta < \pi$ again independent of t such that

$$|\lambda| \left\| (\lambda I - L(t))^{-1} \right\| \leqslant C, \quad \text{for all } \lambda \in \mathbb{C} \text{ with } |\arg(\lambda)| \leqslant \beta. \tag{8.233}$$

For a proof see Theorem 8.8 and its corollaries 8.4 and 8.5. Let $e^{sL(t)}$, $s \geqslant 0$, be the (analytic) semigroup generated by the operator $L(t)$. Then the (unbounded) inverse of the operator $-L(t)$ is given by the strong integral $f \mapsto \int_0^\infty e^{sL(t)} f\, ds$. From (8.228) it follows that for $\mu \in M_0\left(\mathbb{R}^d\right)$ and $\lambda > 0$ the inequality

$$\lambda \left| \left\langle g, \left(\lambda I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)^{-1} \mu \right\rangle \right| \leqslant \|g\|_\infty \operatorname{Var}(\mu), \tag{8.234}$$

holds whenever the function g is of the form $g = \lambda f - L(t) f$, with $f \in D(L(t))$. Here $M_0(\mathbb{R}^d)$ is the space of all complex Borel measures μ on

$\mathbb{R}^d$ with the property that $\mu(\mathbb{R}^d) = 0$. Suppose that $\mathrm{Var}\left(e^{sL(t)^*}\mu\right) \leqslant c(t)e^{-2\omega(t)s}\mathrm{Var}(\mu)$ for all $\mu \in M_0\left(\mathbb{R}^d\right)$ and $s \geqslant 0$. Then for $\Re\lambda \geqslant \omega(t)$, $g \in C_0\left(\mathbb{R}^d\right)$ and $\mu \in M_0(\mathbb{R}^d)$ we have

$$\begin{aligned}&(\lambda - 2\omega(t))\left\langle g, \left((\lambda - 2\omega(t))\, I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)^{-1}\mu\right\rangle \\ &= (\lambda - 2\omega(t))\int_0^\infty \left\langle g, e^{-s\left((\lambda-2\omega(t))I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)}\mu\right\rangle ds, \end{aligned}\tag{8.235}$$

and hence, if $|\lambda - 2\omega(t)| \leqslant 2\omega(t)$ we have

$$\begin{aligned}&|\lambda - 2\omega(t)|\left|\left\langle g, \left((\lambda - 2\omega(t))\, I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)^{-1}\mu\right\rangle\right| \\ &\leqslant |\lambda - 2\omega(t)|\int_0^\infty \left|\left\langle g, e^{-s\left((\lambda-2\omega(t))I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)}\mu\right\rangle\right| ds \\ &\leqslant |\lambda - 2\omega(t)|\int_0^\infty e^{-s(\Re\lambda - 2\omega(t))}\mathrm{Var}\left(e^{sL(t)^*|_{M_0(\mathbb{R}^d)}}\mu\right) ds\, \|g\|_\infty \\ &\leqslant c(t)\,|\lambda - 2\omega(t)|\int_0^\infty e^{-s(\Re\lambda - 2\omega(t))}e^{-2s\omega(t)}ds\mathrm{Var}(\mu)\ \|g\|_\infty \\ &= c(t)\frac{|\lambda - 2\omega(t)|}{\Re\lambda}\|g\|_\infty\ \mathrm{Var}(\mu) \leqslant 2c(t)\,\|g\|_\infty\ \mathrm{Var}(\mu)\,. \end{aligned}\tag{8.236}$$

In view of (8.233), (8.234) and (8.236) it makes sense to consider the largest $\omega(t)$ with the property that for all functions $g \in C_0\left(\mathbb{R}^d\right)$, and all Borel measures $\mu \in M_0\left(\mathbb{R}^d\right)$ the complex-valued function

$$\lambda \mapsto \lambda\left\langle g, \left(\lambda I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)^{-1}\mu\right\rangle$$

extends to a bounded holomorphic function on all half-planes of the form

$$\left\{\lambda \in \mathbb{C} : \Re\lambda > -2\omega'(t)\right\}$$

with $\omega'(t) < \omega(t)$. In follows that there exists a constant $c(t)$ such that for all functions $g \in C_b(E)$ and $\mu \in M_0\left(\mathbb{R}^d\right)$ the following inequality holds:

$$|\lambda|\left|\left\langle g, \left(\lambda I|_{M_0(\mathbb{R}^d)} - L(t)^*|_{M_0(\mathbb{R}^d)}\right)^{-1}\mu\right\rangle\right| \leqslant c(t)\,\|g\|_\infty \mathrm{Var}(\mu)\,, \quad \Re\lambda \geqslant -\omega(t).$$

The following definition is to be compared with the definitions 8.4 and 9.14 in Chapter 9.

Definition 8.5. The number $2\omega(t)$ is called the $M(E)$-spectral gap of the operator $L(t)^*$.

Next let $P(\tau, x; t, B)$ be the transition probability function of the process

$$\left\{\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : t \geqslant \tau\right), \left(\mathbb{R}^d, \mathcal{B}\right)\right\}$$

generated by the operators $L(t)$. Suppose that, for every $\tau \in (0, \infty)$ and every Borel probability measure on $\mathbb{R}^d$, the following condition is satisfied:

$$\lim_{t\to\infty} \frac{c(t)}{\omega(t)} \int_{\mathbb{R}^d} \operatorname{Var}\left(\frac{\partial}{\partial t} P(\tau, x; t, \cdot)\right) d\mu(x) = 0.$$

Let μ be any Borel probability measure on $\mathbb{R}^d$. Put $\mu(t) = Y(\tau, t)^* \mu$, where $Y(\tau, t) f(x) = \mathbb{E}_{\tau,x}\left[f(X(t))\right]$, $f \in C_0(\mathbb{R}^d)$. Then $\dot{\mu}(t) = L(t)^*\mu(t)$. Moreover,

$$\lim_{t\to\infty} \frac{c(t)}{\omega(t)} \operatorname{Var}\left(\dot{\mu}(t)\right) = 0.$$

We will show this. With the above notation we have:

$$\begin{aligned}
&\operatorname{Var}\left(\dot{\mu}(t)\right) \\
&= \sup\left\{\left|\frac{d}{dt}\langle f, \mu(t)\rangle\right| : f \in C_0(\mathbb{R}^d),\ \|f\|_\infty = 1\right\} \\
&= \sup\left\{\left|\frac{\partial}{\partial t}\langle Y(\tau, t) f, \mu\rangle\right| : f \in C_0(\mathbb{R}^d),\ \|f\|_\infty = 1\right\} \\
&= \sup\left\{\left|\frac{\partial}{\partial t}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(y) P(\tau, x; t, dy)\, d\mu(x)\right| : f \in C_0(\mathbb{R}^d),\ \|f\|_\infty = 1\right\} \\
&= \sup\left\{\left|\int_{\mathbb{R}^d} f(y) \frac{\partial}{\partial t}\int_{\mathbb{R}^d} P(\tau, x; t, dy)\, d\mu(x)\right| : f \in C_0(\mathbb{R}^d),\ \|f\|_\infty = 1\right\} \\
&= \operatorname{Var}\left(\frac{\partial}{\partial t}\int_{\mathbb{R}^d} P(\tau, x; t, \cdot)\, d\mu(x)\right) \leqslant \int_{\mathbb{R}^d} \operatorname{Var}\left(\frac{\partial}{\partial t} P(\tau, x; t, \cdot)\right) d\mu(x).
\end{aligned} \tag{8.237}$$

If the probability measure $B \mapsto P(\tau, x; t, B)$ has density $p(\tau, x; t, y)$, then the total variation of the measure $B \mapsto \dfrac{\partial}{\partial t} P(\tau, x; t, \cdot)$ is given by

$$\operatorname{Var}\left(\frac{\partial}{\partial t} P(\tau, x; t, \cdot)\right) = \int_{\mathbb{R}^d} \left|\frac{\partial}{\partial t} p(\tau, x; t, y)\right| dy.$$

If there exists a unique $P(E)$-valued function $t \mapsto \pi(t)$ such that $L(t)^*\pi(t) = 0$, then the system $L(t)^*\mu(t) = \dot{\mu}(t)$ is ergodic. This assertion follows from Theorem 8.5 below.

In order to perform some explicit computations we next assume that $d = 1$. It is assumed that the coefficient $a(t, x)$ is strictly positive on $\mathbb{R}$. Moreover, by hypothesis we assume that there exists a function $B(t, x)$

such that $b(t,x) = a(t,x)\dfrac{\partial}{\partial x}B(t,x)$ and such that $\displaystyle\int_{-\infty}^{\infty} e^{-2B(t,\eta)}d\eta < \infty$. The adjoint $K(t)$ of $L(t)$ acts on a subspace of the dual space of $C_0(\mathbb{R}^d)$ which may be identified with the space of all complex Borel measures on $\mathbb{R}^d$. Formally, $K(t)\mu$ is given by

$$K(t)\mu = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(a(t,\cdot)\mu\right) - \frac{\partial}{\partial x}\left(b(t,\cdot)\mu\right).$$

Let the time-dependent measure $\mu(t)$ have the property that $K(t)\mu(t) = 0$. Then the family of measures $\mu(t)$ has density $\varphi(t,x)$ given by

$$\varphi(t,x) = C_1(t)\frac{e^{2B(t,x)}}{a(t,x)} + C_2(t)\int_a^x \frac{e^{2B(t,x)-2B(t,\eta)}}{a(t,x)}d\eta, \tag{8.238}$$

where $t \mapsto C_j(t)$, $j = 1, 2$, are some functions which only depend on time. In order to be sure that for every t the measure $\mu(t)$ belongs to $M(\mathbb{R})$ and is non-trivial we make additional hypotheses on the coefficients. If both integrals

$$\int_{\infty}^{\infty} \frac{e^{2B(t,x)}}{a(t,x)}dx \quad \text{and} \quad \int_{-\infty}^{\infty}\int_a^x \frac{e^{2B(t,x)-2B(t,\eta)}}{a(t,x)}d\eta\, dx \tag{8.239}$$

are finite, then the function $x \mapsto \varphi(t,x)$ belongs to $L^1(\mathbb{R})$ no matter how the constants $C_1(t)$ and $C_2(t)$ are chosen. The requirement $\int_{-\infty}^{\infty}\varphi(t,x)dx = 1$ does not make them unique. We have uniqueness of solutions in $M(\mathbb{R})$ to the eigenvalue problem $K(t)\mu(t) = 0$ and $\mu(t,\mathbb{R}) = 1$ provided either one of the following conditions is satisfied:

$$\int_{-\infty}^{\infty} \frac{e^{2B(t,x)}}{a(t,x)}dx < \infty \quad \text{and} \quad \int_{-\infty}^{\infty}\int_a^x \frac{e^{2B(t,x)-2B(t,\eta)}}{a(t,x)}d\eta\, dx = \infty, \quad \text{or} \tag{8.240}$$

$$\int_{-\infty}^{\infty} \frac{e^{2B(t,x)}}{a(t,x)}dx = \infty \quad \text{and} \quad \int_{-\infty}^{\infty}\int_a^x \frac{e^{2B(t,x)-2B(t,\eta)}}{a(t,x)}d\eta\, dx < \infty. \tag{8.241}$$

In the cases (8.240) and (8.241) we have respectively

$$\mu(t,B) = C_1(t)\int_B \frac{e^{2B(t,x)}}{a(t,x)}dx \quad \text{and}$$
$$\mu(t,B) = C_2(t)\int_B\int_a^x \frac{e^{2B(t,x)-2B(t,\eta)}}{a(t,x)}d\eta\, dx,$$

where the constants $C_1(t)$ and $C_2(t)$ are chosen in such a way that the total mass $\mu(t,\mathbb{R}) = 1$. The operators $L(t)$ generate a diffusion in the sense that there exists a time-inhomogeneous Markov process

$$\left\{\left(\Omega, \mathcal{F}_t^\tau, \mathbb{P}_{\tau,x}\right), \left(X(t) : t \geqslant \tau\right), \left(\mathbb{R}, \mathcal{B}\right)\right\}$$

such that

$$\frac{\partial}{\partial s}\mathbb{E}_{\tau,x}\left[f\left(X(s)\right)\right] = \mathbb{E}_{\tau,x}\left[L(s)f\left(X(s)\right)\right], \quad f \in D(L(s)),$$

where $0 \leqslant \tau < s \leqslant \infty$. We put $Y(\tau,t)\, f(x) = \mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right]$, $f \in C_b(\mathbb{R})$. Then, under appropriate conditions on the coefficients $a(t,x)$ and $b(t,x)$ the operators $Y(\tau,t)$ leave the space $C_0(\mathbb{R})$ invariant, and hence the adjoint operators $Y(\tau,t)^*$ are mappings from $M(\mathbb{R})$ to $M(\mathbb{R})$. For a given probability measure $\mu(\tau)$ the measure-valued function $\mu(t) := Y(\tau,t)^*\mu(\tau)$ satisfies

$$\begin{aligned}\frac{\partial}{\partial t}\langle f, \mu(t)\rangle &= \frac{\partial}{\partial t}\left\langle f, Y(\tau,t)^*\mu(\tau)\right\rangle = \frac{\partial}{\partial t}\langle Y(\tau,t)\, f, \mu(\tau)\rangle \\ &= \frac{\partial}{\partial t}\int Y(\tau,t)\, f(x)\mu(\tau,dx) = \frac{\partial}{\partial t}\int \mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right]\mu(\tau,dx) \\ &= \int \mathbb{E}_{\tau,x}\left[L(t)f\left(X(t)\right)\right]\mu(\tau,dx) \\ &= \langle Y(\tau,t)\, L(t)f, \mu(\tau)\rangle = \langle f, L(t)^*\mu(t)\rangle .\end{aligned}$$

Let $f \in D(L(t))$. From (8.228) it follows that for all $\lambda \in \mathbb{C}$ with $\Re\lambda \geqslant 0$ the following inequality holds

$$\inf_{\alpha\in\mathbb{C}} |\lambda| \left\|f - \alpha\mathbf{1}\right\|_\infty \leqslant 2 \inf_{\alpha\in\mathbb{C}} \left\|\left(\lambda I - L(t)\right) f - \alpha\mathbf{1}\right\|_\infty . \tag{8.242}$$

If $\lim\limits_{t\to\infty} \dfrac{c(t)}{\omega(t)}\mathrm{Var}\left(L(t)^*\mu(t)\right) = 0$, then the equation $L(t)^*\mu(t) = \dot{\mu}(t)$ is ergodic, provided that $\mathrm{Var}\left(e^{sL(t)^*}\mu\right) \leqslant c(t)e^{-2s\omega(t)}\mathrm{Var}\left(\mu\right)$ for all $\mu \in M_0(E)$. This assertion follows from Theorem 8.5 below, by observing that the dual of the space $C_0(\mathbb{R})$ endowed with the quotient norm $\|f\| := \inf\limits_{\alpha\in\mathbb{C}} \left\|f - \alpha\mathbf{1}\right\|_\infty$ is the space $M_0(\mathbb{R})$.

For explicit formulas for invariant measures for (certain) Ornstein-Uhlenbeck semigroups we refer the reader to [Da Prato and Zabczyk (1992b)] Theorems 11.7 and 11.11, and to [Metafune *et al.* (2002b)]. For some recent regularity and smoothing results see [Bogachev *et al.* (2006)].

8.4 Ergodicity in the non-stationary case

We begin with a relevant definition.

Definition 8.6. The system (8.6) is called ergodic, if there exists a unique solution $\pi(t)$ to the equation $K(t)\pi(t) = 0$, with $\pi(t) \in P(E)$, such that

$$\lim_{t\to\infty} \mathrm{Var}\left(\mu(t) - \pi(t)\right) = 0 \tag{8.243}$$

for all solutions $\mu(t) \in P(E)$ to the equation $\dot{\mu}(t) = K(t)\mu(t)$.

Remark 8.8. Fix $t \in \mathbb{R}$ and let $K(t)$ be a Kolmogorov operator with 0 as an isolated point in its spectrum. Then 0 is a dominant eigenvalue of $K(t)$, and let $P(t) : M(E) \to M(E)$ be the Dunford projection on the generalized eigen-space corresponding to the eigenvalue 0 with eigen-vector $\pi(t) \in P(E)$. If the eigenvalue 0 has multiplicity 1, then $P(t)$ projects the space $M(E)$ onto the one-dimensional subspace $\mathbb{C}\pi(t)$. Since 0 is a dominant eigenvalue of $K(t)$, a key spectral estimate of the following form is valid:

$$\left|\left\langle f, e^{sK(t)}\left(I - P(t)\right)\mu\right\rangle\right| \leqslant c(t)e^{-2\omega(t)s}\left\|f\right\|_{\infty}\operatorname{Var}(\mu), \ f \in C_b(E), \ \mu \in M(E), \tag{8.244}$$

where $\omega(t)$ is strictly positive, $\|f\|_{\infty}$ is the supremum-norm of $f \in C_b(E)$, $\operatorname{Var}(\mu)$ is the total variation norm of $\mu \in M(E)$, and where $c(t)$ is some finite constant.

A $P(E)$-valued function $\pi(t)$ for which $K(t)\pi(t) = 0$ is called a stationary or invariant $P(E)$-valued function of the system in (8.6). In addition to (8.6) we assume that the continuous function $\pi(t)$ with values in $P(E)$ satisfies $K(t)\pi(t) = 0$, and we suppose that this function is uniquely determined.

Theorem 8.5. *Let the function $t \mapsto \mu(t)$ satisfy (8.6); i.e. $\dot{\mu}(t) = K(t)\mu(t)$, $t > t_0$, or more precisely $\frac{d}{dt}\langle f, \mu(t)\rangle = \langle f, K(t)\mu(t)\rangle$, $f \in C_b(E)$. In addition, suppose that there exist strictly positive functions $t \mapsto \omega(t)$ and $t \mapsto c(t)$ possessing the following properties:*

(i) For every $t \geqslant t_0$ there exists a real number with $\Re\lambda > -\omega(t)$ such that

$$\left(\lambda I - K(t)\right)\left(D\left(K(t)\right)\right) = M(E); \tag{8.245}$$

(ii) The following identity holds true:

$$\lim_{t\to\infty}\frac{c(t)}{\omega(t)}\operatorname{Var}\left(\dot{\mu}(t)\right) = \lim_{t\to\infty}\frac{c(t)}{\omega(t)}\operatorname{Var}\left(K(t)\mu(t)\right) = 0; \tag{8.246}$$

(iii) The inequality

$$|\lambda|\operatorname{Var}(\mu) \leqslant c(t)\operatorname{Var}\left(\lambda\mu - K(t)\mu\right), \tag{8.247}$$

holds for all $\mu \in D\left(K(t)\right)$ and all $\lambda \in \mathbb{C}$ with $\Re\lambda > -\omega(t)$.

Then there exists a $P(E)$-valued function $t \mapsto \pi(t)$ such that

$$\lim_{t\to\infty}\operatorname{Var}\left(\mu(t) - \pi(t)\right) = 0,$$

and such that $K(t)\pi(t) = 0$; i.e. the system in (8.6) is ergodic.

Remark 8.9. The inequality in (8.247) is only required on the union of the right half-plane $\{\lambda \in \mathbb{C} : \Re\lambda > 0\}$ and the circular disc $\{\lambda \in \mathbb{C} : |\lambda| \leqslant \omega(t)\}$. This will follow from the proof of Theorem 8.5.

Remark 8.10. Let $g \in C_b(E)$ be such that

$$\left|\left\langle g, \left(-K(t)|_{M_0(E)}\right)^{-1} \mu\right\rangle\right| \leqslant \frac{c(t)}{\omega(t)} \|g\|_\infty \operatorname{Var}(\mu), \quad \mu \in M_0(E), \tag{8.248}$$

where the constants $c(t)$ and $\omega(t)$ satisfy (8.246). Then

$$\lim_{t\to\infty} \langle g, \mu(t) - \pi(t)\rangle = 0. \tag{8.249}$$

If the collection of functions g satisfying (8.248) for an appropriate choice of $c(t)$ and $\omega(t)$ satisfying (8.246) is dense in $C_b(E)$, then (8.249) holds for $g \in C_b(E)$.

The following proposition has some independent interest; it says that an operator which has the properties (i) and (iii) of Theorem 8.5 generates a bounded analytic weak*-continuous semigroup in $M_0(E)$ with exponential decay. For $\omega > 0$ we define the open subset $\widetilde{\Pi}_\omega$ of $\mathbb{C}$ by $\widetilde{\Pi}_\omega = \{\lambda \in \mathbb{C} : \Re\lambda > 0\} \bigcup \{\lambda \in \mathbb{C} : |\lambda| < \omega\}$.

Proposition 8.6. *Let K be a sectorial sub-Kolmogorov operator for which there exist constants ω and c such that $(\lambda I - K)\, D(K) = M(E)$ for some $\lambda \in \widetilde{\Pi}_\omega$ and such that*

$$|\lambda| \operatorname{Var}(\mu) \leqslant c\operatorname{Var}(\lambda\mu - K\mu) \tag{8.250}$$

for all $\lambda \in \widetilde{\Pi}_\omega$ and for all $\mu \in D(K)$. Then the operator K generates a weak-continuous bounded analytic semigroup $\left\{e^{tK} : |\arg(t)| \leqslant \alpha\right\}$. On the range of the operator K this analytic semigroup has exponential decay as $t \to \infty$.*

Proof. [Proof of Proposition 8.6.] We consider the subset Π_ω of $\widetilde{\Pi}_\omega$ defined by

$$\Pi_\omega = \left\{\lambda \in \widetilde{\Pi}_\omega : \lambda \neq 0,\ (\lambda I - K)\, D(K) = M(E)\right\}. \tag{8.251}$$

First suppose that λ_0 belongs to Π_ω. Put $R(\lambda_0) = (\lambda_0 I - K)^{-1}$, and define the operators $R(\lambda)$, $|\lambda - \lambda_0| < c^{-1}|\lambda_0|$, by $R(\lambda) = \sum_{j=0}^\infty (\lambda_0 - \lambda)^j R(\lambda_0)^{j+1}$. From (8.250) it follows that the operators $R(\lambda)$ are well defined and that $(\lambda I - K)\, R(\lambda) = I$ for $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < c^{-1}|\lambda_0|$. Hence the set Π_ω is an open subset of the punctured subset $\widetilde{\Pi}_\omega\backslash\{0\}$. Next let λ_n, $n \in \mathbb{N}$, be a sequence in Π_ω with limit λ_0 in the punctured open subset $\widetilde{\Pi}_\omega \backslash \{0\}$. For $n \in \mathbb{N}$ so large that $|\lambda_0 - \lambda_n| < c^{-1} |\lambda_n|$ we have

$$(\lambda_0 I - K) \sum_{j=0}^{\infty} (\lambda_n - \lambda_0)^j R(\lambda_n)^{j+1} = I,$$

where we wrote $R(\lambda_n) = (\lambda_n I - K)^{-1}$. It follows that the punctured set $\Pi_\omega \backslash \{0\}$ is open and closed in the connected punctured open set $\widetilde{\Pi}_\omega \backslash \{0\}$. Since the latter is topologically connected and since by assumption Π_ω is non-empty it follows that for every $\lambda \in \widetilde{\Pi}_\omega$, $\lambda \neq 0$, the range of the operator $\lambda I - K$ coincides with $M(E)$. As above we put $R(\lambda) = (\lambda I - K)^{-1}$, $\lambda \in \widetilde{\Pi}_\omega$. Inequality (8.250) implies that $\left\| (\lambda I - K)^{-1} \right\| \leqslant c$, $\lambda \in \widetilde{\Pi}_\omega$. From the arguments in the proofs of Theorem 8.7 and Corollary 8.3 it follows that the resolvent $R(\lambda)$ extends to a sectorial region of the form $\Pi_{\omega,\beta} := \widetilde{\Pi}_\omega \bigcup \{\lambda \in \mathbb{C} : |\arg(\lambda)| \leqslant \beta\}$, where $\frac{1}{2}\pi < \beta < \pi$, and the norm of the of the resolvent $R(\lambda)$ satisfies an estimate of the form:

$$|\lambda| \, \|R(\lambda)\| \leqslant c', \quad \lambda \in \Pi_{\omega,\beta}. \tag{8.252}$$

Put

$$P = \frac{1}{2\pi i} \int_{|\lambda|=\omega} (\lambda I - K)^{-1} \, d\lambda \quad \text{and} \quad A = -\frac{1}{2\pi i} \int_{|\lambda|=\omega} \frac{1}{\lambda} (\lambda I - K)^{-1} \, d\lambda. \tag{8.253}$$

Then we have

$$\begin{aligned} KP &= \frac{1}{2\pi i} \int_{|\lambda|=\omega} (\lambda I - (\lambda I - K)) (\lambda I - K)^{-1} \, d\lambda \\ &= \frac{1}{2\pi i} \int_{|\lambda|=\omega} \lambda(\lambda I - K)^{-1} d\lambda - \frac{1}{2\pi i} \int_{|\lambda|=\omega} (\lambda I - K)(\lambda I - K)^{-1} d\lambda = 0, \end{aligned}$$

and

$$\begin{aligned} KA &= \frac{1}{2\pi i} \int_{|\lambda|=\omega} \frac{1}{\lambda} (\lambda I - K - \lambda I) (\lambda I - K)^{-1} \, d\lambda \\ &= \frac{1}{2\pi i} \int_{|\lambda|=\omega} \frac{1}{\lambda} d\lambda I - \frac{1}{2\pi i} \int_{|\lambda|=\omega} (\lambda I - K)^{-1} \, d\lambda = I - P. \end{aligned} \tag{8.254}$$

It follows that $R(K)$, the range of K, is weak*-closed and that $I - P$ is a continuous linear projection from $M(E)$ onto $R(K)$ with null space $R(P) = N(K)$. From Theorem 8.3 it follows that K generates a weak*-continuous sub-Kolmogorov semigroup $\left\{e^{tK} : t \geqslant 0\right\}$ in $M(E)$. By (8.252) we see that this semigroup is analytic. Since the set $\Pi_{\omega,\beta}$ contains a half-plane of the form $\{\lambda \in \mathbb{C} : \Re\lambda \geqslant -\omega_0\}$ where $\omega > \omega_0 > 0$ the representation in (8.258) with $\omega' = -\omega_0$ and $\ell = 1$ can be used to show the exponential decay of the semigroup $\left\{e^{tK} : t \geqslant 0\right\}$ on the range of K.

This completes the proof of Proposition 8.6. □

Suppose that $|\lambda| \operatorname{Var}(\mu) \leqslant c\operatorname{Var}(\lambda\mu - K\mu)$ for $\Re\lambda \geqslant -\omega$, $\mu \in D(K)$. Then the operator $I - P$ can be written as

$$I - P = K\frac{1}{2\pi i}\int_{-\omega-i\infty}^{-\omega+i\infty}\frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda = -K\int_0^\infty e^{sK} ds. \tag{8.255}$$

On the range of K (which coincides with the range of $I - P$) the operator A has the representation:

$$A = \frac{1}{2\pi i}\int_{-\omega-i\infty}^{-\omega+i\infty}\frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda = -\int_0^\infty e^{sK}\left(I - P\right) ds. \tag{8.256}$$

On the space $M(E)$ the operator e^{tK} can be represented by

$$\frac{t^\ell}{\ell!}e^{tK} = \frac{1}{2\pi i}\int_{\omega'-i\infty}^{\omega'+i\infty} e^{t\lambda}\left(\lambda I - K\right)^{-\ell-1} d\lambda, \quad \omega' > 0, \ \ell \geqslant 1. \tag{8.257}$$

On the range of K the operator e^{tK} has the representation:

$$\frac{t^\ell}{\ell!}e^{tK} = \frac{1}{2\pi i}\int_{\omega'-i\infty}^{\omega'+i\infty} e^{t\lambda}\left(\lambda I - K\right)^{-\ell-1} d\lambda, \quad \omega' > -\omega, \ \ell \geqslant 1. \tag{8.258}$$

Notice that by (8.254) the operator K has a bounded inverse on its range. It follows that the function $\lambda \mapsto (\lambda I - K)^{-1}$ restricted to $R(K)$ is holomorphic in a neighborhood of $\lambda = 0$.

Remark 8.11. We may say that the condition $\sup_{t>0}\left\|tLe^{tL}\right\| < \infty$ is kind of an analytic maximum principle.analytic maximum principle. In this remark only, suppose that E is locally compact and second countable. Let L be the generator of a Feller-Dynkin semigroup. Fix $t > 0$ and choose $x \in E$ in such a way that $\left|e^{tL}f(x)\right| = \left\|e^{tL}f\right\|_\infty$. Then we have $\Re\left(\overline{e^{tL}f(x)}tLe^{tL}f(x)\right) \leqslant 0$. Next assume that the operator L is such that the corresponding Feller-Dynkin semigroup has an integral $p(t,x,y)$ with respect to a reference measure $dm(y)$. This means that the semigroup e^{tL} is given by $e^{tL}f(x) = \int p\left(t,x,y\right)f(y)dm(y)$. Then L generates a bounded analytic semigroup if and only if

$$\sup_{t>0}\sup_{x\in E}\int_E\left|\frac{t\partial p\left(t,x,y\right)}{\partial t}\right| dm(y) = \sup_{t>0}\sup_{x\in E}\int_E\left|tL\,p\left(t,\cdot,y\right)(x)\right| dm(y) < \infty.$$

This is the case if and only if for some $\alpha \in \left(0, \frac{1}{2}\pi\right)$ an inequality of the form

$$\sup_{t\in\mathbb{C}:|\arg(t)|\leqslant\alpha}\ \sup_{x\in E}\int\left|p\left(t,x,y\right)\right| dm(y) < \infty$$

holds.

For the moment we only suppose that the operator K generates a bounded analytic weak*-continuous semigroup on $M(E)$. Let $\gamma_r : \left[-\frac{1}{2}\pi, \frac{1}{2}\pi\right]$, $0 < r < \infty$, be a parametrization of the semi-circle $\gamma_r(\vartheta) = re^{i\vartheta}$, $-\frac{1}{2}\pi \leqslant \vartheta \leqslant \frac{1}{2}\pi$. Then by Cauchy's theorem the following equality of sums of integrals holds for $0 < r < R < \infty$:

$$\begin{aligned} &\frac{1}{\pi i}\int_{-iR}^{-ir} \frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda + \frac{1}{\pi i}\int_{\gamma_r} \frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda \\ &\quad + \frac{1}{\pi i}\int_{ir}^{iR} \frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda \\ &= \frac{1}{\pi i}\int_{\gamma_R} \frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda. \end{aligned} \tag{8.259}$$

By using the parameterizations $\xi \mapsto -i\xi$, $R > \xi > r$, and $\xi \mapsto -i\xi$, $r < \xi < R$ and letting R tend to ∞ we obtain:

$$\frac{2}{\pi}\int_r^{\infty} \left(\xi^2 I + K^2\right)^{-1} d\xi = \frac{1}{\pi i}\int_{\gamma_r} \frac{1}{\lambda}\left(\lambda I - K\right)^{-1} d\lambda. \tag{8.260}$$

It follows that

$$\begin{aligned} &\frac{2}{\pi}(-K)\int_r^{\infty} \left(\xi^2 I + K^2\right)^{-1} d\xi = \frac{2}{\pi}\int_r^{\infty} (-K)\left(\xi^2 I + K^2\right)^{-1} d\xi \\ &= \frac{1}{\pi i}\int_{\gamma_r} \frac{1}{\lambda}\left(\lambda I - K - \lambda I\right)\left(\lambda I - K\right)^{-1} d\lambda \\ &= \frac{1}{\pi i}\int_{\gamma_r} \frac{1}{\lambda} d\lambda I - \frac{1}{\pi i}\int_{\gamma_r} \left(\lambda I - K\right)^{-1} d\lambda \\ &= I - \frac{1}{\pi i}\int_{\gamma_r} \left(\lambda I - K\right)^{-1} d\lambda. \end{aligned} \tag{8.261}$$

From (8.261) we also obtain:

$$\begin{aligned} &K\left(\frac{2}{\pi}\int_r^{\infty} (-K)\left(\xi^2 I + K^2\right)^{-1} d\xi - I\right) \\ &= \frac{1}{\pi i}\int_{\gamma_r} 1 d\lambda\, I - \frac{1}{\pi i}\int_{\gamma_r} \lambda\left(\lambda I - K\right)^{-1} d\lambda \\ &= \frac{2r}{\pi} I - \frac{1}{\pi i}\int_{\gamma_r} \lambda\left(\lambda I - K\right)^{-1} d\lambda. \end{aligned} \tag{8.262}$$

We formulate these results in the form of a proposition.

Proposition 8.7. *Put*

$$Q_r = \frac{2}{\pi}(-K)\int_r^{\infty} \left(\xi^2 I + K^2\right)^{-1} d\xi \quad \textit{and} \quad P_r = \frac{1}{\pi i}\int_{\gamma_r} \left(\lambda I - K\right)^{-1} d\lambda. \tag{8.263}$$

Then $I = Q_r + P_r$, $R(P_r) \subset D(K)$, *and*

$$K\left(I - \frac{2}{\pi}\int_r^\infty (-K)\left(\xi^2 I + K^2\right)^{-1} d\xi\right)$$
$$= KP_r = \frac{1}{\pi i}\int_{\gamma_r} \lambda\left(\lambda I - K\right)^{-1} d\lambda - \frac{2r}{\pi} I. \tag{8.264}$$

Moreover, the following inequality is valid:

$$\left\| KP_r \right\| \leqslant r \sup_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]} \left\| re^{i\vartheta}\left(re^{i\vartheta}I - K\right)^{-1}\right\| + \frac{2r}{\pi}. \tag{8.265}$$

Definition 8.7. A linear operator $Q : M(E) \to M(E)$ is called sequentially weak*-closed if its graph $\{(\mu, Q\mu) : \mu \in M(E)\}$ is sequentially weak* closed in $M(E) \times M(E)$. This means that for any sequence $(\mu_n)_{n\in\mathbb{N}}$ which itself converges to μ for the $\sigma(M(E), C_b(E))$-topology, and for which the sequence $(Q\mu_n)_{n\in\mathbb{N}}$ converges to $\nu \in M(E)$ with respect to the $\sigma(M(E), C_b(E))$-topology the equality $\nu = Q\mu$ follows.

In the following proposition we collect a number of alternative ways to represent the operators Q and P. Recall that the projection operator P is called a Dunford projection.

Proposition 8.8. *Let K be a sub-Kolmogorov operator which generates a weak*-continuous semigroup $\left\{e^{sK} : s \geqslant 0\right\}$ in $M(E)$. Put $R(\lambda) = (\lambda I - K)^{-1}$, $\Re\lambda > 0$. The following assertions are true:*

(1) Suppose that the weak-limit*

$$Q\mu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{\lambda\downarrow 0}\left(-K\right)R(\lambda)\mu$$

exists for all $\mu \in M(E)$. In addition, suppose that the operator Q is sequentially weak-closed. Then Q is a projection from $M(E)$ onto the weak*-sequential closure of the space $R(K)$. Its zero space is $N(K)$, and the projection $P = I - Q$ on $N(K)$ is given by*

$$P\mu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{\lambda\downarrow 0}\lambda R(\lambda)\mu.$$

(2) Suppose that the weak-limit*

$$Q\mu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{t\uparrow\infty}\left(-K\right)\int_0^t e^{sK}\mu ds$$

exists for all $\mu \in M(E)$. In addition, suppose that the operator Q is sequentially weak-closed. Then Q is a projection from $M(E)$ onto the* *weak*-sequential closure of the space $R(K)$. Its zero space is $N(K)$, and the projection $P = I - Q$ on $N(K)$ is given by*

$$P\mu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{t\to\infty} e^{tK}\mu,$$

provided that $\sigma\left(M(E), C_b(E)\right)\text{-}\lim_{t\to\infty} Ke^{tK}\mu = 0$ for all $\mu \in D(K)$.

(3) *Suppose that the semigroup generated by K is bounded and analytic. In addition, assume that the weak*-limit*

$$Q\mu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{r\downarrow 0} \frac{2}{\pi}(-K)\int_r^\infty \left(\xi^2 I + K^2\right)^{-1}\mu d\xi$$

exists for all $\mu \in M(E)$, and suppose that the operator Q is sequentially weak-closed. Then Q is a projection from $M(E)$ onto the weak*-sequential closure of the space $R(K)$. Its zero space is $N(K)$, and the projection $P = I - Q$ on $N(K)$ is given by*

$$P\mu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{r\downarrow 0} \frac{1}{\pi i}\int_{\gamma_r} (\lambda I - K)^{-1}\mu\, d\lambda, \quad \mu \in M(E).$$

Here γ_r is the curve $\gamma_r(\vartheta) = re^{i\vartheta}$, $-\frac{1}{2}\pi \leqslant \vartheta \leqslant \frac{1}{2}\pi$.

(4) *Suppose that 0 is an isolated point of the spectrum of K and that in a neighborhood of 0 the following inequality holds for a finite constant C, for all $\mu \in D(K)$ and for all $\lambda \in \mathbb{C}$ in a (small) disc around 0:*

$$|\lambda|\,\mathrm{Var}(\mu) \leqslant C\mathrm{Var}\left(\lambda\mu - K\mu\right). \tag{8.266}$$

Then the range of K is weak-closed, and $M(E) = R(K) + N(K)$. More precisely, put*

$$Q\mu = \frac{1}{2\pi i}(-K)\int_{\widetilde{\gamma}_r} \frac{1}{\lambda}(\lambda I - K)^{-1}\mu\, d\lambda, \quad \textit{and}$$
$$P\mu = \frac{1}{2\pi i}\int_{\widetilde{\gamma}_r} (\lambda I - K)^{-1}\mu\, d\lambda$$

where $\mu \in M(E)$. Here $\widetilde{\gamma}_r$ stands for the full circle: $\widetilde{\gamma}_r(\vartheta) = re^{i\vartheta}$, $-\pi \leqslant \vartheta \leqslant \pi$, and for $|\lambda| \leqslant r$ the inequality in (8.266) holds. Then Q is a weak-continuous projection mapping from $M(E)$ onto $R(K)$, and $P = I - Q$ is weak*-continuous projection mapping from $M(E)$ onto $N(K)$. Moreover, $I = Q + P$.*

Remark 8.12. If the operator K is the weak*-generator of a bounded analytic semigroup $\left\{e^{tK} : t \geqslant 0\right\}$. Then the families $\left\{e^{tK} : t \geqslant 0\right\}$ and $\left\{tKe^{tK} : t \geqslant 0\right\}$ are uniformly bounded. It follows that $\lim_{t\to\infty} \mathrm{Var}\left(Ke^{tK}\mu\right) =$

0, and hence the assumptions of assertion (3) entail those of (2). The identity

$$\lambda R(\lambda)\,\mu = \lambda \int_0^\infty e^{-\lambda t} e^{tK} \mu\, dt = \int_0^\infty e^{-t} e^{\lambda^{-1} tK} \mu\, dt$$

shows that assertion (1) is a consequence of (2). Finally, by residue-calculus and the hypothesis in assertion (4) we also have

$$\frac{1}{2\pi i} \int_{\widetilde{\gamma}_r} (\lambda I - K)^{-1} \mu\, d\lambda = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{\lambda \downarrow 0} \lambda R(\lambda)\mu.$$

It follows that the conditions in assertion (4) imply those of (1).

Remark 8.13. Let $(\mu, \nu) \in M(E) \times M(E)$ be such that there exists a sequence $(\mu_n)_{n\in\mathbb{N}} \subset M(E)$ together with a sequence $(\lambda_n)_{n\in\mathbb{N}} \subset (0, \infty)$ which decreases to 0 if n tends to ∞ such that

$$(\mu, \nu) = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} \left(\mu_n, \mu_n - \lambda R(\lambda)\mu_n\right).$$

Then it is assumed that the graph of the operator Q contains the pair (μ, ν). Let the sequence $(\mu_n, \mu_n - \lambda_n R(\lambda_n)\,\mu_n)$ tend to (μ, ν) for the weak*-topology. First we show that $Q\nu = Q\mu$. By assumption we know that

$$\begin{aligned} &\sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} \left(\mu - \lambda_n R(\lambda_n)\,\mu_n\right) \\ &= \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} \left(-K R(\lambda_n)\,\mu_n\right) = Q\mu. \end{aligned} \tag{8.267}$$

We also have:

$$\sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} \left(\mu_n - \mu - \lambda_n R(\lambda_n)\,(\mu_n - \mu)\right) = \nu - Q\mu. \tag{8.268}$$

Since μ_n converges to μ in the weak* sense the equality in (8.268) implies:

$$\sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} \left(-\lambda_n R(\lambda_n)\,(\mu_n - \mu)\right) = \nu - Q\mu. \tag{8.269}$$

In addition, we have

$$\lim_{n\to\infty} \operatorname{Var}\left(K\lambda_n R(\lambda_n)\,(\mu_n - \mu)\right) = \lim_{n\to\infty} \operatorname{Var}\left(\left(\lambda_n^2 R(\lambda_n) - \lambda_n\right)(\mu_n - \mu)\right) = 0 \tag{8.270}$$

and hence, since the operator K is sequentially weak* closed we infer

$$K\left(\nu - Q\mu\right) = 0.$$

But we also have $N(K) = N(Q)$ and thus $Q\left(\nu - Q\mu\right) = 0$. Since $Q^2 = Q$ we see $Q\nu = Q\mu$. Fix $N \in \mathbb{N}$. Using the equalities $\lambda_n R(\lambda_n)\,(\nu - Q\nu) = \nu - Q\nu$ and $Q\nu = Q\mu$ we obtain the identities:

$$\frac{1}{N+1}\left(\mu_n - \mu - \left(\lambda_n R(\lambda_n)\right)^{N+1}(\mu_n - \mu)\right) - (\nu - Q\nu)$$

$$= \frac{1}{N+1} \sum_{j=0}^{N} \left(\lambda_n R\left(\lambda_n\right)\right)^j \left\{\left(I - \lambda_n R\left(\lambda_n\right)\right)\left(\mu_n - \mu\right) - \left(\nu - Q\nu\right)\right\}$$

$$= \frac{1}{N+1} \sum_{j=0}^{N} \left(\lambda_n R\left(\lambda_n\right)\right)^j \left\{\left(I - \lambda_n R\left(\lambda_n\right)\right)\left(\mu_n - \mu\right) - \left(\nu - Q\mu\right)\right\}.$$

Hence, if we assume from the start that

$$\sigma\left(M(E), C_b(E)\right) \text{-} \lim_{n\to\infty} \left(\lambda_n R\left(\lambda_n\right)\left(\mu_n\right)\right) = 0$$

whenever $\lambda_n \downarrow 0$ and $\sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} \mu_n = 0$, then $R(Q)$ is the weak*-closure of $R(K)$.

Proof. [Proof of Proposition 8.8.] Proof of assertion (1). Let $\mu \in M(E)$. First we notice the equalities $\mu + KR(\lambda)\mu = \lambda R(\lambda)\mu \in D(K)$, and $\lim_{\lambda\downarrow 0} K\left(\lambda R(\lambda)\mu\right) = \lim_{\lambda\downarrow 0}\left(\lambda^2 R(\lambda)\mu - \lambda\mu\right) = 0$. The latter limit is taken with respect to the variation norm. In addition, we see that $P\mu := \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{\lambda\downarrow 0} \lambda R(\lambda)\mu$ exists. Since the graph of K is sequentially weak*-closed, it follows that $P\mu$ belongs to $D(K)$ and $KP\mu$. Hence, we see that the measure $\mu - Q\mu$ belongs to $N(K)$ then. Consequently, if $Q\mu = 0$, then $\mu = \mu - Q\mu \in N(K)$. If $K\mu = 0$, then

$$Q\mu = \lim_{\lambda\downarrow 0} (-K)\left(\lambda R\left(\lambda\right)\mu\right) = -\lim_{\lambda\downarrow 0}\left(\lambda R\left(\lambda\right)K\mu\right) = 0.$$

The previous arguments show the equalities of spaces: $(I - Q)\, M(E) = N(K) = N(Q)$. It follows that $Q\,(I - Q) = 0$, and thus $Q = Q^2$. From the definition of Q it follows that $R(Q)$, the range of Q, is contained in the sequential weak*-closure of $R(K)$. Conversely, let $\nu = \sigma\left(M(E), C_b(E)\right)\text{-}\lim_{n\to\infty} K\mu_n$, where $\left(\mu_n\right)_{n\in\mathbb{N}}$ is a sequence in $D(K)$. Then $Q\left(K\mu_n - \nu\right) = Q\left(K\mu_n\right) - \nu + \nu - Q\nu = K\mu_n - \nu + \nu - Q\nu$, which converges for the weak*-topology to $\nu - Q\nu$. It follows that the pair $(0, \nu - Q\nu)$ belongs to the sequential weak*-closure of the graph of Q, and consequently $\nu = Q\nu$.

Proof of assertion (2). In the proof of this assertion we use the identity $\mu + K\int_0^t e^{sK}\mu ds = e^{tK}\mu$ instead of $\mu + KR\left(\lambda\right)\mu = \lambda R\left(\lambda\right)\mu$. Then we let t tend to ∞.

Proof of assertion (3). In the proof of this assertion we employ the identity

$$\mu + K\int_r^\infty \left(\xi^2 I + K^2\right)^{-1}\mu\, d\xi = \frac{1}{\pi i}\int_{\gamma_r} \left(\lambda I - K\right)^{-1}\mu\, d\lambda.$$

Then we let $r > 0$ tend to 0.

Proof of assertion (4). Here we have the identity:

$$\mu+\frac{1}{2\pi i}K\int_{\widetilde{\gamma}_r}\frac{1}{\lambda}\left(\lambda I-K\right)^{-1}\mu\,d\lambda=\frac{1}{2\pi i}\int_{\widetilde{\gamma}_r}\left(\lambda I-K\right)^{-1}\mu\,d\lambda.$$

Hence, here we have

$$Q\mu=\frac{1}{2\pi i}\left(-K\right)\int_{\widetilde{\gamma}_r}\frac{1}{\lambda}\left(\lambda I-K\right)^{-1}\mu\,d\lambda,\text{ and }P\mu=\frac{1}{2\pi i}\int_{\widetilde{\gamma}_r}\left(\lambda I-K\right)^{-1}\mu\,d\lambda.$$

Essentially speaking this proves assertion (4).

This completes the proof of Proposition 8.8. □

In all these cases we prove that $(I-Q)\,M(E)=N(K)=N(Q)$, and $Q\,(K\mu)=K\,(Q\mu)=K\mu$ for $\mu\in D(K)$. Consequently, $Q^2=Q$. If $Q\mu=0$, then $\mu=\mu-Q\mu\in N(K)$, and hence $N(Q)\subset N(K)$. Conversely, if $\mu\in D(K)$ is such that $K\mu=0$, then the definition of Q implies $Q\mu=0$.

Theorem 8.6. *Suppose that the operator K generates a bounded analytic weak*-continuous semigroup on $M(E)$, and that for every $f\in C_b(E)$ and $\mu\in M(E)$ the integral*

$$\frac{2}{\pi}\int_0^\infty\left\langle f,(-K)\left(\xi^2I+K^2\right)^{-1}\mu\right\rangle d\xi \tag{8.271}$$

exists as an improper Riemann integral. Suppose that for every $\mu\in M(E)$, the family of measures $\left\{\lambda\left(\lambda I-K\right)^{-1}\mu:\,\Re\lambda>0\right\}$ is $\mathcal{T}_\beta$-equi-continuous. Then for every $\mu\in M(E)$, the functional

$$f\mapsto\int_0^\infty\left\langle f,(-K)\left(\xi^2I+K^2\right)^{-1}\mu\right\rangle d\xi,\quad f\in C_b(E) \tag{8.272}$$

is continuous on $(C_b(E),\mathcal{T}_\beta)$, and hence it can be identified with a measure. In addition, it is assumed that for every $f\in C_b(E)$ the equality

$$\lim_{n\to\infty}\int_0^\infty\left\langle f,(-K)\left(\xi^2I+K^2\right)^{-1}\mu_n\right\rangle d\xi=\int_0^\infty\left\langle f,(-K)\left(\xi^2I+K^2\right)^{-1}\mu\right\rangle d\xi$$

holds whenever $(\mu_n:n\in\mathbb{N})$ is a sequence in $M(E)$ which converges with respect to the $\sigma\left(M(E),C_b(E)\right)$-topology to a measure $\mu\in M(E)$, i.e.

$$\lim_{n\to\infty}\langle g,\mu_n\rangle=\langle g,\mu\rangle\quad\text{for all } g\in C_b(E).$$

For $\mu\in M(E)$ let $Q\mu$ denote the measure corresponding to the functional:

$$f\mapsto\frac{2}{\pi}\int_0^\infty\left\langle f,(-K)\left(\xi^2I+K^2\right)^{-1}\mu\right\rangle d\xi=\langle f,Q\mu\rangle\,.$$

Then for every $\mu\in M(E)$ the measure $\mu-Q\mu$ belongs to $D(K)$ and $K\left(\mu-Q\mu\right)=0$. Moreover, $R(Q)$ is the weak sequential closure of $R(K)$, and $Q^2=Q$. In addition $I-Q$ sends positive measures to positive measures, and $R\left(Q\right)\bigcap N(Q)=\{0\}$. If $\langle\mathbf{1},K\mu\rangle=0$ for all $\mu\in D(K)$, then $I-Q$ sends the convex set of probability measures on $\mathcal{E}$ to itself.*

Proof. [Proof of Theorem 8.6.] As in Proposition 8.7 we introduce the operators

$$Q_r = \frac{2}{\pi}(-K)\int_r^\infty \left(\xi^2 I + K^2\right)^{-1} d\xi \quad \text{and} \quad P_r = \frac{1}{\pi i}\int_{\gamma_r} (\lambda I - K)^{-1}\, d\lambda. \tag{8.273}$$

Then $I = Q_r + P_r$. Notice that, for given $\mu \in M(E)$, the collection $\left\{\lambda\,(\lambda I - K)^{-1}\mu : \Re\lambda > 0\right\}$ is $\mathcal{T}_\beta$-equi-continuous. As a consequence we see that the functional in (8.272) belongs to $M(E)$. The proof of Theorem 8.6 can be completed as the proof of Proposition 8.8. □

Proof. [Proof of Theorem 8.5.] Let $\mu(t)$ be as in (8.6), and let $\pi(t)$ satisfy $K(t)\pi(t) = 0$. It follows that $\dot{\mu}(t) = K(t)\mu(t)$ belongs to $M_0(E)$. Since the spectrum of the operator $K(t)|_{M_0(E)}$ is contained in the complement of a circle sector of the form

$$\{\lambda \in \mathbb{C} : \Re\lambda \geqslant -\omega(t) : |\arg(\lambda + \omega(t))| \leqslant \beta\}$$

with $\frac{1}{2}\pi < \beta < \pi$, we have:

$$\begin{aligned}
&(I - P(t))\,\mu(t)\\
&= \frac{1}{2\pi i}K(t)\int_{-\omega(t)-i\infty}^{-\omega(t)+i\infty} \frac{1}{\lambda}\,(\lambda I - K(t))^{-1}\, d\lambda\mu(t)\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{-\omega(t)+i\xi}\left((-\omega(t)+i\xi)\, I|_{M_0(E)} - K(t)|_{M_0(E)}\right)^{-1}\\
&\qquad \left(K(t)|_{M_0(E)}\right)(\mu(t)-\pi(t))\, d\xi \qquad (8.274)\\
&= \left(K(t)|_{M_0(E)}\right)^{-1}\left(K(t)|_{M_0(E)}\right)(\mu(t)-\pi(t))\\
&= \mu(t) - \pi(t). \qquad (8.275)
\end{aligned}$$

From (8.275) we see that $P(t)\mu(t) = \pi(t)$ and hence $K(t)P(t)\mu(t) = 0$. Using (8.247) and (8.274) as a norm estimate we obtain the following one:

$$\begin{aligned}
&\operatorname{Var}\left((I - P(t))\,\mu(t)\right)\\
&\leqslant \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{|-\omega(t)+i\xi|}\operatorname{Var}\Big(\left((-\omega(t)+i\xi)\, I|_{M_0(E)} - K(t)|_{M_0(E)}\right)^{-1}\\
&\qquad \left(K(t)|_{M_0(E)}\right)(\mu(t)-\pi(t))\Big)\, d\xi\\
&\leqslant \frac{c(t)}{2\pi}\int_{-\infty}^{\infty} \frac{1}{|-\omega(t)+i\xi|^2}\, d\xi \operatorname{Var}\left(\left(K(t)|_{M_0(E)}\right)(\mu(t)-\pi(t))\right)\\
&= \frac{c(t)}{2\omega(t)}\operatorname{Var}\left(K(t)\mu(t)\right) = \frac{c(t)}{2\omega(t)}\operatorname{Var}\left(\dot{\mu}(t)\right). \qquad (8.276)
\end{aligned}$$

The estimate in (8.276) and (8.246) entails the following result

$$\lim_{t\to\infty} \operatorname{Var}\left(\left(I-P(t)\right)\mu(t)\right) = \lim_{t\to\infty} \operatorname{Var}\left(\mu(t)-\pi(t)\right) = 0.$$

This essentially proves Theorem 8.5. □

Remark 8.14. In this remark we give an alternative representation of the operator $P(t)$. Since the measure $K(t)\mu(t)$ belongs to $M_0(E)$, from (8.247) it follows that

$$\frac{1}{2\pi i}K(t)\int_{\omega(t)-i\infty}^{\omega(t)+i\infty}\left(\lambda I-K(t)\right)^{-1}\frac{1}{\lambda}d\lambda\,\mu(t)=0, \tag{8.277}$$

and hence, by Cauchy's theorem,

$$\begin{aligned}
&\left(I-P(t)\right)\mu(t)\\
&=\frac{1}{2\pi i}K(t)\int_{-\omega(t)-i\infty}^{-\omega(t)+i\infty}\left(\lambda I-K(t)\right)^{-1}\frac{1}{\lambda}d\lambda\,\mu(t)\\
&\qquad-\frac{1}{2\pi i}K(t)\int_{\omega(t)-i\infty}^{\omega(t)+i\infty}\left(\lambda I-K(t)\right)^{-1}\frac{1}{\lambda}d\lambda\,\mu(t)\\
&=-\frac{1}{2\pi i}\int_{\{|\lambda|=\omega(t)\}}K(t)\left(\lambda I-K(t)\right)^{-1}\frac{1}{\lambda}d\lambda\,\mu(t)\\
&=\frac{1}{2\pi i}\int_{\{|\lambda|=\omega(t)\}}\frac{1}{\lambda}d\lambda\,\mu(t)-\frac{1}{2\pi i}\int_{\{|\lambda|=\omega(t)\}}\left(\lambda I-K(t)\right)^{-1}d\lambda\,\mu(t)\\
&=\mu(t)-\frac{1}{2\pi i}\int_{\{|\lambda|=\omega(t)\}}\left(\lambda I-K(t)\right)^{-1}d\lambda\,\mu(t),
\end{aligned}$$

and consequently,

$$P(t)\mu(t)=\frac{1}{2\pi i}\int_{\{|\lambda|=\omega(t)\}}\left(\lambda I-K(t)\right)^{-1}d\lambda\,\mu(t). \tag{8.278}$$

From residue calculus it follows that $P(t)\mu(t)=\lim_{\lambda\downarrow 0}\lambda\left(\lambda I-K(t)\right)^{-1}\mu(t)$. Since the operator $K(t)$ has the Kolmogorov property, we see that for $\lambda>0$ the operator $\lambda\left(\lambda I-K(t)\right)^{-1}$ sends positive measures to positive measures, and hence $P(t)\mu(t)$ is a positive Borel measure. By the same argument $\langle\mathbf{1},P(t)\mu(t)\rangle=1$.

The following corollary is applicable if $K(t)=L(t)^*$, where the operators satisfy the analytic maximum principle. The latter means that

$$\sup_{s>0}\sup_{t>0}\left\|sK(t)e^{sK(t)}\right\|<\infty,$$

and that the operators $L(t)$ satisfy the maximum principle. Such densely defined operators in $C_b(E)$ generate bounded analytic semigroups $e^{sL(t)}$ where s belongs to a sector with angle opening independent of t. It follows that the operators $\lambda I - L(t)$ are invertible for all $\lambda \in \mathbb{C}$ with $|(\arg \lambda)| \leqslant \beta$ with $\frac{1}{2}\pi < \beta < \pi$, and where for some constant C (independent of t) the inequality $|\lambda| \left\| (\lambda I - L(t))^{-1} \right\| \leqslant C$ holds for $\lambda \in \mathbb{C}$ with $|(\lambda)| \leqslant \beta$.

Corollary 8.3. *Let the family $K(t)$, $t \geqslant 0$, be a family of generators of weak*-continuous semigroups in $M(E)$ with the property that the operators $e^{sK(t)}$, $s \geqslant 0$, $t \geqslant 0$, map positive measures to positive measures and each operator $K(t)$ has the property that*

$$|\lambda| \operatorname{Var}(\mu) \leqslant C \operatorname{Var}(\lambda\mu - K(t)\mu), \quad \Re\lambda > 0, \ \mu \in D(K(t)), \tag{8.279}$$

where C is a constant which does not depend t. Suppose that the constants $\omega(t)$ and $c(t)$ are such that one of the following conditions

$$\operatorname{Var}\left(e^{sK(t)}\mu\right) \leqslant c(t) e^{-2\omega(t)s} \operatorname{Var}(\mu), \quad \textit{for } s > 0 \textit{ or} \tag{8.280}$$

$$|\lambda| \operatorname{Var}(\mu) \leqslant c(t) \operatorname{Var}(\lambda\mu - K(t)\mu), \quad \textit{for all } \lambda \in \mathbb{C} \textit{ such that } |\lambda| \leqslant \omega(t) \tag{8.281}$$

is satisfied for all $\mu \in M_0(E) \bigcap D(K(t))$. Let $t \mapsto \mu(t)$ be a solution to the equation $\dot{\mu}(t) = K(t)\mu(t)$, $t \geqslant 0$, with $\mu(t) \in P(E)$. If (8.246) is satisfied, then the system $\dot{\mu}(t) = K(t)\mu(t)$ is ergodic, provided that there exists a unique function $\pi(t) \in P(E)$ such that $K(t)\pi(t) = 0$.

Proof. There exists $\frac{1}{2}\pi < \beta < \pi$ such that $|\lambda| \left\| (\lambda I - K(t))^{-1} \right\| \leqslant C$ for $\lambda \in \mathbb{C}$ with $|(\lambda)| \leqslant \beta$, with C independent of t: see Theorem 8.7 and the corollaries and 8.4 and 8.5. For $\mu(t) - \pi(t) \in M_0(E) \bigcap D(K(t))$ such that $K(t)(\mu(t) - \pi(t)) = \dot{\mu}(t)$ and $\lambda \in \mathbb{C}$ such that $|\lambda - 2\omega(t)| \leqslant \dfrac{\omega(t)}{2c(t)}$, and such that $\Re\lambda \geqslant \omega(t)$, and for $\mu \in M_0(E)$ we have

$$\mu(t) - \pi(t) = \int_0^\infty e^{-s(\lambda I - 2\omega(t)I - K(t))} \left((\lambda - 2\omega(t)) (\mu(t) - \pi(t)) - \dot{\mu}(t) \right) ds,$$

and hence for such λ

$$\operatorname{Var}(\mu(t) - \pi(t)) \leqslant c(t) \frac{|\lambda - 2\omega(t)|}{\Re\lambda} \leqslant \frac{1}{2} \operatorname{Var}(\mu(t) - \pi(t)) + \frac{c(t)}{\omega(t)} \operatorname{Var}(\dot{\mu}(t)). \tag{8.282}$$

An easy application of Theorem 8.5 then completes the proof of Corollary 8.3. □

Families of semigroups $\left\{e^{sK(t)} : s \geqslant 0\right\}_{t \geqslant t_0}$ which satisfy (b) of the following theorem are called uniformly bounded and uniformly holomorphic families of operator semigroups: cf. [Blunck (2002)]. The next result will be used with $A(t) = 2\omega I \big|_{M_0(E)} + K(t) \big|_{M_0(E)}$: see Corollary 8.6 below.

Theorem 8.7. *Let $A(t)$, $t \geqslant t_0$, be a family of closed linear operators, each of which has a dense domain in a Banach space $(X, \|\cdot\|)$. Suppose that, for every $t \geqslant t_0$, and for every $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$, the inverses $(\lambda I - A(t))^{-1}$ exist and are bounded. Then the following assertions are equivalent:*

(a) $\sup\limits_{t \geqslant t_0} \sup\limits_{\Re\lambda > 0} |\lambda| \left\| (\lambda I - A(t))^{-1} \right\| < \infty$;

(b) $\sup\limits_{s>0} \sup\limits_{t \geqslant t_0} \left\| sA(t)e^{sA(t)} \right\| < \infty$ *and* $\sup\limits_{s>0} \sup\limits_{t \geqslant t_0} \left\| e^{sA(t)} \right\| < \infty$.

Proof. Most standard proofs for one generator A can be adapted to include a family of operators $A(t)$, $t \geqslant t_0$: (see e.g. [Van Casteren (1985)], page 84, or [Pazy (1983a)] Theorem 5.2 and formula (5.16)). Another thorough discussion can be found in Chapter II section 4 of [Engel and Nagel (2000)]. □

It is also a consequence of the following theorem. For convenience, and because we need to keep track of the constants an outline of the proof is included.

Theorem 8.8. *Let K be the generator of a strongly continuous semigroup with the property that for $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$ the inverse $(\lambda I - K)^{-1}$ exists as a bounded linear operator. Then the following assertions are true:*

(i) *If, for some finite constant C, the inequality*

$$|\lambda| \left\| (\lambda I - K)^{-1} \right\| \leqslant C \text{ holds for all } \lambda \in \mathbb{C} \text{ with } \Re\lambda > 0, \tag{8.283}$$

then

$$\left\| e^{tK} \right\| \leqslant \frac{e}{2} C^2 \text{ and } \left\| tKe^{tK} \right\| \leqslant eC^2 (1 + C) \text{ for all } t > 0. \tag{8.284}$$

(ii) *If there exist finite constants C_1 and C_2 such that*

$$\left\| e^{tK} \right\| \leqslant C_1 \text{ and } \left\| tKe^{tK} \right\| \leqslant C_2, \text{ for all } t > 0, \tag{8.285}$$

then

$$|\lambda| \left\| (\lambda I - K)^{-1} \right\| \leqslant C \text{ holds for all } \lambda \in \mathbb{C} \text{ with } \Re\lambda > 0. \tag{8.286}$$

Here the constant C is given by $C = 2 (C_2 e + 1) \left(C_1 + \dfrac{e}{\sqrt{2\pi}} C_2 \right)$.

Representations as in (8.287) and (8.288) below can be found in [Blunck (2001)], [Eisner (2005)], and [Eisner and Zwart (2007)].

Proof. Assertion (i) follows from the representations:

$$te^{tK} = \frac{1}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty} \lambda^2 e^{t\lambda}\left(\lambda I - K\right)^{-2}\frac{1}{\lambda^2}d\lambda; \tag{8.287}$$

$$\begin{aligned}\frac{1}{2}t^2Ke^{tK} &= \frac{1}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty} \lambda^3 e^{t\lambda}\left(\lambda I - K\right)^{-3}\frac{1}{\lambda^2}d\lambda \\ &\quad - \frac{1}{2\pi i}\int_{\omega-i\infty}^{\omega+i\infty} \lambda^2 e^{t\lambda}\left(\lambda I - K\right)^{-2}\frac{1}{\lambda^2}d\lambda,\end{aligned} \tag{8.288}$$

together with the choice $\omega = \dfrac{1}{t}$.

The proof of assertion (ii) is somewhat more delicate. At first we fix $t_0 > 0$ and we consider $t > 0$ with the property that

$$|t - t_0| \leqslant \frac{t_0}{C_2 e + 1}. \tag{8.289}$$

We notice the inequality

$$t \geqslant t_0\frac{C_2 e}{C_2 e + 1}, \tag{8.290}$$

whenever t satisfies (8.289). Moreover, for $n \geqslant 0$ we have the representation

$$e^{tK} = \sum_{\ell=0}^{n}\frac{\left(t-t_0\right)^{\ell}}{\ell!}K^{\ell}e^{t_0K} + \frac{1}{n!}\int_{t_0}^{t}\left(t-s\right)^n K^{n+1}e^{sK}ds. \tag{8.291}$$

The remainder term in (8.291) can be estimated as follows:

$$\begin{aligned}&\left\|\frac{1}{n!}\int_{t_0}^{t}(t-s)^n K^{n+1}e^{sK}ds\right\| \\ &\leqslant \frac{1}{n!}\left|\int_{t_0}^{t}(t-s)^n\frac{(n+1)^{n+1}}{s^{n+1}}\left\|\left(\frac{sK}{n+1}e^{\frac{sK}{n+1}}\right)^{n+1}\right\|ds\right| \\ &\leqslant \frac{(n+1)^{n+1}C_2^{n+1}}{n!}\left|\int_{t_0}^{t}\frac{(t-s)^n}{\left(\min\left(t,t_0\right)\right)^{n+1}}ds\right| \\ &\leqslant \frac{(n+1)^{n+1}C_2^{n+1}}{(n+1)!}\frac{|t-t_0|^{n+1}}{\left(\min\left(t,t_0\right)\right)^{n+1}}\end{aligned}$$

(employ (8.289) and (8.290))

$$\leqslant \frac{(n+1)^{n+1}}{(n+1)!}\frac{1}{C_2^{n+1}e^{n+1}}C_2^{n+1}$$

(use Stirling's formula: $(n+1)! \geqslant \sqrt{2\pi(n+1)}e^{-n-1}(n+1)^{n+1}$)

$$\leqslant \frac{1}{\sqrt{2\pi(n+1)}}. \tag{8.292}$$

This inequality clearly shows that the remainder term converges to 0 uniformly for t and t_0 satisfying: $|t-t_0| \leqslant \dfrac{t_0}{C_2 e+1}$. From (8.291) we see that for $t \in \mathbb{C}$ chosen in such a way that $|t-t_0| \leqslant \dfrac{t_0}{C_2 e+1}$ the semigroup e^{tK} can be represented as:

$$e^{tK} = e^{t_0 K} + \sum_{\ell=1}^{\infty} \frac{(t-t_0)^{\ell}}{\ell!} K^{\ell} e^{t_0 K}. \tag{8.293}$$

From (8.293) it follows that

$$\begin{aligned}\left\|e^{tK}-e^{t_0K}\right\| &\leqslant \sum_{\ell=1}^{\infty} \frac{|t-t_0|^{\ell}}{\ell!}\left\|K^{\ell}e^{t_0K}\right\| \\ &= \sum_{\ell=1}^{\infty}\left(\frac{1}{C_2e+1}\right)^{\ell}\frac{\ell^{\ell}}{\ell!}\left\|\left(\frac{t_0K}{\ell}e^{\frac{t_0K}{\ell}}\right)^{\ell}\right\| \leqslant \sum_{\ell=1}^{\infty}\frac{1}{(C_2e+1)^{\ell}}\frac{\ell^{\ell}}{\ell!}C_2^{\ell}\end{aligned}$$

(again we employ Stirling's formula $\ell! \geqslant \sqrt{2\pi\ell}e^{-\ell}\ell^{\ell}$)

$$\leqslant \sum_{\ell=1}^{\infty}\frac{(C_2e)^{\ell}}{(C_2e+1)^{\ell}}\frac{1}{\sqrt{2\pi\ell}} \leqslant \sum_{\ell=1}^{\infty}\frac{(C_2e)^{\ell}}{(C_2e+1)^{\ell}}\frac{1}{\sqrt{2\pi}} = \frac{e}{\sqrt{2\pi}}C_2. \tag{8.294}$$

Consequently, by our assumption $\left\|e^{t_0K}\right\| \leqslant C_1$, for all $t_0>0$, we get

$$\left\|e^{tK}\right\| \leqslant C_1 + \frac{e}{\sqrt{2\pi}}C_2, \tag{8.295}$$

whenever $t \in \mathbb{C}$ is chosen in such a way that (8.289) is satisfied for some $t_0>0$. If we choose $\frac{1}{3}\pi > \alpha > 0$ in such a way that $\dfrac{1}{2C_2e+2} = \sin\left(\dfrac{1}{2}\alpha\right)$, and if $|\arg(t)| \leqslant \alpha$, then t satisfies: $|t-t_0| \leqslant \dfrac{t_0}{C_2e+1}$, with $t_0=|t|$. Hence, the norm of e^{tK} satisfies (8.295). For $\lambda \in \mathbb{C}$ such that $-\frac{1}{2}\pi+\frac{1}{2}\alpha < \arg(\lambda) < \frac{1}{2}\pi+\frac{1}{2}\alpha$ we have:

$$(\lambda I - K)^{-1} = e^{-\frac{i}{2}\alpha}\int_0^{\infty}\exp\left(-\lambda e^{-\frac{i}{2}\alpha}sI + e^{-\frac{i}{2}\alpha}sK\right)ds, \tag{8.296}$$

and hence

$$\begin{aligned}
&|\lambda| \left\|(\lambda I - K)^{-1}\right\| \\
&\leqslant \int_0^\infty \left|\exp\left(-\lambda e^{-\frac{i}{2}\alpha} s\right)\right| ds \left\|\exp\left(e^{-\frac{i}{2}\alpha} sK\right)\right\| ds \\
&\leqslant |\lambda| \int_0^\infty \exp\left(-|\lambda| \cos\left(\arg(\lambda) - \frac{1}{2}\alpha\right) s\right) ds \left(C_1 + \frac{e}{\sqrt{2\pi}} C_2\right) \\
&= \frac{1}{\cos\left(\arg(\lambda) - \frac{1}{2}\alpha\right)} \left(C_1 + \frac{e}{\sqrt{2\pi}} C_2\right). \qquad (8.297)
\end{aligned}$$

By the same token we also get, for $\lambda \in \mathbb{C}$ such that $-\frac{1}{2}\pi - \frac{1}{2}\alpha < \arg(\lambda) < \frac{1}{2}\pi - \frac{1}{2}\alpha$,

$$(\lambda I - K)^{-1} = e^{\frac{i}{2}\alpha} \int_0^\infty \exp\left(-\lambda e^{\frac{i}{2}\alpha} sI + e^{\frac{i}{2}\alpha} sK\right) ds, \qquad (8.298)$$

and hence

$$\begin{aligned}
&|\lambda| \left\|(\lambda I - K)^{-1}\right\| \\
&\leqslant \int_0^\infty \left|\exp\left(-\lambda e^{\frac{i}{2}\alpha} s\right)\right| ds \left\|\exp\left(e^{\frac{i}{2}\alpha} sK\right)\right\| ds \\
&\leqslant |\lambda| \int_0^\infty \exp\left(-|\lambda| \cos\left(\arg(\lambda) + \frac{1}{2}\alpha\right) s\right) ds \left(C_1 + \frac{e}{\sqrt{2\pi}} C_2\right) \\
&= \frac{1}{\cos\left(\arg(\lambda) + \frac{1}{2}\alpha\right)} \left(C_1 + \frac{e}{\sqrt{2\pi}} C_2\right). \qquad (8.299)
\end{aligned}$$

From (8.297) and (8.299) we infer:

$$|\lambda| \left\|(\lambda I - K)^{-1}\right\| \leqslant \frac{1}{\cos\left(|\arg(\lambda)| - \frac{1}{2}\alpha\right)} \left(C_1 + \frac{e}{\sqrt{2\pi}} C_2\right), \qquad (8.300)$$

for $-\frac{1}{2}\pi - \frac{1}{2}\alpha < \arg(\lambda) < \frac{1}{2}\pi + \frac{1}{2}\alpha$. Inequality (8.286) in Theorem 8.8 follows from (8.300) with $\lambda \in \mathbb{C}$ such that $|\arg(\lambda)| < \frac{1}{2}\pi$.

This completes the proof of Theorem 8.8. □

An inspection of the proof of assertion (ii) in Theorem 8.8, in particular inequality (8.300), yields the following result, which says that the resolvent family of a bounded analytic semigroup is bounded in a sector with an opening which is larger than the open right half-plane.

Corollary 8.4. *Let the hypotheses and notation be as in Theorem 8.8. Choose the angle* $\frac{1}{3}\pi > \alpha > 0$ *in such a way that* $\sin\left(\frac{1}{2}\alpha\right) = \dfrac{1}{2C_2 e + 2}$. *Choose* $0 \leqslant \beta < \frac{1}{2}\alpha$. *Then*

$$|\lambda| \left\|(\lambda I - K)^{-1}\right\| \leqslant \frac{1}{\sin\left(\frac{1}{2}\alpha - \beta\right)} \left(C_1 + \frac{e}{\sqrt{2\pi}} C_2\right), \quad |\arg \lambda| \leqslant \frac{1}{2}\pi + \beta. \qquad (8.301)$$

The result in Corollary 8.4 extends to uniformly bounded and uniformly analytic semigroups. Notice that (8.303) is equivalent to an inequality of the form ($t \geqslant t_0$):

$$|\lambda| \left\| (\lambda I - A(t))^{-1} \right\| \leqslant C, \text{ for } \lambda \in \mathbb{C} \text{ with } \Re\lambda > 0, \tag{8.302}$$

where the constants C and C_1, C_2 are related in an explicit manner: see Theorem 8.8.

Corollary 8.5. *Let $A(t)$, $t \geqslant t_0$, be a family of closed densely defined linear operators. Suppose there exist finite constants C_1 and C_2 such that*

$$\left\| e^{sA(t)} \right\| \leqslant C_1 \text{ and } \left\| sA(t)e^{sA(t)} \right\| \leqslant C_2, \text{ for all } s > 0 \text{ and for all } t \geqslant t_0. \tag{8.303}$$

Choose $0 < \alpha < \frac{1}{3}\pi$ in such a way that $\sin\left(\frac{1}{2}\alpha\right) = \dfrac{1}{2C_2e + 2}$. Fix $0 \leqslant \beta < \frac{1}{2}\alpha$. Then, for all $t \geqslant t_0$, the inequality

$$|\lambda| \left\| (\lambda I - A(t))^{-1} \right\| \leqslant C(\beta) \tag{8.304}$$

is true for all $\lambda \in \mathbb{C}$ with $|\arg\lambda| \leqslant \frac{1}{2}\pi + \beta$. Here the constant $C(\beta)$ is given by $C(\beta) = \dfrac{1}{\sin\left(\frac{1}{2}\alpha - \beta\right)}\left(C_1 + \dfrac{e}{\sqrt{2\pi}}C_2\right)$.

In the following corollary we use Theorem 8.7 and Corollary 8.5 with $A(t) = 2\omega I\,|_{M_0(E)} + K(t)\,|_{M_0(E)}$.

Corollary 8.6. *Let the function $t \mapsto \mu(t)$ solve the equation:*

$$\dot{\mu}(t) = K(t)\mu(t), \quad \mu(t) \in P(E).$$

Suppose that $\lim\limits_{t\to\infty} \mathrm{Var}\,(\dot{\mu}(t)) = 0$, and that there exists $\omega > 0$ such that

$$c := \sup_{s,t>0} \left\| s\,(2\omega I + K(t))\, e^{s(2\omega I + K(t))}\,|_{M_0(E)} \right\| < \infty. \tag{8.305}$$

If, in addition, there exists only one continuous function $t \mapsto \pi(t)$ with values in $P(E)$ such that $K(t)\pi(t) = 0$, then $\lim_{t\to\infty} \mathrm{Var}\,(\mu(t) - \pi(t)) = 0$.

Notice that the operator $(2\omega I + K(t))\, e^{s(2\omega I + K(t))}$ is a mapping from $M_0(E)$ to $M_0(E)$.

Proof. An appeal to Corollary 8.5 together with the hypothesis in inequality (8.305) shows that there exists a finite constant c_1 such that

$$|\lambda| \left\| \left(\lambda I\,|_{M_0(E)} - (2\omega I + K(t))\,|_{M_0(E)}\right)^{-1} \right\| \leqslant c_1 \quad \text{for all } \lambda \text{ with } \Re\lambda > 0. \tag{8.306}$$

The latter result follows in fact from the theory of families of uniform holomorphic semigroups (the inequality (8.306) is uniform in $t > t_0$). Consequently, we obtain:

$$|\lambda - 2\omega| \left\| \left(\lambda I \big|_{M_0(E)} - (2\omega I + K(t)) \big|_{M_0(E)} \right)^{-1} \right\| \leqslant 3c_1$$

for all λ with $\Re\lambda \geqslant \omega$.

The result in Corollary 8.6 then follows from Theorem 8.5. □

Examples of operators L which generate analytic Feller semigroups can be found in [Taira (1997)]. Other valuable sources of information are [Metafune *et al.* (2002a)] and [Taira (1992)].

8.5 Conclusions

In this chapter we discussed some properties of the fundamental operator of the non-stationary, or time-dependent continuous system (8.6). Moreover, in some particular cases, when we deal with a family of Kolmogorov operators $K(t)$, we introduce and prove some efficient criteria for checking ergodicity (Theorem 8.5). This is done by using the Dunford projection on the eigenspace corresponding to the critical eigenvalue 0 of $K(t)$.

The properties of the families of semigroups $\left\{e^{sK(t)} : s \geqslant 0\right\}_{t \geqslant t_0}$ are examined in detail in Theorem 8.7 and Theorem 8.8 as well as in Corollary 8.4 and Corollary 8.5. The obtained results allow us to present Corollary 8.6 providing the ergodicity of non-stationary system in terms of bounded analytic semigroups. In addition, in §9.4 we discuss a rather general situation in which we have a spectral gap: see e.g. Proposition 9.16. Some of this work was based on ideas and concepts of Katilova [Katilova (2008, 2004, 2005)]. What follows next can be found in [Van Casteren (2005a)]. Theorem 8.9 is inspired by ideas in Nagy and Zemanek: see [Nagy and Zemánek (1999)]. The result can also be found in the Ph.-D. thesis of Katilova: see [Katilova (2004)], Theorem 8.9.

Theorem 8.9. *Let M be a bounded linear operator in a Banach space X. By definition the sub-space X_0 of X is the $\|\cdot\|$-closure of the vector sum of the range and zero-space of $I - M$: $X_0 = \overline{R(I-M) + N(I-M)}^{\|\cdot\|}$. Suppose that the spectrum of M is contained in the open unit disc union $\{1\}$. The following assertions are equivalent:*

(i) $\sup_{|\lambda|<1} \left\|(1-\lambda)(I-\lambda M)^{-1} x\right\| < \infty$ *for every* $x \in X_0$;

(ii) $\sup_{n\in\mathbb{N}} \|M^n x\| < \infty$ *and* $\sup_{n\in\mathbb{N}}(n+1)\|M^n (I-M)x\| < \infty$ *for every* $x \in X_0$;

(iii) $\sup_{t>0} \left\|e^{t(M-I)}x\right\| < \infty$ *and* $\sup_{t>0} \left\|t(M-I)e^{t(M-I)}x\right\| < \infty$ *for* $\forall x \in X_0$;

(iv) *There exists* $\frac{1}{2}\pi < \alpha < \pi$ *such that for all* $x \in X_0$:

$$\sup\left\{|\lambda|\left\|(\lambda I - (M-I))^{-1}x\right\| : -\alpha < \arg(\lambda) < \alpha\right\} < \infty;$$

(v) *There exists* $\frac{1}{2}\pi < \alpha < \pi$ *such that for all* $x \in X_0$:

$$\sup\left\{\left\|(I-M)((\lambda+1)I-M)^{-1}x\right\| : -\alpha < \arg(\lambda) < \alpha\right\} < \infty;$$

(vi) *For every* $x \in X_0$ *the following limits exist*

$$Px := \lim_{n\to\infty} M^n x \text{ and } (I-P)x = \lim_{\substack{re^{i\vartheta}\to 1\\ 0<r<1}} (I-M)\left(I - re^{i\vartheta}M\right)^{-1}x;$$

(vii) *For every* $x \in X_0$ *the following limit exists*

$$(I-P)x := \lim_{\substack{re^{i\vartheta}\to 1\\ 0<r<1}} (I-M)\left(I - re^{i\vartheta}M\right)^{-1}x.$$

Moreover, if M *satisfies one of the conditions (i) through (vii), then*

$$X_0 = \overline{R(I-M)}^{\|\cdot\|} + N(I-M).$$

Remark 8.15. The Banach-Steinhaus theorem implies that in (i) through (v) in Theorem 8.9 the vector norms may be replaced with the operator norm restricted to X_0; i.e. the operator M must be restricted to X_0. These assertions (i) through (v) are also equivalent if X_0 is replaced with the space X. This fact will be used in Definition 8.8.

Conditions (a) and (b) of the following corollary from [Arendt *et al.* (2001)] are satisfied, if the space X is reflexive. The closed range condition in (c) has been used by Lin in [Lin (1974)] and in [Lin (1975)]; in the latter reference he also tied it up with Doeblin's ergodicity condition. For a precise formulation of Doeblin's ergodicity condition see item (ii) in Definition 10.8.

Corollary 8.7. *Let* M *be a bounded linear operator in a Banach space* $(X, \|\cdot\|)$. *As in Theorem 8.9 let* X_0 *be the closure in* X *of the sub-space* $R(I-M) + N(I-M)$. *Suppose that, for* $0 < \lambda < 1$, *the inverse operators* $(I-\lambda M)^{-1}$ *exist and are bounded, and that* $\sup_{0<\lambda<1}(1-\lambda)\left\|(I-\lambda M)^{-1}\right\| < \infty$. *If one of the following conditions:*

(a) *the zero space of the operator* $(I - M)^{**}$*, which is a sub-space of the bidual space* X^{**} *is in fact a subspace of* X*;*
(b) *the* $\sigma(X^*, X)$*-closure of* $R((I - M)^*)$ *coincides with its* $\|\cdot\|$*-closure;*
(c) *the range of* $I - M$ *is closed in* X*;*

is satisfied, then the space X_0 *coincides with* X*, and hence all assertions in Theorem 8.9 are equivalent with* X *replacing* X_0*.*

Remark 8.16. If $\sup_{n\in\mathbb{N}} \|M^n\| < \infty$, then $\sup_{0<\lambda<1} (1-\lambda)\left\|(I-\lambda M)^{-1}\right\| < \infty$.

Definition 8.8. An operator M which satisfies the equivalent conditions (i) – (v) of Theorem 8.9 with the space X replacing X_0 is called an analytic operator.

Proof. [Proof of Corollary 8.7.] If the range of $I - M$ is closed, then by the closed range theorem, the range of $I - M$ is weak*-closed and hence (c) implies (b). We will prove that (a) as well as (b) implies $X_0 = X$. First we assume (a) to be satisfied. Pick $x \in X$, and consider

$$x = (I - M)(I - \lambda M)^{-1} x + (1 - \lambda) M (I - \lambda M)^{-1} x = x - x_\lambda + x_\lambda, \tag{8.307}$$

where $x_\lambda = (1-\lambda) M (I - \lambda M)^{-1} x$. Then $\sup_{0<\lambda<1} \|x_\lambda\| < \infty$, and consequently the family x_λ, $0 < \lambda < 1$, has a point of adherence x^{**} in X^{**}; i.e. x^{**} belongs to the $\sigma(X^{**}, X^*)$-closure of the subset $\{x_\lambda : 1 - \eta < \lambda < 1\}$, and this for every $0 < \eta < 1$. Fix $x^* \in X^*$. Then

$$\begin{aligned} &\left|\left\langle (1-\lambda) M (I-\lambda M)^{-1} x, (I-M)^* x^* \right\rangle\right| \\ &= \left|\left\langle (1-\lambda)(I-M)(I-\lambda M)^{-1} x, M^* x^* \right\rangle\right| \\ &\leqslant (1-\lambda)\left\|(I-M)(I-\lambda M)^{-1} x\right\| \|M^* x^*\|. \end{aligned} \tag{8.308}$$

Since $\sup_{0<\lambda<1} (1-\lambda)\left\|(I-\lambda M)^{-1}\right\| < \infty$, the identity

$$(I-M)(I-\lambda M)^{-1} = \frac{1}{\lambda}\left(I - (1-\lambda)(I-\lambda M)^{-1}\right)$$

yields that $\sup_{0<\lambda<1} \left\|(I-M)(I-\lambda M)^{-1}\right\| < \infty$. Consequently, (8.308) implies

$$\begin{aligned} \left\langle x^{**}, (I-M)^* x^* \right\rangle &= \lim_{\lambda\uparrow 1} \left\langle x_\lambda, (I-M)^* x^* \right\rangle \\ &= \lim_{\lambda\uparrow 1} (1-\lambda)\left\langle (I-M)(I-\lambda M)^{-1} x, M^* x^* \right\rangle = 0. \end{aligned}$$

Hence x^{**} annihilates $R\left((I-M)^*\right)$ and so it belongs to the zero space of the operator $(I-M)^{**}$. By assumption this zero space is a subspace of X. We infer that the vector x can be written as $x = x - x_1 + x_1$, where x_1 is a member of $N(I-M)$, and where $x - x_1$ belongs to the weak closure of the range of $I - M$. However this weak closure is the same as the norm-closure of $R(I-M)$. Altogether this shows $X = X_0 =$ $\|\cdot\|$-closure of $R(I-M) + N(I-M)$.

Next we assume that (b) is satisfied. Let x_0^* be an element of X^* which annihilates X_0; i.e. which has the property that $\langle x, x_0^*\rangle = 0$ for all $x \in X_0$. Then x_0^* annihilates $R(I-M)$, and hence it belongs to zero-space of $(I-M)^*$. Since x_0^* also annihilates the zero-space of $I-M$, it belongs to the weak*-closure of $R\left((I-M)^*\right)$. By assumption (b), we see that x_0^* is a member of its norm-closure; i.e. x_0^* belongs to the intersection $N\left((I-M)^*\right) \bigcap \overline{R\left((I-M)^*\right)}^{\|\cdot\|}$. We will show that $x_0^* = 0$. By the Hahn-Banach theorem [Hahn (1958)] it then follows that $X_0 = X$. Since x_0^* belongs to the $\|\cdot\|$-closure of $R\left((I-M)^*\right)$, it follows that

$$x_0^* = \|\cdot\|\text{-}\lim_{\lambda\uparrow 1} (I-M)^* \left((I-\lambda M)^*\right)^{-1} x_0^*. \tag{8.309}$$

To see this we first suppose that $x_0^* = (I-M)^* x_1^*$. Then

$$\begin{aligned}
&(I-M)^* x_1^* - (I-M)^* \left((I-\lambda M)^*\right)^{-1} (I-M)^* x_1^* \\
&= (1-\lambda) M^* \left((I-\lambda M)^*\right)^{-1} (I-M)^* x_1^*.
\end{aligned} \tag{8.310}$$

Since the family $M^* \left((I-\lambda M)^*\right)^{-1} (I-M)^* x_1^*$, $0 < \lambda < 1$, is bounded, we see that (8.309) is a consequence of (8.310) provided x_0^* belongs to the range of $(I-M)^*$. By the uniform boundedness of the family $(I-M)^* \left((I-\lambda M)^*\right)^{-1}$, $0 < \lambda < 1$, the same conclusion is true if x_0^* belongs to the closure of the range of $(I-M)^*$. Since, in addition, x_0^* is a member of $N\left((I-M)^*\right)$, it follows that $x_0^* = 0$. This proves Corollary 8.7.

This completes the proof of Corollary 8.7. □

Proof. [Proof of Theorem 8.9.] (i) $\Longrightarrow$ (ii). Fix $0 < r < 1$. The following representations from Lyubich [Lyubich (1999)] are being used:

$$(n+1)M^n = \frac{1}{2\pi i}\int_{|\lambda|=r} (1-\lambda)^2 (I-\lambda M)^{-2} \frac{d\lambda}{\lambda^{n+1}(1-\lambda)^2}; \tag{8.311}$$

$$\frac{1}{2}(n+1)(n+2)M^n(I-M)$$

$$
\begin{aligned}
&= \frac{1}{2\pi i}\int_{|\lambda|=r} (1-\lambda)^2 (I-M)(I-\lambda M)^{-3} \frac{d\lambda}{\lambda^{n+1}(1-\lambda)^2} \\
&= \frac{1}{2\pi i}\int_{|\lambda|=r} (1-\lambda)^2 (I-\lambda M)^{-2} \frac{1}{\lambda^{n+2}(1-\lambda)^2} d\lambda \\
&\quad - \frac{1}{2\pi i}\int_{|\lambda|=r} (1-\lambda)^3 (I-\lambda M)^{-3} \frac{1}{\lambda^{n+2}(1-\lambda)^2} d\lambda. \qquad (8.312)
\end{aligned}
$$

Put $C := \sup\left\{\left\|(1-\lambda)(I-\lambda M)^{-1}|X_0\right\| : |\lambda|<1\right\}$. From (8.311) we infer

$$
(n+1)\|M^n\| \leqslant \frac{C^2}{r^n}\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\left|1-re^{i\vartheta}\right|^2}d\vartheta = \frac{C^2}{r^n}\frac{1}{1-r^2}. \qquad (8.313)
$$

The choice $r^2 = \dfrac{n}{n+2}$ yields

$$
\|M^n|X_0\| \leqslant \frac{2}{3}eC^2. \qquad (8.314)
$$

In the same spirit from (8.312) we obtain

$$
\frac{1}{2}(n+1)(n+2)\|M^n(M-I)|X_0\| \leqslant \left(C^2+C^3\right)\frac{1}{r^{n+1}}\frac{1}{1-r^2}.
$$

The choice $r^2 = \dfrac{n+1}{n+3}$ yields the inequality:

$$
(n+1)\|M^n(M-I)|X_0\| \leqslant \frac{4e}{3}\left(C^2+C^3\right).
$$

This proves the implication (i) $\Longrightarrow$ (ii).

(ii) $\Longrightarrow$ (iii). The representations (see [Nagy and Zemánek (1999)])

$$
e^{t(M-I)} = e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{k!}M^k \quad\text{and}\quad t(M-I)e^{t(M-I)} = e^{-t}\sum_{k=0}^{\infty}\frac{t^{k+1}}{k!}M^k(M-I)
$$

show that (iii) is a consequence of (ii).

(iii) $\Longrightarrow$ (iv). This is a (standard) result in analytic operator semigroup theory: see e.g. [Van Casteren (1985)], Chapter 5, Theorem 5.1.

(iv) $\Longrightarrow$ (v). The equality

$$
(I-M)\left((\lambda+1)I-M\right)^{-1} = I-\lambda\left(\lambda I-(M-I)\right)^{-1}
$$

shows the equivalence of (iv) and (v).

(v) $\Longrightarrow$ (i). Fix $x \in X_0$. The choice

$$
\lambda = -1+e^{-i\vartheta} = -2i\sin\left(\frac{1}{2}\vartheta\right)e^{-\frac{1}{2}i\vartheta}, \quad |\vartheta| \leqslant 2\alpha,
$$

yields the boundedness of the function

$$\vartheta \mapsto (I - M)\left(I - e^{i\vartheta}M\right)^{-1} x$$

on the interval $[-\alpha, \alpha]$. Since, for $|\lambda| = 1$, $\lambda \neq 1$, the function

$$\lambda \mapsto (I - M)(I - \lambda M)^{-1} x$$

is continuous, it follows that this function is bounded on the unit circle. The maximum modulus theorem shows that this function is bounded on the unit disc, which is assertion (i).

(i) $\Longrightarrow$ (vi). Fix $x \in X_0$. For $0 < r < 1$ and $\vartheta \in \mathbb{R}$ we also have

$$\begin{aligned}
&\left(I - P\left(re^{i\vartheta}\right)\right)(I - M)\, x \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1 - r^2}{1 - 2r\cos(\vartheta - t) + r^2}(I - M)\left(I - e^{it}M\right)^{-1}(I - M)\, x\, dt \\
&= (I - M)\left(I - re^{i\vartheta}M\right)^{-1}(I - M)\, x. \qquad (8.315)
\end{aligned}$$

In (8.315) we use the continuity of the boundary function

$$e^{it} \mapsto (I - M)\left(I - e^{it}M\right)^{-1}(I - M)\, x \qquad (8.316)$$

to show that

$$\lim_{re^{i\vartheta}\to 1,\, 0\leqslant r<1}\left(I - P\left(re^{i\vartheta}\right)\right)(I - M)\, x = (I - P)(I - M)\, x = (I - M)\, x \qquad (8.317)$$

exists, and that $I - P$ is a bounded projection on X_0. From (i) it follows that the function $\lambda \mapsto (I - M)(I - \lambda M)^{-1} x$ is uniformly bounded on the unit disc, and hence that the limit in (8.317) exists for all y in the closure of $R(I - M)$. In addition, for such vectors y we have $(I - P)\, y = y$. The limit in (8.317) trivially exists for $x \in X$ such that $Mx = x$, and hence we conclude that the limit in (i) exists for all $x \in X_0$, because $x = (I-P)x+Px$, where $(I - P)x$ belongs to the closure of the range of $I - M$ and where

$$\begin{aligned}
Px = x - (I - P)x &= x - \lim_{\lambda\uparrow 1}(I - M)(I - \lambda M)^{-1} x \\
&= \lim_{\lambda\uparrow 1}(1 - \lambda)\, M\, (I - \lambda M)^{-1} x. \qquad (8.318)
\end{aligned}$$

From (8.318) it follows that $(I - M)\, Px = 0$. In addition, from (ii), which is equivalent to (i), we see that $\lim_{n\to\infty} M^n y = 0$ for all y in the range of $I - M$; here we use the boundedness of the sequence $(n + 1)\, M^n\, (I - M)$, $n \in \mathbb{N}$. The boundedness of the sequence M^n, $n \in \mathbb{N}$, then yields $\lim_{n\to\infty} M^n y = 0$ for $y \in R(I - P)$, because the range of $I - M$ is dense in the range of $I - P$. An arbitrary $x \in X_0$ can be written as $x = (I - P)\,x + Px$. From the previous arguments it follows that $\lim_{n\to\infty} M^n x = Px$. Fix $x \in X_0$. Altogether this shows the implication (v) $\Longrightarrow$ (vi), provided we show the continuity of the function in (8.316) in the sense that $\lim_{t\to 0} (I - M)\left(I - e^{it}M\right)^{-1}(I - M)\,x = (I - M)\,x$. However, this follows from the identity

$$\begin{aligned}&(I - M)\left(I - e^{it}M\right)^{-1}(I - M)\,x - (I - M)\,x\\&= \left(e^{it} - 1\right)(I - M)\left(I - e^{it}M\right)^{-1} Mx,\end{aligned}$$

together with the uniform boundedness (in $0 < |t| \leqslant \pi$) of the family of operators:

$$(I - M)\left(I - e^{it}M\right)^{-1}.$$

In the latter we use the implication (v) $\Longrightarrow$ (i).

The implication (vi) $\Longrightarrow$ (vii) being trivial there remains to be shown that (vii) implies (i). For this purpose we fix $x \in X_0$ and we consider the continuous function on the closed unit disc, defined by

$$F(\lambda)x := \begin{cases} (I - M)\,(I - \lambda M)^{-1}\,x & \text{for } |\lambda| \leqslant 1,\ \lambda \neq 1,\\ (I - P)x = \lim\limits_{\substack{\lambda\to 1\\ |\lambda|<1}} (I - M)\,(I - \lambda M)^{-1}\,x & \text{for } \lambda = 1.\end{cases}$$

From (vii) it follows that the function $F(\lambda)x$ is well-defined and continuous. Hence it is bounded. The theorem of Banach-Steinhaus then implies (i) completing the proof of Theorem 8.9. □

For more recent results about stability and asymptotic behavior of linear semigroups the reader is referred to [van Neerven (1996)] or to [Eisner (2010)]. Books on operator semigroups are e.g. [Pazy (1983b)], [Goldstein (1985)], [Engel and Nagel (2000)], [Balakrishnan (2000)], [Kantorovitz (2010)].

8.6 Another characterization of generators of analytic semigroups

Let L be a closed linear operator with domain $D(L)$ and $R(L)$ in a Banach space $(X, \|\cdot\|)$ with topological dual $(X^*, \|\cdot\|)$. Suppose that $D(L)$ is dense and that there exists $\lambda \in \mathbb{C}$, $\Re\lambda > 0$ such that $(\lambda I - L) = X$. We want to give a characterization of generators of bounded analytic semigroups purely in terms of dual elements and arguments of complex numbers of the form

$\dfrac{\langle Lx, x^*\rangle}{\langle x, x^*\rangle}$, $x \in D(L)$, $x^* \in X^*$, $\langle x, x^*\rangle \neq 0$. For a concise notation we introduce the following subsets and quantities. Fix $0 < \eta < 1$. Put

$$S_1(x,\eta) = \{x^* \in X^* : \|x^*\| \leqslant 1, |\langle x, x^*\rangle| \geqslant \eta \|x\|\}, \quad x \in X. \tag{8.319}$$

For brevity write, for $x \in D(L)$, $\langle x, x^*\rangle \neq 0$,

$$f_L(x, x^*) = \frac{\langle Lx, x^*\rangle}{\langle x, x^*\rangle}, \quad \text{and} \quad q_L(x,\eta) = \frac{\inf\limits_{x^*\in S_1(x,\eta)} |f_L(x,x^*)|}{\sup\limits_{x^*\in S_1(x,\eta)} |f_L(x,x^*)|}. \tag{8.320}$$

If $L = 0$, then by definition $q_L(x,\eta) = 0$. In addition, the following quantities are introduced ($x \in D(L)$):

$$\alpha_1(x,\eta) = \inf_{x^*\in S_1(x,\eta)} \arg f_L(x,x^*), \quad \alpha_2(x,\eta) = \sup_{x^*\in S_1(x,\eta)} \arg f_L(x,x^*),$$

and

$$\beta_L(x,\eta) = \max\left(\frac{1}{2}(\alpha_2(x,\eta) - \alpha_1(x,\eta)), \alpha_2(x,\eta) - \frac{\pi}{2}, -\alpha_1(x,\eta) - \frac{\pi}{2}\right). \tag{8.321}$$

The following result follows from Lemma 8.9 below and standard results on generation of bounded analytic semigroups.

Theorem 8.10. *Let L be a closed linear operator with dense domain $D(L)$ in a Banach space $(X, \|\cdot\|)$ with dual $(X^*, \|\cdot\|)$. Fix $\eta \in (0,1)$, and let $S_1(x,\eta)$, $x \in X$, be as in (8.319), and define the quantities $q_L(x,\eta)$ and $\beta_L(x,\eta)$, $x \in D(L)$, as in (8.320) and (8.321) respectively. Suppose that there exists $\lambda \in \mathbb{C}$, $\Re\lambda > 0$, such that $R(\lambda I - L) = X$. Then the following assertions are equivalent:*

(i) The operator L generates a bounded analytic semigroup;
(ii) There exists a $\delta(\eta) > 0$ such that for all $x \in D(L)$ the inequality in (8.337) holds:

$$\max\left(\frac{1 - q_L(x,\eta)}{1 + q_L(x,\eta)}, \sin\frac{1}{2}\beta_L(x,\eta)\right) \geqslant \delta(\eta); \tag{8.322}$$

(iii) The following inequality holds:

$$\inf_{x\in X,\,\|x\|=1} \inf_{\Re\lambda>0} \sup_{x^*\in S_1(x,\eta)} \left|1 - \frac{\langle \lambda Lx, x^*\rangle}{\langle x, x^*\rangle}\right| > 0. \tag{8.323}$$

(iv) There exists a strictly finite constant C such that the following inequality holds for all $x \in D(L)$ and all $\lambda \in \mathbb{C}$, $\Re\lambda \geqslant 0$:

$$|\lambda| \, \|x\| \leqslant C \, \|\lambda x - Lx\| . \tag{8.324}$$

We need some elementary results for complex numbers and functions.

Lemma 8.7. *Let $w \neq 0$ be a complex number. Then the following inequalities hold:*

$$\begin{aligned} \max\left(\left|1-|w|\right|, \frac{1}{2}\left|1-\frac{w}{|w|}\right|\right) &\leqslant |1-w| \leqslant \left|1-|w|\right| + \left|1-\frac{w}{|w|}\right| \\ &\leqslant 2\max\left(\left|1-|w|\right|, \left|1-\frac{w}{|w|}\right|\right). \end{aligned} \tag{8.325}$$

Proof. Put $w = |w| \, e^{i\vartheta}$, $-\pi \leqslant \vartheta \leqslant \pi$. Then

$$\left|1-\frac{w}{|w|}\right| = 2\left|\sin\frac{1}{2}\vartheta\right|. \tag{8.326}$$

By writing $|1-w|^2 = 1 - 2\,|w|\cos\frac{1}{2}\vartheta + |w|^2$ the first inequality follows by squaring both sides and using (8.326). The second inequality from the equality

$$|1-w| = \left|2i\sin\frac{1}{2}\vartheta + (1-|w|)\, e^{i\frac{1}{2}\vartheta}\right|,$$

which can be checked easily. The third inequality in (8.325) being trivial this completes the proof of Lemma 8.7. □

Lemma 8.8. *Let α_1 and α_2 be real numbers such that $-\pi \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \pi$. Put $\beta = \max\left(\frac{1}{2}(\alpha_2-\alpha_1), \alpha_2 - \frac{1}{2}\pi, -\alpha_1 - \frac{1}{2}\pi\right)$. Then*

$$\inf_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]} \sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}|t-\vartheta| = \inf_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]} \sup_{t\in\{\alpha_1,\alpha_2\}} \sin\frac{1}{2}|t-\vartheta| = \sin\frac{1}{2}\beta. \tag{8.327}$$

Proof. First we write

$$\begin{aligned} M := & \inf_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]} \sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}|t-\vartheta| \\ = & \inf_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]} \max\left(\sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}(t-\vartheta), \sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}(\vartheta-t)\right). \end{aligned} \tag{8.328}$$

From (8.328) we deduce

$$M \geqslant \inf_{\vartheta \in \left[-\frac{1}{2}\pi, \frac{1}{2}\pi\right]} \max\left(\sin\frac{1}{2}\left(\alpha_2 - \vartheta\right), \sin\frac{1}{2}\left(\vartheta - \alpha_1\right)\right). \tag{8.329}$$

Then we distinguish cases: $\alpha_1 + \alpha_2 \leqslant -\pi$, $-\pi \leqslant \alpha_1 + \alpha_2 \leqslant \pi$, and $\pi \leqslant \alpha_1 + \alpha_2$. If $-2\pi \leqslant \alpha_1 + \alpha_2 \leqslant -\pi$, then from (8.329) we get $M \geqslant \sin\frac{1}{2}\left(-\alpha_1 - \frac{1}{2}\pi\right)$ with $\vartheta = -\frac{1}{2}\pi$, if $-\pi \leqslant \alpha_1 + \alpha_2 \leqslant \pi$ then $M \geqslant \sin\frac{1}{4}\left(\alpha_2 - \alpha_1\right)$ with $\vartheta = \frac{1}{2}\left(\alpha_1 + \alpha_2\right)$, and finally, if $\pi \leqslant \alpha_1 + \alpha_2 \leqslant 2\pi$, then it turns out that $M \geqslant \sin\frac{1}{2}\left(\alpha_2 - \frac{1}{2}\pi\right)$ with $\vartheta = \frac{1}{2}\pi$. This shows $M \geqslant \sin\frac{1}{2}\beta$. In order to obtain an upper bound we write:

$$\begin{aligned}
M_1 :&= \max\left(\sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}\left(t - \frac{1}{2}\pi\right), \sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}\left(\frac{1}{2}\pi - t\right)\right); \\
M_2 :&= \max\left(\sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}\left(t - \frac{1}{2}\left(\alpha_1 + \alpha_2\right)\right),\right. \\
&\qquad \left.\sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}\left(\frac{1}{2}\left(\alpha_1 + \alpha_2\right) - t\right)\right); \\
M_3 :&= \max\left(\sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}\left(t + \frac{1}{2}\pi\right), \sup_{t\in[\alpha_1,\alpha_2]} \sin\frac{1}{2}\left(-\frac{1}{2}\pi - t\right)\right),
\end{aligned} \tag{8.330}$$

and notice that $M_1 = \sin\frac{1}{2}\left(\alpha_2 - \frac{1}{2}\pi\right)$ if $\alpha_1 + \alpha_2 \geqslant \pi$, $M_2 = \sin\frac{1}{4}\left(\alpha_2 - \alpha_1\right)$ if $-\pi \leqslant \alpha_1 + \alpha_2 \leqslant \pi$, and $M_3 = \sin\frac{1}{2}\left(-\alpha_1 - \frac{1}{2}\pi\right)$ if $\alpha_1 + \alpha_2 \leqslant -\pi$. It follows that $M \leqslant \max\left(M_1, M_2, M_3\right)$. This concludes the proof of Lemma 8.8. □

Lemma 8.9. *Put* $S_1(x,\eta) = \{x^* \in X^* : \|x^*\| \leqslant 1, |\langle x, x^*\rangle| \geqslant \eta\|x\|\}$, $0 < \eta \leqslant 1$, $x \in X$. *The notation as in (8.319), (8.320), and (8.321) is in use. Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence in* $D(L)$. *Then* $\lim_{n\to\infty}\beta_L(x_n,\eta) = 0$ *if and only if*

$$-\frac{\pi}{2} \leqslant \liminf_{n\to\infty} \alpha_1(x_n,\eta) = \limsup_{n\to\infty} \alpha_2(x_n,\eta) \leqslant \frac{\pi}{2}. \tag{8.331}$$

Finally, put

$$\begin{aligned}
\delta_L(\eta) = \inf_{\Re\lambda>0} \inf_{x\in D(L),\,\|x\|=1} \sup_{x^*\in S_1(x,\eta)} \max\Big(&\left|1 - |\lambda f_L(x,x^*)|\right|, \\
&\sin\frac{1}{2}\left|\arg\left(\lambda f_L(x,x^*)\right)\right|\Big).
\end{aligned} \tag{8.332}$$

Then

$$\inf_{\Re\lambda>0}\ \sup_{x^*\in S_1(x,\eta)} \left|1-\left|\lambda f_L\left(x,x^*\right)\right|\right| = \frac{1-q_L(x,\eta)}{1+q_L(x,\eta)}, \quad \textit{and} \tag{8.333}$$

$$\begin{aligned} &\inf_{\Re\lambda>0}\ \sup_{x^*\in S_1(x,\eta)} \sin\frac{1}{2}\left|\arg\left(\lambda f_L\left(x,x^*\right)\right)\right| \\ &= \inf_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]}\ \sup_{x^*\in S_1(x,\eta)} \sin\frac{1}{2}\left|\vartheta+\arg\left(f_L\left(x,x^*\right)\right)\right| = \sin\frac{1}{2}\beta_L\left(x,\eta\right). \end{aligned} \tag{8.334}$$

Moreover,

$$\begin{aligned} \delta_L(\eta) &= \inf_{x\in D(L),\,\|x\|=1} \max\left(\frac{1-q_L(x,\eta)}{1+q_L(x,\eta)},\right. \\ &\qquad \left. \inf_{\vartheta\in\left[-\frac{1}{2}\pi,\frac{1}{2}\pi\right]}\ \sup_{x^*\in S_1(x,\eta)} \sin\frac{1}{2}\left|\vartheta+\arg\left(f_L\left(x,x^*\right)\right)\right|\right) \\ &= \inf_{x\in D(L),\,\|x\|=1} \max\left(\frac{1-q_L(x,\eta)}{1+q_L(x,\eta)}, \sin\frac{1}{2}\beta_L\left(x,\eta\right)\right), \end{aligned} \tag{8.335}$$

and the following inequalities hold for all $x\in D(L)$ and for $\lambda\in\mathbb{C}$, $\Re\lambda\geqslant 0$:

$$\begin{aligned} \frac{1}{2}\left(1-\eta\right)\delta_L(\eta)\eta\left\|x\right\| &\leqslant \frac{1}{2}\left(1-\eta\right)\left\|x-\lambda Lx\right\| \\ &\leqslant \sup_{x^*\in S_1(x,\eta)} \left|\left\langle x-\lambda Lx, x^*\right\rangle\right| \leqslant \left\|x-\lambda Lx\right\|. \end{aligned} \tag{8.336}$$

In addition, $\delta_L(\eta)>0$ if and only if there exists $\delta>0$ such that for all $x\in D(L)$ the following inequality holds:

$$\max\left(\frac{1-q_L(x,\eta)}{1+q_L(x,\eta)}, \sin\frac{1}{2}\beta_L\left(x,\eta\right)\right) \geqslant \delta. \tag{8.337}$$

Proof. The (in-)equalities in (8.331) are easy consequences of (8.321). The equality in (8.333) is an exercise on inequalities, and so is the first equality in (8.334). The second equality in (8.334) follows from (8.327) in Lemma 8.8. The equalities (8.333) and (8.334) yield the equalities in (8.335).

Next let $x\in D(L)$ and $\lambda\in\mathbb{C}$ be such that $\Re\lambda\geqslant 0$. The second inequality in (8.336) is trivial and so is the first one when $\eta=1$. So assume that $0<\eta<1$. Choose $x_0^*\in X^*$ in such a way that $\left|\left\langle x, x_0^*\right\rangle\right| = \frac{1}{2}\left(1+\eta\right)\left\|x\right\|$ and $\left\|x_0^*\right\| = \frac{1}{2}\left(1+\eta\right)$. By the Hahn-Banach theorem such a linear functional exists. If $y^*\in X^*$ is such that $\left\|y^*\right\|\leqslant\frac{1}{2}\left(1-\eta\right)$, then for $\vartheta\in\left[-\pi,\pi\right]$ we have

$$\left\|e^{i\vartheta}x_0^*+y^*\right\| \leqslant \left\|x_0^*\right\|+\left\|y^*\right\| \leqslant \frac{1}{2}\left(1+\eta\right)+\frac{1}{2}\left(1-\eta\right) = 1. \tag{8.338}$$

In addition, again for $\vartheta \in [-\pi,\pi]$ and $y^* \in X^*$ with $\|y^*\| \leqslant \frac{1}{2}(1-\eta)$, we have

$$\left|\left\langle x, e^{i\vartheta}x_0^* + y^*\right\rangle\right| \geqslant |\langle x, x_0^*\rangle| - |\langle x, y^*\rangle| \geqslant \frac{1}{2}(1+\eta)\|x\| - \frac{1}{2}(1-\eta)\|x\| = \eta\|x\|. \tag{8.339}$$

From (8.338) and (8.339) it follows that all vectors of the form $e^{i\vartheta}x_0 + y^*$, $\vartheta \in [-\pi,\pi]$, $\|y^*\| \leqslant \frac{1}{2}(1-\eta)$, belong to the set $S_1(x,\eta_L)$. Then we have

$$\begin{aligned}&\sup_{x^*\in S_1(1,\eta)} |\langle x - \lambda Lx, x^*\rangle| \\ &\geqslant \sup_{\vartheta\in[-\pi,\pi]}\ \sup_{\|y^*\|\leqslant\frac{1}{2}(1-\eta)} \Re\left\langle x - \lambda Lx, e^{i\vartheta}x_0^* + y^*\right\rangle\end{aligned}$$

(by the right choice of ϑ)

$$\geqslant \sup_{\|y^*\|\leqslant\frac{1}{2}(1-\eta)} \Re\langle x - \lambda Lx, y^*\rangle = \frac{1}{2}(1-\eta)\|x - \lambda Lx\|. \tag{8.340}$$

The inequality in (8.340) completes the proof of (8.336). Since the assertion in (8.337) is trivial this completes the proof of Lemma 8.9. □

Proof. [Proof of Theorem 8.10.] The equivalence of the assertions (i) and (iv) is a standard result in the theory of analytic semigroups: see e.g. [Van Casteren (1985)], page 84, or [Pazy (1983a)] Theorem 5.2 and formula (5.16). Another thorough discussion can be found in Chapter II section 4 of [Engel and Nagel (2000)]. The equivalence of the assertions (ii) and (iii) is a consequence of the inequalities (8.325) in Lemma 8.7. The implication (ii) $\Longrightarrow$ (iii) is follows from inequality (8.335) in Lemma 8.9.

Finally, the proof of Theorem 8.10 is completed by showing the implication (iii) $\Longrightarrow$ (iv). To this end put

$$\delta = \inf_{x\in D(L)}\ \inf_{\Re\lambda>0}\ \sup_{x^*\in S_1(x,x^*)} \left|1 - \frac{\lambda\langle Lx, x^*\rangle}{\langle x, x^*\rangle}\right|.$$

Then by (8.323) in (iv) $\delta > 0$, and from the first inequality in (8.336) in Lemma 8.9 it follows that

$$\delta\eta\|x\| \leqslant \|x - \lambda Lx\|, \quad x \in D(L), \quad \Re\lambda > 0. \tag{8.341}$$

The inequality in (8.341) is equivalent to (8.324) with $C = \dfrac{1}{\delta\eta}$ and $\dfrac{1}{\lambda}$ instead of λ. Therefore the proof of Theorem 8.10 is now complete. □

In the following proposition we prove a triviality result. The following characterization of the zero operator does not seem to be known.

Proposition 8.9. *Let L be a closed linear operator with dense domain $D(L)$ in a Banach space $(X, \|\cdot\|)$ with dual $(X^*, \|\cdot\|)$. Put*

$$S_1(x,\eta) = \{x^* \in X^*, \|x^*\| \leqslant 1, |\langle x, x^*\rangle| \geqslant \eta \|x\|\}, \quad x \in X, \quad 0 < \eta < 1. \tag{8.342}$$

Then the following assertions are equivalent:

(i) The operator L is trivial: $L = 0$;

(ii) There exists η in the open interval $(0,1)$ such that the following inequality holds:

$$\inf_{x \in D(L), \|x\|=1} \inf_{\lambda \geqslant 0} \sup_{x^* \in S_1(x,\eta)} \left| 1 - \frac{|\langle \lambda Lx, x^*\rangle|}{|\langle x, x^*\rangle|} \right| > 0, \tag{8.343}$$

and there exists $\lambda \in \mathbb{C}$, $\lambda \neq 0$ such that $R(\lambda I - L) = X$.

From the proof of Lemma 8.9 it follows that (8.343) is equivalent to (see (8.337)):

$$\inf_{x \in D(L)} \frac{1 - q_L(x,\eta)}{1 + q_L(x,\eta)} > 0. \tag{8.344}$$

Proof. The implication (i) $\Longrightarrow$ (ii) being trivial, we only consider the implication (ii) $\Longrightarrow$ (i). Consider the subset $\Lambda := \{\lambda \in \mathbb{C}\backslash\{0\} : (\lambda I - L) D(L) = X\}$. By assumption the set $\Lambda \neq \varnothing$. Let $\delta > 0$ be a strictly positive lower bound of the expression in (8.343). Then for $\lambda \in \mathbb{C}$, $x \in D(L)$, there exists $x^* \in S_1(x,\eta)$ such that we have:

$$\begin{aligned} \|x - \lambda Lx\| &\geqslant |\langle x, x^*\rangle - \langle \lambda Lx, x^*\rangle| = \left| 1 - \frac{\langle \lambda Lx, x^*\rangle}{\langle x, x^*\rangle} \right| |\langle x, x^*\rangle| \\ &\geqslant \left| 1 - \frac{|\langle \lambda Lx, x^*\rangle|}{|\langle x, x^*\rangle|} \right| |\langle x, x^*\rangle| \geqslant \delta\eta \|x\|. \end{aligned} \tag{8.345}$$

From (8.345) it follows that

$$\|\lambda x - Lx\| \geqslant |\lambda| \eta\delta \|x\|, \quad \text{for all } x \in D(L) \text{ and all } \lambda \in \mathbb{C}. \tag{8.346}$$

Let $\lambda_0 \neq 0$ be such that $(\lambda_0 I - L) D(L) = X$, and define the operator $R(\lambda_0) : X \to X$ by $R(\lambda_0)(\lambda_0 I - L)x = x$, $x \in D(L)$. Then by (8.346) for $\lambda = \lambda_0$ we see $|\lambda_0| \|R(\lambda_0)\| \leqslant \dfrac{1}{\delta\eta}$. For $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < |\lambda_0| \eta\delta$ we define the operator $R(\lambda) : X \to X$ by

$$R(\lambda)x = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k R(\lambda_0)^{k+1} x.$$

Then $(\lambda I - L) R(\lambda)x = x$ for all $x \in X$, and $R(\lambda)(\lambda I - L)x = x$ for all $x \in D(L)$. In other words: $R(\lambda) = (\lambda I - L)^{-1}$. It follows that Λ is an open subset of $\mathbb{C}\backslash\{0\}$. Next let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in Λ which converges to $\lambda \in \Lambda$. Then for large enough n we have $|\lambda_n - \lambda| < |\lambda_n| \,\delta\eta$. Since $\lambda_n \in \Lambda$, the above argument with λ_n instead of λ_0 shows that $R(\lambda I - L) = X$. It follows that Λ is closed in $\mathbb{C}\backslash\{0\}$ as well. Consequently, $\Lambda = \mathbb{C}\backslash\{0\}$. So that for every $\lambda \in \mathbb{C}$, $\lambda \neq 0$, we have $(\lambda I - L)\, D(L) = X$. From (8.346) it follows that the family of operators $\left\{(I-\lambda L)^{-1} : \lambda \in \mathbb{C}\backslash\{0\}\right\}$ is uniformly bounded: $\left\|(I-\lambda L)^{-1}\right\| \leqslant \dfrac{1}{\delta\eta}$. Next fix $x \in X$ and $x^* \in X^*$. As a consequence the function $f_{x,x^*} : \lambda \mapsto \left\langle (I-\lambda L)^{-1} x, x^*\right\rangle$ is a bounded holomorphic function on $\mathbb{C}\backslash\{0\}$. By the classical theory about holomorphic functions it follows that the function f_{x,x^*} extends to a bounded holomorphic function on $\mathbb{C}$. By Liouville's theorem this function is constant. All this means that:

$$f_{x,x^*}(\lambda) = \left\langle (I-\lambda L)^{-1} x, x^*\right\rangle = \lim_{\lambda\to 0} \left\langle (I-\lambda L)^{-1} x, x^*\right\rangle, \quad \lambda \in \mathbb{C}. \tag{8.347}$$

Next we identify the limit in (8.347). To this end we first assume that $x \in D(L)$. Then we have

$$\begin{aligned}\left\|(I-\lambda L)^{-1} x - x\right\| &\leqslant |\lambda| \left\|(I-\lambda L)^{-1} Lx\right\| \leqslant |\lambda| \left\|(I-\lambda L)^{-1}\right\| \|Lx\| \\ &\leqslant \frac{|\lambda|}{\eta\delta} \|Lx\|,\end{aligned} \tag{8.348}$$

so that $\lim_{\lambda\to 0} f_{x,x^*}(\lambda) = \langle x, x^*\rangle$, $x \in D(L)$, $x^* \in X^*$. Since

$$\lim_{\lambda\to 0} \left\langle (I-\lambda L)^{-1} x, x^*\right\rangle = \langle x, x^*\rangle \quad \text{for all } x \in D(L),$$

and the family operators $\left\{(I-\lambda L)^{-1} : \lambda \in \mathbb{C}\backslash\{0\}\right\}$ is uniformly bounded (or equi-continuous), it follows that for every x in the closure of $D(L)$ and every $x^* \in X^*$ the function $\lambda \mapsto \left\langle (I-\lambda L)^{-1} x, x^*\right\rangle$ equals the constant $\langle x, x^*\rangle$. So, since by assumption $D(L)$ is dense in X we infer $\left\langle (I-\lambda L)^{-1} x, x^*\right\rangle = \langle x, x^*\rangle$ for all $\lambda \in \mathbb{C}$, $x \in X$, and $x^* \in X^*$. As a consequence we see that $(I-\lambda L)^{-1} = I$ for all $\lambda \in \mathbb{C}$. Then $Lx = 0$ for $x \in D(L)$. Since L is closed with dense domain it necessarily follows that $L = 0$.

This concludes the proof of Proposition 8.9. □

8.7 A version of the Bismut-Elworthy formula

In this section we want to present a version of the Bismut-Elworthy formula for derivatives of Feller propagators applied to a bounded continuous function. In the infinite-dimensional setting we introduce the following Feller propagator (compare with (1.138)):

$$Q(\tau,t)f(x) = \mathbb{E}_{\tau,x}\left[f\left(X(t)\right)\right] = \mathbb{E}\left[f\left(X^{\tau,x}(t)\right)\right] \tag{8.349}$$

where $X(t) = X^{\tau,x}(t)$ is a unique weak solution to the equation (compare with (1.23) and with (1.139))

$$X(t) = x + \int_\tau^t b\left(s, X(s)\right)\, ds + \int_\tau^t \sigma\left(s, X(s)\right)\, dW_H(s), \quad t \geqslant \tau. \tag{8.350}$$

where $X(t) = X^{\tau,x}(t)$ is a unique weak solution to the equation (compare with (1.23) and with (1.139)). Like in Chapter 1 we assume that the process $t \mapsto W_H(t)$ is a cylindrical Brownian motion, that $\sigma(s,x) : H \to E$ is a family of linear operator from the Hilbert H to the Banach space, and that $b(s,x)$ takes its values in the Banach space E. In addition to the stochastic differential equation satisfied by the E-valued process $t \mapsto X^{\tau,x}(t)$, $t \geqslant \tau$, we consider the corresponding flow $F : (t,x) \mapsto X^{\tau,x}(t)$, $t \geqslant \tau$, and the corresponding velocity process $t \mapsto V^{\tau,v}(t)$, $t \geqslant \tau$, defined by $V^{\tau,v}(t) = \langle v, DF(t,\cdot)\rangle$, $v \in E$, and $t \geqslant \tau$. The velocity process satisfies the following stochastic integral equation:

$$\begin{aligned} V^{\tau,v}(t) = v &+ \int_\tau^t D\sigma\left(s,\cdot\right)\left(X^{\tau,x}(s)\right)\left(V^{\tau,v}(s)\right)\, dW_H(s) \\ &+ \int_\tau^t Db\left(s,\cdot\right)\left(X^{\tau,x}(s)\right)\left(V^{\tau,v}(s)\right)\, ds, \end{aligned} \tag{8.351}$$

where $t \geqslant \tau$, $v \in E$, and $x \in E$. We also introduce the propagator $\delta Q\left(\tau,t\right)$, $0 \leqslant \tau \leqslant t \leqslant T$, by

$$\left\langle v, \delta Q\left(\tau,t\right)\varphi(x)\right\rangle = \mathbb{E}\left[\left\langle V^{\tau,v}(t), \varphi\left(X^{\tau,x}(t)\right)\right\rangle\right]. \tag{8.352}$$

Indeed, uniqueness of solutions to equation (8.350) and (8.351) implies the propagator property of the family $\delta Q\left(\tau,t\right)$, $0 \leqslant \tau \leqslant t \leqslant T$. More precisely, for $\rho \leqslant \rho' \leqslant s$ we have:

$$\begin{aligned} &\left\langle v, \delta Q\left(\rho,\rho'\right)\delta Q\left(\rho',s\right)\varphi(x)\right\rangle \\ &= \mathbb{E}\left[\left\langle V^{\rho,v}\left(\rho'\right), \delta Q\left(\rho',s\right)\varphi\left(X^{\rho,v}\left(\rho'\right)\right)\right\rangle\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\left\langle V^{\rho',V^{\rho,v}(\rho')}\left(s\right), \varphi\left(X^{\rho',X^{\rho,v}(\rho')}\left(s\right)\right)\right\rangle\right]\right] \end{aligned}$$

(uniqueness of solutions)

$$= \mathbb{E}\left[\mathbb{E}\left[\left\langle V^{\rho,v}\left(s\right), \varphi\left(X^{\rho,v}\left(s\right)\right)\right\rangle\right]\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left\langle V^{\rho,v}(s), \varphi\left(X^{\rho,v}(s)\right)\right\rangle\right] \\
&= \left\langle v, \delta Q(\rho, s)(x)\right\rangle . \qquad (8.353)
\end{aligned}
$$

In addition, for $0 \leqslant \rho \leqslant s \leqslant T$, x, $v \in E$, we have

$$
\langle v, DQ(\rho, s)\varphi(x)\rangle = \langle v, \delta Q(\rho, s) D\varphi(x)\rangle . \qquad (8.354)
$$

For this equality we refer to the literature: see [Li (1994)]. Next we apply Itô's lemma to the process $s \mapsto Q(s,t)\varphi\left(X^{\tau,x}(s)\right) - Q(\tau,t)\varphi(x)$ to obtain:

$$
\begin{aligned}
&Q(s,t)\varphi\left(X^{\tau,x}(s)\right) - Q(\tau,t)\varphi(x) \\
&= Q(s,t)\varphi\left(X^{\tau,x}(s)\right) - Q(\tau,t)\varphi\left(X^{\tau,x}(\tau)\right) \\
&= -\int_\tau^s L(\rho)Q(\rho,t)\varphi\left(X^{\tau,x}(\rho)\right) d\rho \\
&\quad + \int_\tau^s \left\langle dX^{\tau,x}(\rho), DQ(\rho,t)\left(X^{\tau,x}(\rho)\right)\right\rangle \\
&\quad + \frac{1}{2}\int_\tau^s \operatorname{Tr}\left(\sigma\left(\rho, X^{\tau,x}(\rho)\right)^* D^2 Q(\rho,t)\varphi\sigma\left(\rho, X^{\tau,x}(\rho)\right)\right) d\rho \\
&= \int_\tau^s \left\langle \sigma\left(\rho, X^{\tau,x}(\rho)\right) dW_H(\rho), DQ(\rho,t)\varphi\left(X^{\tau,x}(\rho)\right)\right\rangle . \qquad (8.355)
\end{aligned}
$$

Notice that the equality

$$
\frac{\partial}{\partial s}Q(s,t)\varphi(x) = -L(s)Q(s,t)f(x) \qquad (8.356)
$$

is a consequence of the Theorem 1.16. The reader should compare the equality in (8.356) with the equalities in (5.55). Here the operator $L(s)$ is the same as the one in (1.143), i.e.

$$
L(s)f(s,x) = \langle b(t,x), Df(t,x)\rangle + \frac{1}{2}\operatorname{Tr}\left(\sigma(s,x)^* D^2 f(s,x)\sigma(s,x)\right). \qquad (8.357)
$$

Let $s \uparrow t$ in both sides of (8.355) to obtain:

$$
\begin{aligned}
&\varphi\left(X^{\tau,x}(t)\right) - Q(\tau,t)\varphi(x) \\
&= \int_\tau^t \left\langle \sigma\left(\rho, X^{\tau,x}(\rho)\right) dW_H(\rho), DQ(\rho,t)\varphi\left(X^{\tau,x}(\rho)\right)\right\rangle . \qquad (8.358)
\end{aligned}
$$

Next we assume that the stochastic integral in the right-hand side of (8.358) is a martingale. Then we calculate:

$$
\begin{aligned}
&\mathbb{E}\left[\varphi\left(X^{\tau,x}(t)\right)\int_\tau^t \left\langle dW_H(\rho), \sigma\left(\rho, X^{\tau,x}(\rho)\right)^{-1} V^{\tau,v}(\rho)\right\rangle_H\right] \\
&= \mathbb{E}\left[\left(\varphi\left(X^{\tau,x}(t)\right) - Q(\tau,t)\varphi(x)\right)\int_\tau^t \left\langle dW_H(\rho), \sigma\left(\rho, X^{\tau,x}(\rho)\right)^{-1} V^{\tau,v}(\rho)\right\rangle_H\right]
\end{aligned}
$$

$$= \mathbb{E}\left[\int_\tau^t \langle \sigma\left(\rho, X^{\tau,x}(\rho)\right) dW_H(\rho), DQ\left(\rho,t\right)\varphi\left(X^{\tau,x}(\rho)\right)\rangle\right.$$
$$\left.\times \int_\tau^t \left\langle dW_H(\rho), \sigma\left(\rho, X^{\tau,x}(\rho)\right)^{-1} V^{\tau,v}(\rho)\right\rangle_H\right]$$
$$= \mathbb{E}\left[\int_\tau^t \left\langle \sigma\left(\rho, X^{\tau,x}(\rho)\right)^{-1} V^{\tau,v}(\rho), \sigma\left(\rho, X^{\tau,x}(\rho)\right)^* DQ\left(\rho,t\right)\varphi\left(X^{\tau,x}(\rho)\right)\right\rangle_H\right]$$
$$= \mathbb{E}\left[\int_\tau^t \langle V^{\tau,v}(\rho), DQ\left(\rho,t\right)\varphi\left(X^{\tau,x}(\rho)\right)\rangle\, d\rho\right]. \tag{8.359}$$

From (8.354) it follows that the expression in (8.359) can be rewritten as

$$\mathbb{E}\left[\int_\tau^t \langle V^{\tau,v}(\rho), DQ\left(\rho,t\right)\varphi\left(X^{\tau,x}(\rho)\right)\rangle\, d\rho\right]$$
$$= \mathbb{E}\left[\int_\tau^t \langle V^{\tau,v}(\rho), \delta Q\left(\rho,t\right)D\varphi\left(X^{\tau,x}(\rho)\right)\rangle\, d\rho\right]$$

(definition of the operator $\delta Q(\tau,\rho)$ together with Fubini's theorem)

$$= \int_\tau^t \langle v, \delta Q\left(\tau,\rho\right)\delta Q\left(\rho,t\right)D\varphi(x)\rangle\, d\rho$$

(propagator property (8.353))

$$= \int_\tau^t \langle v, \delta Q\left(\tau,t\right)D\varphi(x)\rangle\, d\rho$$
$$= (t-\tau)\langle v, DQ\left(\tau,t\right)\varphi(x)\rangle. \tag{8.360}$$

In the final equality in (8.360) we again made an appeal to (8.354). From (8.359) and (8.360) we deduce:

$$\mathbb{E}\left[\varphi\left(X^{\tau,x}(t)\right)\int_\tau^t \left\langle dW_H(\rho), \sigma\left(\rho, X^{\tau,x}(\rho)\right)^{-1} V^{\tau,v}(\rho)\right\rangle_H\right]$$
$$= (t-\tau)\langle v, DQ\left(\tau,t\right)\varphi(x)\rangle. \tag{8.361}$$

Although the derivation of the formula in (8.361) was not rigorous, we presented it because of its importance. The invertibility of the operators $\sigma(s,x)$ can be relaxed. The real requirement is that the velocity process $t \mapsto V^{\tau,v}(t)$ is such that it belongs to the range of $\sigma\left(t, X^{\tau,x}(t)\right)$: $V^{\tau,v}(t) = \sigma\left(t, X^{\tau,x}(t)\right)\widetilde{V}^{\tau,v}(t)$ where $t \mapsto \widetilde{V}^{\tau,v}(t)$ is an adapted H-valued process such that $\int_\tau^t \mathbb{E}\left[\left\|\widetilde{V}^{\tau,v}(s)\right\|_H^2\right] ds < \infty$.

Definition 8.9. The equality in (8.361) is known as the Bismut-Elworthy formula.

The Bismut-Elworthy formula has many applications. There are also versions in the context of Brownian motion on a manifold. Versions of the Bismut-Elworthy formula with higher order derivatives exist and can be used to prove that certain Feller type semigroups are analytic: see e.g. [Cerrai (2001)] Chapter 3 and Chapter 6. For a formulation and and a proof in the infinite-dimensional context the reader is referred to [Da Prato *et al.* (1995)]. Proofs for the finite-dimensional case can be found in [Bismut (1981a, 1984)] and in [Elworthy and Li (1994)]. The reader is also referred to [Li (1994)]. For an application of the Bismut-Elworthy formula to Backward Stochastic Differential Equations in control theory see e.g. [Fuhrman and Tessitore (2002, 2004)].

Chapter 9

Coupling methods and Sobolev type inequalities

In this chapter we begin with a discussion of a coupling method by Chen and Wang. We want to establish a spectral gap related to solutions of stochastic differential equations: see Theorem 9.1. In addition we want to include results which do not depend on the matrix $\sigma(t,x)$ (diffusion coefficient) which is such that the matrix $a(t,x) = \sigma(t,x)\sigma(t,x)^*$ is positive-definite. We have a Poincaré inequality in mind: see Proposition 9.10, and Definition 9.15. Related inequalities are (tight) logarithmic Sobolev inequalities: see Definition 9.17, and Proposition 9.11. Another feature of this chapter is the use of the first iterated squared gradient operator, and the abstract Hessian: see the equalities in (9.224), (9.235), and (9.236). In Theorem 9.20 a relationship is established between a spectral gap and an iterated squared gradient inequality of the form (9.226).

9.1 Coupling methods

In this section we want to apply a coupling method to prove the following theorem, which is due to Chen and Wang: see [Chen and Wang (1997)] Theorem 4.13. The operator $L = (L(t))_{t\geqslant 0}$ is of the form:

$$L(t)f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,x)\frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(t,x)\frac{\partial f(x)}{\partial x_i}. \tag{9.1}$$

The matrix $a(t,x) = (a_{i,j}(t,x))_{i,j=1}^{d}$ is supposed to be positive definite, The functions $x \mapsto a_{i,j}(t,x)$, $t \geqslant 0$, belong to $C^2\left(\mathbb{R}^d\right)$, and the functions $(t,x) \mapsto a_{i,j}(t,x)$ are continuous. In addition, $b(t,x)$ is of the form

$$b_i(t,x) = \frac{1}{2}\sum_{j=1}^{d}\left(a_{i,j}(t,x)\frac{\partial V(t,x)}{\partial x_j} + \frac{\partial a_{i,j}(t,x)}{\partial x_j}\right). \tag{9.2}$$

Here for every $t \geqslant 0$ the function $x \mapsto V(t,x)$ is a member of $C^2\left(\mathbb{R}^d\right)$ and has the property that $Z(t) := \int e^{V(t,x)}dx < \infty$; moreover, the function $(t,x) \mapsto V(t,x)$ is continuous on $[0,\infty)\times\mathbb{R}^d$. Let μ_t be the probability measure with density $Z(t)^{-1}e^{V(t,x)}$ with respect to the d-dimensional Lebesgue measure. Then μ_t is an invariant measure for $L(t)$ and the semigroup $e^{sL(t)}$ generated by $L(t)$, provided such a semigroup exists. Let us check this. Let $L(t)^*$ be the (formal) adjoint of $L(t)$. We notice

$$
\begin{aligned}
&2L(t)^*f(x)\\
&= \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i\partial x_j}\left(a_{i,j}(t,x)f(x)\right) + 2\sum_{i=1}^d \frac{\partial}{\partial x_i}\left(b_i(t,x)f_i(x)\right)\\
&= \sum_{i,j=1}^d a_{i,j}(t,x)\frac{\partial^2 f(x)}{\partial x_i\partial x_j} + 2\left(\sum_{i,j=1}^d \frac{\partial a_{i,j}(t,x)}{\partial x_i}\frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^d b_i(t,x)\frac{\partial f(x)}{\partial x_i}\right)\\
&\quad + \left(\sum_{i,j=1}^d \frac{\partial^2 a_{i,j}(t,x)}{\partial x_i\partial x_j} - 2\sum_{i=1}^d \frac{\partial b_i(t,x)}{\partial x_i}\right)f(x),
\end{aligned}
$$

and hence

$$
\begin{aligned}
&2L(t)^*\left(e^{V(t,x)}\right)\\
&= e^{V(t,x)}\sum_{i,j=1}^d a_{i,j}(t,x)\frac{\partial^2 V(t,x)}{\partial x_i\partial x_j} + e^{V(x)}\sum_{i,j=1}^d a_{i,j}(t,x)\frac{\partial V(t,x)}{\partial x_i}\frac{\partial V(t,x)}{\partial x_j}\\
&\quad + 2e^{V(t,x)}\left(\sum_{i,j=1}^d \frac{\partial a_{i,j}(t,x)}{\partial x_i}\frac{\partial V(t,x)}{\partial x_j} - \sum_{i=1}^d b_i(x)\frac{\partial V(t,x)}{\partial x_i}\right)\\
&\quad + e^{V(t,x)}\left(\sum_{i,j=1}^d \frac{\partial^2 a_{i,j}(t,x)}{\partial x_i\partial x_j} - 2\sum_{i=1}^d \frac{\partial b_i(t,x)}{\partial x_i}\right).
\end{aligned}
\tag{9.3}
$$

From (9.3) in conjunction with (9.2) we see $L(t)^*\left(e^V(t,\cdot\right) = 0$, and consequently,

$$
Z(t)\int L(t)f\,d\mu = \int \left(L(t)f(x)\right)\,e^{V(t,x)}dx = \int f(x)\,L^*e^{V(t,\cdot)}(x)dx = 0.
$$

Note that we used the symmetry of the matrix $a(t,x) = \left(a_{i,j}(t,x)\right)_{i,j=1}^d$.

In the following theorem 9.1 we consider the time-homogeneous case, i.e. the operator L does not depend on the time t. It is not clear how to get such a result in the time-dependent case. It is assumed that the coefficients $a(x)$ and $b(x)$ are such that the martingale problem is uniquely solvable for for L, and that the corresponding Markov process is irreducible in the sense that the transition probability measures $B \mapsto P(t,x,B)$, $B \in \mathcal{B}_{\mathbb{R}^d}$, $t > 0$, $x \in \mathbb{R}^d$, are equivalent, i.e. all of them have the same null-sets. In fact this is a stronger notion than the standard notion of irreducibility. However, if all functions of the form $(t,x) \mapsto P(t,x,B)$, $B \in \mathcal{E}$, are continuous, then these two notions coincide: see Lemma 9.1 below.

Definition 9.1. A time-homogeneous Markov process with state space E and probability transition function $P(t,x,\cdot)$ is called irreducible if $P(t,x,U) > 0$ for all $(t,x) \in (0,\infty) \times E$ and all non-empty open subsets U of E.

Lemma 9.1. *Let $(t,x,B) \mapsto P(t,x,B)$ be a transition probability function with the property that for every $(t,B) \in (0,\infty) \times \mathcal{E}$ the function $x \mapsto P(t,x,B)$ is lower semi-continuous. Then all measures $P(t,x,\cdot)$, $(t,x) \in (0,\infty) \times E$, are equivalent if and only if, for every non-void open subset U and every $(t,x) \in (0,\infty) \times E$, $P(t,x,U) > 0$.*

Proof. First suppose that for every non-void open subset U the quantity $P(t,x,U)$ is strictly positive for all pairs $(t,x) \in (0,\infty) \times E$. Let $(t_0,x_0,B) \in (0,\infty) \times E \times \mathcal{E}$ be such that $P(t_0,x_0,B) = 0$. Fix $s \in (0,t_0)$. Then $0 = P(t_0,x_0,B) = \int P(s,y,B)\, P(t_0 - s,x,dy)$, and hence the function $y \mapsto P(s,y,B)$ is $P(t_0 - s,x,\cdot)$-almost everywhere zero. Assume that there exists $y_0 \in E$ and $\varepsilon > 0$ such that $P(s,y_0,B) > \varepsilon > 0$, and put $U_\varepsilon = \{y \in E : P(s,y,B) > \varepsilon\}$. Then U_ε is a non-void open subset of E. Moreover,

$$
\begin{aligned}
0 = P(t_0,x_0,B) &= \int P(s,y,B)\, P(t_0 - s,x_0,dy) \\
&\geqslant \int_{U_\varepsilon} P(s,y,B)\, P(t_0 - s,x_0,dy) \geqslant \varepsilon P(t_0 - s,x_0,U_\varepsilon) > 0 \qquad (9.4)
\end{aligned}
$$

where in the final step of (9.4) we used our initial hypothesis. Anyway, our assumption that $P(t_0,x_0,B) = 0$ leads to a contradiction with the assertion that all transition probabilities of the form $P(t,x,U)$, $(t,x) \in (0,\infty) \times E$, U open, $U \neq \varnothing$, are strictly positive.

Next assume that all measures $P(t,x,\cdot)$, $(t,x) \in (0,\infty) \times E$, are equivalent, and assume that for some non-empty open subset U of E the quantity $P(t,x,U) = 0$. Then, by our assumption we may and will assume that $x \in U$, and that we may choose $t > 0$ as close to zero as we please. By the normality we have $1 = \lim_{t\downarrow 0} P(t,x,U) = 0$. Again we end up with a contradiction.

This completes the proof of Lemma 9.1. □

Let $(X(t), \mathbb{P}_x)$ be a Markov process with the Feller property. Among other things this implies that $\lim_{t\downarrow 0} \mathbb{P}_x [X(t) \in U] = 1$ for all open subsets U of E, and for all $x \in U$. If all probability measures $B \mapsto P(t,x,B) = \mathbb{P}_x [X(t) \in B]$, $B \in \mathcal{E}$, have the same null-sets, then the corresponding time-homogeneous Markov process (with the Feller property) is irreducible in the sense of Definition 9.1. To this end, assume that there exists a non-void open subset U of E such that $P(t,x,U) = 0$. Since all measures $P(t,x,\cdot)$, $(t,x) \in (0,\infty)\times E$ have the same negligible sets, we may and will assume that $x \in U$ and t is as close to zero as we please. Since $\lim_{t\downarrow 0} \mathbb{P}_x [X(t) \in U] = 1$, this leads to a contradiction, and hence our Markov process is irreducible, provided all transition probability measures $P(t,x,\cdot)$ have the same null-sets. The proof of the following theorem will be given at the end of §9.3.

Theorem 9.1. *Suppose that there exists $\overline{a} > 0$ such that $\langle a(x)\xi, \xi\rangle \leqslant \overline{a} |\xi|^2$ for all x, $\xi \in \mathbb{R}^d$. Let $a(x) = \sigma(x)\sigma(x)^*$ and put*

$$-\gamma = \sup_{x\neq y \in \mathbb{R}^d} \frac{\operatorname{Tr}(\sigma(x)-\sigma(y))(\sigma(x)-\sigma(y))^* + 2\langle b(x)-b(y), x-y\rangle}{|x-y|^2}. \tag{9.5}$$

(Here as elsewhere $\operatorname{Tr}(A)$ stands for the trace of the matrix or trace class operator A.) Then the following inequality holds for all globally Lipschitz functions $f : \mathbb{R}^d \to \mathbb{R}$, all $x \in \mathbb{R}^d$, and all $t \geqslant 0$:

$$e^{tL} |f|^2 (x) - \left|e^{tL} f(x)\right|^2 \leqslant \frac{\overline{a}\left(1-e^{-\gamma t}\right)}{\gamma} e^{tL} |\nabla f|^2 (x). \tag{9.6}$$

If $\gamma = 0$, then $\dfrac{1-e^{-\gamma t}}{\gamma}$ is to be interpreted as t.

In the next corollary we write:

$$\lambda_{\min}(a) = \inf\left\{\langle a(x)\xi,\xi\rangle : (x,\xi) \in \mathbb{R}^d \times \mathbb{R}^d, |\xi| = 1\right\}. \tag{9.7}$$

Corollary 9.1. *In addition to the hypotheses in Theorem 9.1 suppose that $\gamma > 0$. Then the diffusion generated by L is mixing in the sense that $\lim_{t\to\infty} \int \left|e^{tL} f\right|^2 d\mu = \left|\int f d\mu\right|^2$, and the spectral gap of L satisfies*

$$\operatorname{gap}(L) \geqslant \gamma \frac{\lambda_{\min}(a)}{\overline{a}}. \tag{9.8}$$

Proof. [Proof of Corollary 9.1.] Let μ be the invariant probability measure corresponding to the generator L. The fact that the diffusion generated by L is ergodic follows from results in [Chen and Wang (2003)]: see Theorem 9.2 below. The mixing property is a consequence of assertion (ii) in

Theorem 9.2. Since

$$e^{tL}|f|^2 - \left|e^{tL}f\right|^2 = \int_0^t e^{(t-s)L}\Gamma_1\left(e^{sL}\overline{f}, e^{sL}f\right)ds, \tag{9.9}$$

we see that

$$\int |f|^2 d\mu - \int \left|e^{tL}f\right|^2 d\mu = \int_0^t \int \Gamma_1\left(e^{sL}\overline{f}, e^{sL}f\right)d\mu\, ds \tag{9.10}$$

where we used the L-invariance of the measure μ several times. From (9.9) it follows that $\lim_{t\to\infty}\int\left|e^{tL}f\right|^2 d\mu$ exists. It is not clear that this limit is equal to $\left|\int f d\mu\right|^2$. The equality in (9.9) is an immediate consequence of equality (9.158) in the proof of Theorem 9.1 below. We wrote

$$\Gamma_1(f,g) = \langle a\nabla f, \nabla g\rangle = \sum_{i,j=1}^d a_{i,j}\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}. \tag{9.11}$$

By taking the limit as $t\to\infty$ in (9.9) we obtain

$$\int |f|^2 d\mu - \left|\int f\, d\mu\right|^2 = \int_0^\infty \int \Gamma_1\left(e^{sL}\overline{f}, e^{sL}f\right)d\mu\, ds. \tag{9.12}$$

The result in Corollary 9.1 is a consequence of (9.6), (9.7), (9.11), and (9.12). □

In the proof of Corollary 9.1 we used a result on ergodicity. The following result can be found as Theorem 4 in [Maslowski and Seidler (1998)]. It is applicable in our situation. For its proof we refer the reader to [Stettner (1994)] and [Seidler (1997)]. A general discussion about this kind of properties can be found in [Maslowski and Seidler (1998)]. For convenience we insert an outline of a proof. We need the following definition: compare with property (a) in Proposition 9.1 below.

Definition 9.2. Let D be a subspace of $C_b(E)$. It is said that D almost separates compact and closed sets, if for every compact subset K and closed subset F such that $K\bigcap F=\varnothing$ there exist a constant $\alpha>0$ and a function $u\in D$ such that $\alpha\leqslant u(x)-u(y)$ for all $x\in K$ and all $y\in F$.

Remark 9.1. If the linear subspace D contains the constant functions, and is closed under taking finite maxima, then D almost separates compact and closed subsets if and only for every closed subset F of E, and every $x\in E\backslash F$ there exists a function $u\in D$ such that $u(x)>\sup_{y\in D}u(y)$. Let F be closed subset of E. First suppose that D almost separates compact subsets not intersecting F. Since a set consisting of one singleton $x\in E\backslash F$ is compact, there exists a function u and a constant $\alpha > 0$ such that $\alpha < u(x) - u(y)$, for all $y \in F$. Then $u(x) > \alpha + \sup_{y \in F} u(y) > \sup_{y \in F} u(y)$. Conversely, let K and F be compact and closed subset of E which do not intersect. Suppose that for every $x \in K$ there exists a function $u_x \in D$ such that $u_x(x) > \sup_{y \in F} u_x(y)$. Then by subtracting the constant $\alpha_x = \sup_{y \in F} u_x(y)$ we see that $v_x := u_x - \alpha_x$ satisfies $v_x(x) > 0 \geqslant \sup_{y \in F} v_x(y)$. By compactness there exist finitely many functions $v_j := v_{x_j}$, $1 \leqslant j \leqslant N$, such that

$$\max_{1 \leqslant j \leqslant N} v_j(x) \geqslant \alpha > 0 \geqslant \sup_{y \in F} \max_{1 \leqslant j \leqslant N} v_j(y), \quad x \in K, \tag{9.13}$$

where $\alpha = \inf_{x \in K} \max_{1 \leqslant j \leqslant N} v_j(x)$, which is strictly positive real number. It follows that $0 < \alpha \leqslant \max_{1 \leqslant j \leqslant N} v_j(x) - \max_{1 \leqslant j \leqslant N} v_j(y)$, $x \in K$, $y \in F$.

Definition 9.3. Consider the Markov process in (9.14) below. Let the family of time-translation operators have the property that $X(s) \circ \vartheta_t = X(s+t)$ $\mathbb{P}_x$-almost surely for all x, and are such that $\vartheta_{s+t} = \vartheta_s \circ \vartheta_t$ for all $s, t \in [0, \infty)$. Its tail or asymptotic σ-field $\mathcal{T}$ is defined by $\mathcal{T} = \bigcap_{t>0} \vartheta_t^{-1} \mathcal{F}$. In fact an event A belongs to $\mathcal{T}$ if and only if for every $t > 0$ there exists an event $A_t \in \mathcal{F}$ such that $A = \vartheta_t^{-1} A_t$, or what amounts to the same $\mathbf{1}_A = \mathbf{1}_{A_t} \circ \vartheta_t$.

Remark 9.2. In fact we may assume that $A_t \in \mathcal{T}$. The reason being that $A = \vartheta_{s+t}^{-1} A_{s+t} = \vartheta_t^{-1} \left(\vartheta_s^{-1} A_{s+t} \right)$, and hence for A_t we may choose $A_t = \bigcap_{s>0} \vartheta_s^{-1} A_{s+t}$.

For more details on the notion of strong Feller property see Definitions 2.5 and 2.16.

Theorem 9.2. *Let*

$$\left\{ (\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in E}, (X(t), t \geqslant 0), (E, \mathcal{E}) \right\} \tag{9.14}$$

be a time-homogeneous Markov process on a Polish space E with a transition probability function $P(t, x, \cdot)$, $t \geqslant 0$, $x \in E$, which is conservative in the sense that $P(t, x, E) = 1$ for all $t \geqslant 0$ and $x \in E$. Assume that the process $X(t)$ is strong Feller in the sense that for all Borel subsets B of E the function $(t, x) \mapsto P(t, x, B)$ is continuous on $(0, \infty) \times E$. In addition, suppose that all measures $B \mapsto P(t, x, B)$, $B \in \mathcal{E}$, $t > 0$, $x \in E$, are equivalent, and that the process has an invariant probability measure μ. In addition suppose that the domain of the generator L of the Markov process almost separates compact and closed subsets. Then the following assertions are true:

(i) For every $f \in L^1(E,\mu)$ and every $x \in E$ the equality

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f\left(X(s)\right) ds = \int_E f d\mu \tag{9.15}$$

holds $\mathbb{P}_x$-almost surely;

(ii) For every $x \in E$ the following equality holds:

$$\lim_{t\to\infty} \mathrm{Var}\left(P\left(t,x,\cdot\right) - \mu\right) = 0. \tag{9.16}$$

In particular, both assertion (i) and (ii) imply that the invariant measure μ is unique.

Remark 9.3. In Theorem 10.12 in Chapter 10 it will be shown that the Markov process (9.14) admits a σ-finite invariant measure provided that this process satisfies the conditions of Theorem 9.2, and that it is topologically recurrent. The Markov process (9.14) is called topologically recurrent if every non-empty open subset is recurrent. In addition, in Corollary 10.5 a condition will be formulated which implies that this invariant measure is in fact finite, and hence may be taken to be a probability measure.

The equality in (9.15) is known as the strong law of large numbers or the pointwise ergodic theorem of Birkhoff. In (9.16) $\mathrm{Var}(\nu)$ stands for the variation norm of the measure ν. The property in (ii) is stronger than the weak and strong mixing property. If the process in (9.14) has property (ii), then it is said to be ergodic. There exist stronger notions of ergodicity: see e.g. [Chen (2005)]. The property in (ii) is closely related to the fact that in the present situation the tail σ-field is trivial. Mixing properties are heavily used in ergodic theory: see e.g. [Meyn and Tweedie (1993b)]. Suppose that there exists a (reference) measure m on $\mathcal{E}$ and a measurable function $(t,x,y) \mapsto p(t,x,y)$, $(t,x,y) \in (0,\infty) \times E \times E$, which is strictly positive such that for every $(t,x,B) \in (0,\infty) \times E \times \mathcal{E}$ the equality $P(t,x,B) = \int p(t,x,y)dm(y)$ holds. Then $P(t,x,A) = 0$ if and only if $m(A) = 0$, and so all measures $P(t,x,\cdot)$ have the same null-sets. For a proof of Theorem 9.2 the reader is referred to the cited literature. We will also include a proof, which is based on work by Seidler [Seidler (1997)]: see Theorem 10.12.

Lemma 9.2 says that property (ii) in Theorem 9.2 is stronger than the strong mixing property, which can be phrased as follows: for every f and $g \in L^2\left(E,\mu\right)$ we have

$$\lim_{t\to\infty} \mathbb{E}_\mu\left[f\left(X(t)\right) g\left(X(0)\right)\right] = \lim_{t\to\infty} \int_E \left(e^{tL} f\right)(x) g(x) d\mu(x) = \int_E f d\mu \int_E g d\mu. \tag{9.17}$$

Here $\mathbb{E}_{\mu}[F] = \int_E \mathbb{E}_x[F]\, d\mu(x)$, $F \in L^{\infty}(\Omega, \mathcal{F})$. Notice that by Cauchy-Schwarz' inequality and by the L-invariance of the probability measure μ we have:

$$\left(\int \left|e^{tL} f(x) g(x)\right| \, d\mu(x)\right)^2 \leqslant \int \left|e^{tL} f(x)\right|^2 d\mu(x) \cdot \int |g(x)|^2 \, d\mu(x)$$
$$\leqslant \int e^{tL} |f|^2 \, d\mu \cdot \int |g|^2 \, d\mu = \int |f|^2 \, d\mu \cdot \int |g|^2 \, d\mu < \infty$$

whenever f, $g \in L^2(E, \mu)$.

Lemma 9.2. *Suppose that μ is an L-invariant probability measure which, for each $x \in E$, satisfies (9.16) in Theorem 9.2. Then*

$$\lim_{t\to\infty} e^{tL} f(x) e^{tL} g(x) = \int f \, d\mu \cdot \int g d\mu \quad \textit{and}$$
$$\lim_{t\to\infty} \int e^{tL} f(x) \cdot e^{tL} g(x) d\mu(x) = \int f(x) \, d\mu(x) \int g(x) \, d\mu(x) \tag{9.18}$$

for all f and $g \in C_b(E)$.

Proof. Let the functions f and g belong to $C_b(E)$. The second equality in (9.18) is a consequence of the first one and the dominated convergence theorem of Lebesgue. The first equality is a consequence of the following equalities and (9.16) in Theorem 9.2:

$$\begin{aligned}
&\left|e^{tL} f(x) \cdot e^{tL} g(x) - \int f(y) \, d\mu(y) \cdot \int g(y) d\mu(y)\right| \\
&\leqslant \left|e^{tL} f(x) - \int f(y) d\mu(y)\right| \cdot \left|e^{tL} g(x)\right| \\
&\quad + \left|\int f(y) d\mu(y)\right| \cdot \left|e^{tL} g(x) - \int g(y) d\mu(y)\right| \\
&= \left|\int f(y) P(t, x, dy) - \int f(y) d\mu(y)\right| \cdot \left|e^{tL} g(x)\right| \\
&\quad + \left|\int f(y) d\mu(y)\right| \cdot \left|\int g(y) P(t, x, dy) - \int g(y) d\mu(y)\right| \\
&\leqslant 2 \|f\|_{\infty} \|g\|_{\infty} \operatorname{Var}(P(t, x, \cdot) - \mu).
\end{aligned} \tag{9.19}$$

The right-hand side of (9.19) together with (9.16) completes the proof of Lemma 9.2. □

In case the Markov process in Theorem 9.2 originates from a Feller-Dynkin semigroup with a locally compact state space, then the following proposition is automatically true. In case we are dealing with a Polish state space, we need the extra condition that the domain of the generator has the property described in property (a) in Proposition 9.1 below. This property says that, up to any $\varepsilon > 0$, the domain of L separates disjoint compact and closed sets. In the locally compact and a strong Markov process originating from a Dynkin-Feller semigroup, it is only required that $C_0(E)$ has this property, which is automatically the case. Since by assumption $P(t,x,E) = 1$ there is no need to consider $E^{\triangle}$: see final assertion in Theorem 2.9.

Proposition 9.1. *Let K be a compact subset of E and let U be an open subset of E such that $K \subset U$. Let τ_{U^c} be the hitting time of $E\backslash U$: $\tau_{U^c} = \inf\{s > 0 : X(s) \in E\backslash U\}$. Assume that the generator L has the following separation property:*

(a) For every $x \in K$ there exist a function $u = u_x \in D(L)$ such that $u(x) > \sup_{y \in U^c} u(y)$.

Then

$$\lim_{t \downarrow 0} \sup_{x \in K} \mathbb{P}_x \left[\tau_{U^c} \leqslant t\right] = 0. \tag{9.20}$$

In Proposition 9.2 below we will give alternative formulations for (9.20).

Proof. Since K, and since the domain of L contains the constant functions there exist finitely many functions $u_j \in D(L)$, $1 \leqslant j \leqslant N$, and a constant $\alpha > 0$ such that

$$0 < \alpha \leqslant \inf_{x \in K} \max_{1 \leqslant j \leqslant N} u_j(x) - \sup_{y \in U^c} \max_{1 \leqslant j \leqslant N} u_j(y). \tag{9.21}$$

To see this the reader is referred to the arguments leading to (9.13). Choose the constant $\alpha > 0$ and the functions $u_j \in D(L)$, $1 \leqslant j \leqslant N$, satisfying (9.21). Then for $x \in K$ and $1 \leqslant j \leqslant N$ we have

$$\begin{aligned}
&\left(u_j(x) - \sup_{y \in U^c} u_j(y)\right) \mathbb{P}_x \left[\tau_{U^c} \leqslant t\right] \\
&\leqslant \mathbb{E}_x \left[u_j\left(X(0)\right) - u_j\left(X\left(\tau_{U^c}\right)\right), \tau_{U^c} \leqslant t\right] \\
&= \mathbb{E}_x \left[u_j\left(X(t)\right) - u_j\left(X\left(\tau_{U^c}\right)\right), \tau_{U^c} \leqslant t\right] \\
&\quad + \mathbb{E}_x \left[u_j\left(X(0)\right) - u_j\left(X(t)\right), \tau_{U^c} \leqslant t\right] \qquad (9.22) \\
&= \mathbb{E}_x \left[u_j\left(X(t)\right) - u_j\left(X\left(\tau_{U^c} \wedge t\right)\right) - \int_{\tau_{U^c} \wedge t}^t L u_j\left(X(s)\right) ds, \tau_{U^c} \wedge t < t\right] \\
&\quad + \mathbb{E}_x \left[\int_{\tau_{U^c}}^t L u_j\left(X(s)\right) ds, \tau_{U^c} \leqslant t\right] + \mathbb{E}_x \left[u\left(X(0)\right) - u\left(X(t)\right), \tau_{U^c} \leqslant t\right]
\end{aligned}$$

$$= \mathbb{E}_x \left[\mathbb{E}_x \left[u_j \left(X(t) \right) - u_j \left(X \left(\tau_{U^c} \wedge t \right) \right) - \int_{\tau_{U^c} \wedge t}^t L u_j \left(X(s) \right) ds \;\Big|\; \mathcal{F}_{\tau_{U^c} \wedge t} \right], \right.$$
$$\left. \tau_{U^c} \wedge t < t \right] + \mathbb{E}_x \left[\int_{\tau_{U^c}}^t L u_j \left(X(s) \right) ds,\, \tau_{U^c} \leqslant t \right]$$
$$+ \mathbb{E}_x \left[u_j \left(X(0) \right) - u_j \left(X(t) \right),\, \tau_{U^c} \leqslant t \right]$$

(Doob's optional sampling theorem)

$$= \mathbb{E}_x \left[\int_{\tau_{U^c}}^t L u_j \left(X(s) \right) ds,\, \tau_{U^c} \leqslant t \right] + \mathbb{E}_x \left[u_j \left(X(0) \right) - u_j \left(X(t) \right),\, \tau_{U^c} \leqslant t \right]$$
$$\leqslant t \sup_{y \in E} L u_j(y) + \mathbb{E}_x \left[\left| u_j(X(0)) - u_j(X(t)) \right| \right]. \tag{9.23}$$

The choice of $\alpha > 0$ together with (9.23) shows:

$$\alpha \sup_{x \in K} \mathbb{P}_x \left[\tau_{U^c} \leqslant t \right]$$
$$\leqslant t \max_{1 \leqslant j \leqslant N} \sup_{y \in E} L u_j(y) + \max_{1 \leqslant j \leqslant N} \sup_{x \in K} \mathbb{E}_x \left[\left| u_j(X(0)) - u_j(X(t)) \right| \right]. \tag{9.24}$$

We also notice the inequalities ($1 \leqslant j \leqslant N$):

$$\left(\mathbb{E}_x \left[\left| u_j(X(0)) - u_j(X(t)) \right| \right] \right)^2$$
$$\leqslant \mathbb{E}_x \left[\left| u_j(X(0)) - u_j(X(t)) \right|^2 \right]$$
$$= 2 u_j(x) \left(u_j(x) - \mathbb{E}_x \left[u_j \left(X(t) \right) \right] \right) + \mathbb{E}_x \left[u_j \left(X(t) \right)^2 \right] - u_j(x)^2$$
$$= 2 u_j(x) \left(u_j(x) - e^{tL} u_j(x) \right) + e^{tL} \left| u_j \right|^2 (x) - u_j(x)^2. \tag{9.25}$$

Since the semigroup $\left\{ e^{tL} : t \geqslant 0 \right\}$ is $\mathcal{T}_\beta$-continuous from (9.25) and (9.24) we infer that

$$\lim_{t \downarrow 0} \sup_{x \in K} \mathbb{E}_x \left[\left| u_j \left(X(0) \right) - u_j \left(X(t) \right) \right| \right] = 0, \quad 1 \leqslant j \leqslant N. \tag{9.26}$$

From (9.23) and (9.26) it follows that

$$\limsup_{t \downarrow 0} \sup_{x \in K} \mathbb{P}_x \left[\tau_{U^c} \leqslant t \right] \leqslant \varepsilon. \tag{9.27}$$

Hence, since in (9.27) $\varepsilon > 0$ is arbitrary, this concludes the proof of Proposition 9.1. □

Remark 9.4. Suppose that in Proposition 9.1 the state space E is second countable and locally compact. In this case there exists a function $u \in C_0(E)$ such that $\mathbf{1}_K \leqslant u \leqslant \mathbf{1}_U$. Then we use the time-homogeneous strong Markov property to rewrite (9.22) as follows:

$$\mathbb{P}_x \left[\tau_{U^c} \leqslant t \right] = \mathbb{E}_x \left[u \left(X(t) \right),\, \tau_{U^c} \leqslant t \right] + \mathbb{E}_x \left[\mathbf{1} - u \left(X(t) \right),\, \tau_{U^c} \leqslant t \right]$$

$$
\begin{aligned}
&= \mathbb{E}_x\left[u\left(X(t)\right) - u\left(X\left(\tau_{U^c}\right)\right), \tau_{U^c} \leqslant t\right] + \mathbb{E}_x\left[u\left(X\left(\tau_{U^c}\right)\right), \tau_{U^c} \leqslant t\right] \\
&\quad + \mathbb{E}_x\left[\mathbf{1} - u\left(X(t)\right), \tau_{U^c} \leqslant t\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_{X(\tau_{U^c})}\left[u\left(X\left(t-\tau_{U^c}\right)\right) - u\left(X\left(0\right)\right)\right], \tau_{U^c} \leqslant t\right] \\
&\quad + \mathbb{E}_x\left[\mathbf{1} - u\left(X(t)\right), \tau_{U^c} \leqslant t\right] \\
&\leqslant \sup_{y \notin U} \sup_{s \in [0,t]} \mathbb{E}_y\left[u\left(X(s)\right) - u(X(0))\right] + \mathbb{E}_x\left[\left|u\left(X(0)\right) - u\left(X(t)\right)\right|\right] \\
&\leqslant \sup_{s \in [0,t]} \sup_{y \in E} \left(e^{sL}u(y) - u(y)\right) + e^{tL}\left|u(x) - u\right|(x).
\end{aligned} \tag{9.28}
$$

Since, uniformly on E, $e^{tL}u - u$ converges to zero when $t \downarrow 0$, the proof of Proposition 9.1 can be finished as in the non-locally compact case.

Remark 9.4 shows that assertion (i) in Proposition (9.2) automatically holds when the state space E is second countable and locally compact. It also holds when the domain of the generator almost separates points and closed sets (see (a) in Proposition 9.1), and when the functions $x \mapsto \mathbb{P}_x\left[\tau_{U^c}\right]$ are continuous on the open subset U: see also item (i) in Proposition 9.3.

Proposition 9.2. *Let d be a metric on E which is compatible with its Polish topology. Then the following assertions are equivalent:*

(i) For every compact subset K and every open subset U of E such that $K \subset U$ the equality in (9.20) holds, i.e. $\lim_{t \downarrow \infty} \sup_{x \in K} \mathbb{P}_x\left[\tau_{U^c} \leqslant t\right] = 0$ where τ_{U^c} stands for the first hitting time of the complement of U, which is also called the first exit time from U.

(ii) For every compact subset K of E and every $\eta > 0$ the following equality holds:

$$\lim_{t \downarrow 0} \sup_{x \in K} \mathbb{P}_x\left[\sup_{0 < s \leqslant t} d\left(X(s), x\right) \geqslant \eta\right] = 0. \tag{9.29}$$

(iii) For every compact subset K and every open subset U of E such that $K \subset U$, and every sequence $(t_n)_{n \in \mathbb{N}} \subset (0, \infty)$ which decreases to 0, there exists a sequence of open subsets $(U_n)_{n \in \mathbb{N}}$ such that $U_n \supset K$, $n \in \mathbb{N}$, and which has the property that

$$\lim_{n \to \infty} \sup_{x \in U_n} \mathbb{P}_x\left[\tau_{U^c} \leqslant t_n\right] = 0. \tag{9.30}$$

(iv) For every compact subset K of E and every $\eta > 0$ and every sequence $(t_n)_{n \in \mathbb{N}} \subset (0, \infty)$ which decreases to 0 there exists a sequence of open subsets $(U_n)_{n \in \mathbb{N}}$ such that $U_n \supset K$, $n \in \mathbb{N}$, and with the property that:

$$\lim_{n \to \infty} \sup_{x \in U_n} \mathbb{P}_x\left[\sup_{0 < s \leqslant t_n} d\left(X(s), x\right) \geqslant \eta\right] = 0. \tag{9.31}$$

The result in Proposition 9.2 resembles the result in Proposition 4.6: see the proof of Lemma 4.2. In [Seidler (1997)] Seidler employs assumption (9.29) to a great extent. Assertions (iii) and (iv) of Proposition 9.2 show that in [Seidler (1997)] hypothesis (A5) is in fact a consequence of (A4). Proposition 2.4 in [Seidler (1997)] then shows that there exists a compact recurrent subset whenever there exists a point $x_0 \in E$ with the property that every open subset containing x_0 is recurrent. For a precise formulation and a proof the reader is referred to Proposition 9.4 and its proof below. See formula (9.20) in Proposition 9.1 how the almost separation property of the generator L of the Markov process in (9.14) implies that the equivalent conditions in Proposition 9.2 are satisfied.

Proof. As already indicated the proof is in the spirit of the proof of Lemma 4.2. Also note that, since $U_n \supset K$, the implications (iii) $\Longrightarrow$ (i) and (iv) $\Longrightarrow$ (ii) are trivially true.

(i) $\Longrightarrow$ (ii). Let K be a compact subset of E. Fix $\eta > 0$, and consider the open subset U defined by $U = \bigcup_{x\in K} \{y \in E : d(y,x) < \eta\}$. Then $K \subset U$, and $U^c = E\backslash U = \bigcap_{x\in K} \{y : d(y,x) \geqslant \eta\}$. It follows that the event $\{\tau_{U^c} < t\}$ is contained in the event $\left\{\sup_{0<s<t} d\left(X(s),x\right) \geqslant \eta\right\}$ for all $x \in K$. Then for $x \in K$ we have

$$\mathbb{P}_x\left[\tau_{U^c} < t\right] \leqslant \mathbb{P}_x\left[\sup_{0<s<t} d\left(X(s),x\right) \geqslant \eta\right]. \tag{9.32}$$

Then (ii) follows from (i) and (9.32).

(ii) $\Longrightarrow$ (i). Let the compact K and the open subset of E be such that $K \subset U$. Then by compactness there exist points $x_1, \dots, x_n$ in K and strictly positive numbers $\eta_1, \dots, \eta_n$ such that

$$K \subset \bigcup_{j=1}^{n} \{y \in E : d\left(y,x_j\right) < \eta_j\} \subset \bigcup_{j=1}^{n} \{y \in E : d\left(y,x_j\right) < 2\eta_j\} \subset U. \tag{9.33}$$

Put $V = \bigcup_{j=1}^{n} \{y \in E : d\left(y,x_j\right) < 2\eta_j\}$. Then

$$U^c \subset V^c = \bigcap_{j=1}^{n} \{y \in E : d\left(y,x_j\right) \geqslant 2\eta_j\}. \tag{9.34}$$

Let $y \in V^c$ and $x \in K$ be arbitrary. Then by (9.33) there exists $j_x \in \{1, \dots, n\}$ such that $d\left(x, x_{j_x}\right) < \eta_{j_x}$. It follows that

$$2\eta_{j_x} \leqslant d\left(y, x_{j_x}\right) \leqslant d\left(y,x\right) + d\left(x, x_{j_x}\right) < d\left(y,x\right) + \eta_{j_x},$$

and hence $d(y,x) > \eta_{j_x}$. Put $\eta = \min_{1\leqslant j\leqslant n}\eta_j$. Consequently, from (9.34) we infer $U^c \subset \bigcap_{x\in K}\{y\in E : d(y,x) > \eta\}$, and hence for all $x \in K$ the event $\{\tau_{U^c} < t\}$ is contained in $\{\sup_{0<s<t} d(X(s),x) > \eta\}$. Putting these observations together shows

$$\sup_{x\in K}\mathbb{P}_x\left[\tau_{U^c} < t\right] \leqslant \sup_{x\in K}\mathbb{P}_x\left[\sup_{0<s<t} d(X(s),x) > \eta\right], \tag{9.35}$$

and hence by (9.35) assertion (i) follows from (ii).

Fix $\eta > 0$, and put $U_n = \{x\in E : d(x,K) < 2^{-n}\eta\}$. In the proofs of the implications (i) $\Longrightarrow$ (iii) and (ii) $\Longrightarrow$ (iv) we take the sequence $(U_n)_{n\in\mathbb{N}}$.

(i) $\Longrightarrow$ (iii). Let $(t_n)_{n\in\mathbb{N}}$ be a sequence which decreases to 0. Since K is a compact subset of U, it follows that $U_n \subset U$ for n sufficiently large. Assuming that the limit in (9.30) does not vanish, then there exists $\delta > 0$ and a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ together with a sequence $(x_k)_{k\in\mathbb{N}}$, $x_k \in U_{n_k}$, such that

$$\mathbb{P}_{x_k}\left[\tau_{U^c} \leqslant t_{n_k}\right] > \delta. \tag{9.36}$$

Since $x_k \in U_{n_k}$ there exists $x'_k \in K$ such that $d(x_k,x'_k) < 2^{-n_k}\eta$. By compactness of K (and metrizability) there exists a subsequence $\left(x'_{k_\ell}\right)_{\ell\in\mathbb{N}}$ which converges to $x' \in K$. Then by the triangle inequality

$$d\left(x_{k_\ell},x'\right) \leqslant d\left(x_{k_\ell},x'_{k_\ell}\right) + d\left(x'_{k_\ell},x'\right) \leqslant 2^{-n_{k_\ell}}\eta + d\left(x'_{k_\ell},x'\right). \tag{9.37}$$

From (9.37) it follows that the set $K' := \{x'\}\bigcup\left\{x_{n_{k_\ell}} : \ell\in\mathbb{N}\right\}$ is compact. From (9.36) we see that

$$\delta < \sup_{x\in K'}\mathbb{P}_x\left[\tau_{U^c} \leqslant t_{n_{k_\ell}}\right]. \tag{9.38}$$

From assertion (i) it follows that the right-hand side of (9.38) converges to 0, when $\ell \to \infty$. Since the latter is a contradiction we see that assertion (iii) is a consequence of (i).

The proof of the implication (ii) $\Longrightarrow$ (iv) follows the same lines: details are left to the reader.

This completes the proof of Proposition 9.2. □

Proposition 9.3. *Let the notation and hypotheses be as in Proposition 9.1. In particular τ_{U^c} is the first hitting time of the complement of the open set U, and K is a compact subset of U. Let $g \in L^\infty\left([0,\infty)\times E, \mathcal{B}_{[0,\infty)}\otimes\mathcal{E}\right)$ and $t > 0$ be fixed. Then the following assertions are true:*

(i) *The following functions are continuous on* U*:*

$$x \mapsto \mathbb{E}_x \left[g\left(t, X(t)\right), \tau_{U^c} > t \right] \quad \textit{and} \quad x \mapsto \mathbb{E}_x \left[g\left(\tau_{U^c}, X\left(\tau_{U^c}\right)\right)\right].$$

(ii) *Let* K *be a compact subset of* U*. Then the family of measures*

$$\left\{ B \mapsto \mathbb{P}_x \left[\left(\tau_{U^c}, X\left(\tau_{U^c}\right)\right) \in B \right] : x \in K \right\}$$

is tight. Here B *varies over the Borel subsets of* $[0, \infty) \times E$*.*

(iii) *The function* $x \mapsto \mathbb{P}_x \left[\tau_{U^c} < \infty\right]$ *is lower semi-continuous.*

In assertion (iii) the subset U may be an arbitrary Borel subset. In the proof we use the fact that $s + \tau_{U^c} \circ \vartheta_s$ decreases to τ_{U^c} $\mathbb{P}_x$-almost surely when s decreases to 0.

Remark 9.5. A proof similar to the proof of (i) shows that the function $x \mapsto \mathbb{P}_x \left[\tau_{U^c} = \infty\right]$ is continuous on U as well.

Proof. For brevity we write $\tau = \tau_{U^c}$. Let $s \in (0, t)$ be arbitrary (small) and $x \in K$ where K is a fixed compact subset of U.

(i) Then we have

$$\begin{aligned}
&\mathbb{E}_x \left[\mathbb{E}_{X(s)} \left[g\left(t, X(t-s)\right), \tau > t-s \right]\right] - \mathbb{E}_x \left[g\left(t, X(t)\right), \tau > t \right] \\
&= \mathbb{E}_x \left[g\left(t, X(t-s)\right) \circ \vartheta_s, \tau \circ \vartheta_s > t-s \right] - \mathbb{E}_x \left[g\left(t, X(t)\right), \tau > t \right] \\
&= \mathbb{E}_x \left[g\left(t, X(t-s)\right) \circ \vartheta_s, \tau \circ \vartheta_s > t-s, \, \tau > s \right] \\
&\quad + \mathbb{E}_x \left[g\left(t, X(t-s)\right) \circ \vartheta_s, \tau \circ \vartheta_s > t-s, \tau \leqslant s \right] \\
&\quad - \mathbb{E}_x \left[g\left(t, X(t)\right), \tau > t \right]
\end{aligned}$$

(on the event $\{\tau > s\}$ the equality $s + \tau \circ \vartheta_s = \tau$ holds $\mathbb{P}_x$-almost surely)

$$= \mathbb{E}_x \left[g\left(t, X(t-s)\right) \circ \vartheta_s, \tau \circ \vartheta_s > t-s, \tau \leqslant s \right]. \tag{9.39}$$

From (9.39) and the Markov property we infer:

$$\begin{aligned}
&\left| \mathbb{E}_x \left[\mathbb{E}_{X(s)} \left[g\left(t, X(t-s)\right), \tau > t-s \right]\right] - \mathbb{E}_x \left[g\left(t, X(t)\right), \tau > t \right] \right| \\
&\leqslant \left\| g\left(t, \cdot\right) \right\|_\infty \mathbb{P}_x \left[\tau \leqslant s \right].
\end{aligned} \tag{9.40}$$

By Proposition 9.1 the right-hand side of (9.40) converges to zero uniformly on compact subsets of U. Since, by the strong Feller property, the functions $x \mapsto \mathbb{E}_x \left[\mathbb{E}_{X(s)} \left[g\left(t, X(t-s)\right), \tau > t-s \right]\right]$, $s \in (0, t)$, are continuous, we infer that the function $x \mapsto \mathbb{E}_x \left[g\left(t, X(t)\right) : \tau > t \right]$ is continuous as well.

Let $h \in L^\infty\left(E, \mathcal{E}\right)$. We will use the continuity on U of functions of the form $x \mapsto \mathbb{E}_x \left[h\left(X(t)\right), \tau > t \right]$ to prove that the function $x \mapsto$

$\mathbb{E}_x\left[g\left(\tau_{U^c}, X\left(\tau_{U^c}\right)\right)\right]$ is continuous on U. To this end let $x \in K$. We consider the following difference:

$$\begin{aligned}
&\mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau<\infty\right]-\mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right]\right] \\
&=\mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau<\infty\right]-\mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right], \tau>s\right] \\
&\quad-\mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right], \tau \leqslant s\right]
\end{aligned}$$

(Markov property)

$$\begin{aligned}
&=\mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau<\infty\right]-\mathbb{E}_x\left[g\left(s+\tau, X(\tau)\right) \circ \vartheta_s, \tau \circ \vartheta_s<\infty, \tau>s\right] \\
&\quad-\mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right], \tau \leqslant s\right] \\
&=\mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau<\infty\right] \\
&\quad-\mathbb{E}_x\left[g\left(s+\tau \circ \vartheta_s, X(s+\tau \circ \vartheta_s)\right), s+\tau \circ \vartheta_s<\infty, \tau>s\right] \\
&\quad-\mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right], \tau \leqslant s\right]
\end{aligned}$$

(on the event $\{\tau > s\}$ the equality $s + \tau \circ \vartheta_s = \tau$ holds $\mathbb{P}_x$-almost surely)

$$\begin{aligned}
&=\mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau \leqslant s\right] \\
&\quad-\mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right], \tau \leqslant s\right].
\end{aligned} \tag{9.41}$$

By the strong Feller property the functions

$$x \mapsto \mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right]\right], \quad s>0,$$

are continuous. From (9.20) in Proposition 9.1 together with (9.41) we see that, uniformly on the compact subset K, the functions

$$x \mapsto \mathbb{E}_x\left[\mathbb{E}_{X(s)}\left[g\left(s+\tau, X(\tau)\right), \tau<\infty\right]\right], \quad s>0,$$

converge to $x \mapsto \mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau<\infty\right]$ whenever $s \downarrow 0$. Consequently, since K is an arbitrary compact subset of U we see that the function $x \mapsto \mathbb{E}_x\left[g\left(\tau, X(\tau)\right), \tau<\infty\right]$ is continuous on U.

(ii). Let K be a compact subset of the open subset U, and let $\tau = \tau_{U^c}$ the hitting of U^c, the complement of U. In order to prove that the family of $\mathbb{P}_x$-distributions, $x \in K$, of the space-time variable $(\tau, X(\tau))$, is tight, by assertion (a) of Theorem 2.3 it suffices to prove that for every sequence of bounded continuous functions $(f_n : n \in \mathbb{N}) \subset C_b([0,\infty) \times E)$ which decreases pointwise to zero we have:

$$\lim_{n\to\infty} \sup_{x\in K} \mathbb{E}_x\left[f_n\left(\tau, X(\tau)\right), \tau<\infty\right]=0. \tag{9.42}$$

By Dini's lemma and by assertion (i), the equality in (9.42) follows from the pointwise equality:

$$\lim_{n\to\infty} \mathbb{E}_x\left[f_n\left(\tau, X(\tau)\right), \tau<\infty\right]=0. \tag{9.43}$$

The equality in (9.43) follows from Lebesgue's dominated convergence theorem.

(iii) By the Markov property we have the equalities

$$\mathbb{P}_x\left[\tau<\infty\right]=\sup_{s>0}\mathbb{P}_x\left[s+\tau\circ\vartheta_s<\infty\right]=\sup_{s>0}\mathbb{E}_x\left[\mathbb{P}_{X(s)}\left[\tau<\infty\right]\right]. \quad (9.44)$$

Functions of the form $x\mapsto\mathbb{E}_x\left[g(X(s))\right]$, where g is a bounded Borel function, are continuous, and hence by (9.44) the function $x\mapsto\mathbb{P}_x\left[\tau<\infty\right]$ is lower semi-continuous. The same argument works in case τ is the hitting time of a Borel subset of E.

This completes the proof of Proposition 9.3. □

Under the hypotheses of the equivalent properties in Proposition 9.2 it will be shown that there exists a compact recurrent subset, provided all open subsets are recurrent. More precisely we have the following result.

Proposition 9.4. *Suppose that there exists a point $x_0\in E$ such that every open neighborhood of x_0 is recurrent, and suppose that the equivalent properties in Proposition 9.2 are satisfied. In addition, suppose that all probability measures $B\mapsto P(t,x,B)$, $(t,x)\in(0,\infty)\times E$, are equivalent. Then there exists a compact recurrent subset. In fact, the following assertion is true. Fix $t_0>0$, and let K be a compact subset of E with the property that $P(t_0,x_0,K)>0$, and $x_0\notin K$. Then K is recurrent.*

Since the equivalent properties in Proposition 9.2 are satisfied whenever the domain of the generator L almost separates points and closed subsets and if it contains the constant functions, the following corollary is an immediate consequence of Proposition 9.4.

Corollary 9.2. *Suppose that there exists a point $x_0\in E$ such that every open neighborhood of x_0 is recurrent, and suppose that the domain of the generator L almost separates points and closed subsets, and contains the constant functions. In addition, suppose that all probability measures $B\mapsto P(t,x,B)$, $(t,x)\in(0,\infty)\times E$, are equivalent. Then there exists a compact recurrent subset. More precisely, the following statement is true. Fix $t_0>0$, and let K be a compact subset of E with the property that $P(t_0,x_0,K)>0$, and $x_0\notin K$. Then K is recurrent.*

Proof. Let x_0 be as in Proposition 9.4. Fix $t_0>0$, and let K be a compact subset of E with the property that $P(t_0,x_0,K)>0$, and $x_0\notin K$. By inner-regularity of the measure $B\mapsto P(t_0,x_0,B)$ such compact subset K exists, We shall prove that K is recurrent. Let τ_K be the first hitting time of K and let $(U_\ell)_{\ell\in\mathbb{N}}$ be sequence of open neighborhoods of x_0 with respective first hitting times $\tau^{(\ell)}$, $\ell \in \mathbb{N}$. We suppose that this sequence forms a neighborhood base of x_0, and that $\overline{U}_{\ell+1} \subset U_\ell$ where $\overline{U}_{\ell+1}$ stand for the closure of $U_{\ell+1}$. We assume that $U_\ell \bigcap K = \emptyset$. For every $\ell \in \mathbb{N}$ we define the following sequence of stopping times: $\tau_1^{(\ell)} = \tau^{(\ell)}$, and

$$\tau_{n+1}^{(\ell)} = \inf\left\{s > \tau_n^{(\ell)} + 2t_0 : X(s) \in U_\ell\right\}. \tag{9.45}$$

Since the open subset U_ℓ is recurrent, the hitting times $\tau_n^{(\ell)}$ are finite $\mathbb{P}_x$-almost surely for all $x \in E$, and for all $n \in \mathbb{N}$. As in the proof of Lemma 9.3 below we introduce the following sequence of events:

$$A_n^\ell = \left\{\tau_n^{(\ell)} \leqslant \tau_n^{(\ell)} + \tau_K \circ \vartheta_{\tau_n^{(\ell)}} \leqslant \tau_n^{(\ell)} + t_0\right\} = \left\{\tau_K \circ \vartheta_{\tau_n^{(\ell)}} \leqslant t_0\right\}, \tag{9.46}$$

ℓ, $n \in \mathbb{N}$. Then $A_n^{(\ell)} \in \mathcal{F}_{\tau_{n+1}^{(\ell)}}$, and we have

$$\mathbb{P}_x\left[A_n^{(\ell)} \;\middle|\; \mathcal{F}_{\tau_n^{(\ell)}}\right] = \mathbb{E}_x\left[\mathbb{P}_{X(\tau_n^\ell)}\left[\tau_K \leqslant t_0\right]\right] \geqslant \inf_{y\in\overline{U}_\ell} P_y\left[\tau_K \leqslant t_0\right]. \tag{9.47}$$

By assertion (i) in Proposition 9.3 we see that the function $y \mapsto \mathbb{P}_y\left[\tau_K \leqslant t_0\right] = 1 - \mathbb{P}\left[\tau_K > t_0\right]$ is continuous at $y = x_0$. From (9.47) it then follows that

$$\begin{aligned}\mathbb{P}_x\left[A_n^{(\ell)} \;\middle|\; \mathcal{F}_{\tau_n^{(\ell)}}\right] &\geqslant \frac{1}{2}\mathbb{P}_{x_0}\left[\tau_K \leqslant t_0\right] \geqslant \frac{1}{2}\mathbb{P}_{x_0}\left[X\left(t_0\right) \in K\right] \\ &= \frac{1}{2}P\left(t_0, x_0, K\right) > 0 \end{aligned}\tag{9.48}$$

for $\ell \geqslant \ell_0$. From the generalized Borel-Cantelli lemma (or the Borel-Cantelli-Lévy lemma) it then follows that $\mathbb{P}_x\left[\sum_{n=1}^\infty \mathbf{1}_{A_n^{(\ell)}} = \infty\right] = 1$, $\ell \geqslant \ell_0$, and hence the compact subset K is recurrent. For a precise formulation of the Borel-Cantelli-Lévy lemma the reader is referred to [Shiryayev (1984)] Corollary 2 page 486 or to Theorem 9.3 below.

This completes the proof of Proposition 9.4. □

A precise formulation of the generalized Borel-Cantelli lemma reads as follows: see e.g. (the proof of) Corollary 5.29 in [Breiman (1992)].

Theorem 9.3. *Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space and let $(\mathcal{G}_n)_{n\in\mathbb{N}}$ be filtration in $\mathcal{G}$. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events such that $A_n \in \mathcal{G}_{n+1}$, $n \in \mathbb{N}$. Then the following equality of events holds $\mathbb{P}$-almost surely:*

$$\{\omega \in \Omega : \omega \in A_n \text{ infinitely often}\} = \left\{\omega \in \Omega : \sum_{n=1}^\infty \mathbb{P}\left[A_n \;\middle|\; \mathcal{G}_n\right](\omega) = \infty\right\}. \tag{9.49}$$

The following result shows that Proposition 9.4 also holds for Markov chains.

Proposition 9.5. *Let*

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(n), n \in \mathbb{N}), (\vartheta_n, n \in \mathbb{N}), (E, \mathcal{N})\} \tag{9.50}$$

be a Markov chain with the property that all Borel measures $B \mapsto P(1, x, B) = \mathbb{P}_x[X(1) \in B]$, $x \in E$, *are equivalent. In addition suppose that for every Borel subset* B *the function* $x \mapsto P(1, x, B)$ *is continuous. Let there exist a point* $x_0 \in E$ *such that every open neighborhood of* x_0 *is recurrent. Then there exists a compact recurrent subset. In fact, the following assertion is true. Let* K *be a compact subset of* $E\backslash\{x_0\}$ *with the property that* $P(1, x_0, K) > 0$. *Then* K *is recurrent, i.e.* $\mathbb{P}_x\left[\tau_K^1 < \infty\right] = 1$ *for all* $x \in E$.

Here $\tau_K^1 = \inf\{k \geqslant 1 : k \in \mathbb{N}, X(k) \in K\}$.

Proof. The proof can be copied from the proof of Proposition 9.4 with $t_0 = 1$, $\tau_K = \tau_K^1$. A similar convention is used for the hitting times of the open neighborhoods U_ℓ of x_0. Also notice that $\left\{\tau_K^1 \leqslant 1\right\} = \{X(1) = K\}$, and that the function $x \mapsto P(1, x, K)$ is continuous.

These arguments suffice to complete the proof of Proposition 9.5. □

Next we collect some of the results proved so far. The existence of a compact recurrent subset will also be used when we prove the existence of a σ-finite invariant Radon measure: see Theorem 10.12. For the notion of strong Feller property see Definitions 2.5 and 2.16.

Theorem 9.4. *As in Theorem 9.2 let*

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x)_{x\in E}, (X(t), t \geqslant 0), (E, \mathcal{E})\} \tag{9.51}$$

be a time-homogeneous Markov process on a Polish space E *with a transition probability function* $P(t, x, \cdot)$, $t \geqslant 0$, $x \in E$, *which is conservative in the sense that* $P(t, x, E) = 1$ *for all* $t \geqslant 0$ *and* $x \in E$. *Assume that the process* $X(t)$ *is strong Feller in the sense that for all Borel subsets* B *of* E *the function* $(t, x) \mapsto P(t, x, B)$ *is continuous on* $(0, \infty) \times E$. *In addition, suppose that all measures* $B \mapsto P(t, x, B)$, $B \in \mathcal{E}$, $t > 0$, $x \in E$, *are equivalent. Suppose that there exists* $x_0 \in E$ *with the property that all open neighborhoods of* x_0 *are recurrent. In addition assume that the generator of the process almost separate points and closed subsets, in the sense that for every* $x \in U$ *with* U *open there exists a function* $v \in D(L)$ *such that* $v(x) > \sup_{y\in E\backslash U} v(y)$. *Then there exists a compact subset* A *which is* *recurrent. Moreover, every Borel subset B for which $P(t_0, x_0, B) > 0$ for some $(t_0, x_0) \in (0, \infty)$ is recurrent.*

Theorem 10.8 and its companion Theorem 10.9 show that under the hypotheses of Theorem 9.4 a Borel subset is recurrent if and only if it is Harris recurrent. For the notion of the almost separation property the reader may want to see Remark 9.1 following Definition 9.2: see Proposition 9.1 as well.

Proof. In assertion (i) of Proposition 9.1 it is shown that the almost separation implies the very relevant property (9.20) which is somewhat strengthened in Proposition 9.2. Using this property we see that the function $x \mapsto \mathbb{P}_x[\tau_A \leqslant t]$ is continuous on $E \backslash A$ where A is compact. From the proof of Proposition 9.4 it follows that there exists a recurrent compact subset. From Lemma 9.3 below we see that all Borel subsets B for which $P(t_0, x_0, B) > 0$ for some $(t_0, x_0) \in (0, \infty) \times E$ are recurrent.

This completes the proof of Theorem 9.4. □

Again we consider the time-homogeneous Markov process (9.14) in Theorem 9.2. Theorem 10.8 and its companion Theorem 10.9 show that in the context of a strong Markov process with the strong Feller property the collection of recurrent Borel subsets coincides with the collection of Harris recurrent subsets provided that all measures $B \mapsto P(t, x, B)$, $B \in \mathcal{E}$, $(t, x) \in (0, \infty) \times E$, are equivalent.

Definition 9.4. Let A be a Borel subset of E, and τ_A its first hitting time: $\tau_A = \inf\{s > 0 : X(s) \in A\}$. The subset A is called recurrent if

$$\mathbb{P}_x[\tau_A < \infty] = 1 \quad \text{for all } x \in E.$$

The subset A is called Harris recurrent provided

$$\mathbb{P}_x\left[\int_0^\infty \mathbf{1}_A(X(s))\, ds = \infty\right] = 1 \quad \text{for all } x \in E. \tag{9.52}$$

Definition 9.5. Let μ be an invariant measure for the Markov process in (9.14). Then the Markov process is μ-Harris recurrent provided every Borel subset A for which $\mu(A) > 0$ is Harris recurrent. Suppose that all measures $B \mapsto P(t, x, B)$, $B \in \mathcal{E}$, $(t, x) \in (0, \infty) \times E$, are equivalent. Then the corresponding Markov process is called Harris recurrent if every Borel subset for which $P(1, x_0, B) > 0$ for some $x_0 \in E$ is Harris recurrent.

The following theorem says among other things that, if the Markov process possesses a finite invariant measure μ, then there exists a compact recurrent subset K of E such that $\mu(K) > 0$. It is closely related to Theorem

2.1 in [Seidler (1997)]. An adapted version will be employed in the proof of Theorem 10.12 in Chapter 10: see (10.221)–(10.238). In particular the σ-finiteness will be at stake: see the arguments after the (in-)equalities (10.205) and (10.220). Another variant can be found in Theorem 9.11 below. For more details on the notion of strong Feller property see Definitions 2.5 and 2.16.

Theorem 9.5. *Let the Markov process have right-continuous sample paths, be strong Feller, and irreducible. Let $K \subset E$ be a compact subset which is non-recurrent. Then*

$$\sup_{x\in E} \int_0^\infty P(t,x,K)\,dt < \infty, \quad \text{and} \quad \mu(K) = 0 \tag{9.53}$$

for all finite invariant measures μ. If, in addition, $P(1,x_0,K) > 0$ for some $x_0 \in E$, then

$$\mathbb{P}_x\left[\sup\left\{t \geqslant 0,\ X(t) \in K'\right\} < \infty\right] = 1, \quad \text{and} \quad \lim_{t\to\infty} P\left(t,x,K'\right) = 0 \tag{9.54}$$

for all $x \in E$ and all compact subsets K'.

Proof. Let K be a non-recurrent compact subset of E. We begin by showing that

$$\sup_{x\in E} \int_0^\infty P(t,x,K)\,dt < \infty. \tag{9.55}$$

The proof of (9.55) follows the same pattern as the corresponding proof by Seidler in [Seidler (1997)], who in turn follows Khasminskii [Has′minskiĭ (1960)]. Let τ be the first hitting time of K. Since K is non-recurrent there exists $y_0 \notin K$ such that

$$\mathbb{P}_{y_0}[\tau = \infty] = \mathbb{P}_{y_0}[X(t) \notin K \ \text{ for all } t \geqslant 0] > 0.$$

By Remark 9.5 which follows Proposition 9.3 the function $x \mapsto \mathbb{P}_x[\tau = \infty]$ is continuous on $E\backslash K$. Hence there exists an open neighborhood V of y_0 such that

$$\alpha := \inf_{x\in V} \mathbb{P}_x[\tau = \infty] > 0. \tag{9.56}$$

Fix $t_0 > 0$ arbitrary, and choose $y \in K$. Then by the Markov property we have

$$\begin{aligned}
&\mathbb{P}_y\left[\int_0^\infty \mathbf{1}_K(X(t))\,dt < t_0\right] \\
&= \mathbb{E}_y\left[\omega \mapsto \mathbb{P}_{X(t_0)(\omega)}\left[\int_0^{t_0} \mathbf{1}_K(X(t)(\omega))\,dt + \int_0^\infty \mathbf{1}_K(X(t))\,dt < t_0\right]\right]
\end{aligned}$$

$$\geqslant \mathbb{E}_y\left[\omega\mapsto \mathbb{P}_{X(t_0)(\omega)}\left[\int_0^{t_0}\mathbf{1}_K\left(X(t)(\omega)\right)dt<t_0,\ X(t)\notin K\ \text{ for all } t\geqslant 0\right]\right]$$
$$\geqslant \mathbb{P}_y\left[\int_0^{t_0}\mathbf{1}_K\left(X(t)\right)dt<t_0,\ X(t)\notin K\ \text{ for all } t\geqslant t_0\right]$$
$$\geqslant \mathbb{P}_y\left[\int_0^{t_0}\mathbf{1}_K\left(X(t)\right)dt<t_0,\ X\left(t_0\right)\in V,\ X(t)\notin K\ \text{ for all } t\geqslant t_0\right]$$
$$\geqslant \mathbb{E}_y\left[\mathbb{P}_{X(t_0)}\left[\tau=\infty\right],\ X\left(t_0\right)\in V\right]$$

(apply (9.56), the definition of α)

$$\geqslant \alpha P\left(t_0,y,V\right)\geqslant \alpha\inf_{x\in K}P\left(t_0,x,V\right)=:q>0, \tag{9.57}$$

where we used the irreducibility of our Markov process, and the continuity of the function $x\mapsto P\left(t_0,x,V\right)$. Hence we infer

$$\sup_{y\in K}\mathbb{P}_y\left[\int_0^{\infty}\mathbf{1}_K\left(X(t)\right)dt\geqslant t_0\right]\leqslant 1-q. \tag{9.58}$$

Put

$$\kappa=\inf\left\{t>0:\int_0^t\mathbf{1}_K\left(X(s)\right)ds\geqslant t_0\right\}=\inf\left\{t>0:\int_0^t\mathbf{1}_K\left(X(s)\right)ds=t_0\right\}. \tag{9.59}$$

Then κ is a stopping time relative to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, because $X(s)$ is $\mathcal{F}_t$-measurable for all $0\leqslant s\leqslant t$. Moreover, by right-continuity of the process $t\mapsto X(t)$ it follows that $X(\kappa)\in K$ on the event $\{\tau<\infty\}$. Let $y\in E$. By induction we shall prove that

$$\mathbb{P}_y\left[\int_0^{\infty}\mathbf{1}_K\left(X(t)\right)dt>kt_0\right]\leqslant (1-q)^{k-1},\ k\in\mathbb{N},\ k\geqslant 1. \tag{9.60}$$

To this end we put

$$\alpha_k=\sup_{x\in K}\mathbb{E}_x\left[\int_0^{\infty}\mathbf{1}_K\left(X(s)\right)ds\geqslant kt_0\right]. \tag{9.61}$$

If x belongs to K, then by the Markov property we have:

$$\mathbb{P}_x\left[\int_0^{\infty}\mathbf{1}_K\left(X(s)\right)ds>(k+1)t_0\right]=\mathbf{P}_x\left[\int_{\kappa}^{\infty}\mathbf{1}_K\left(X(s)\right)ds>kt_0,\ \kappa<\infty\right]$$
$$=\mathbb{E}_x\left[\mathbb{P}_{X(\kappa)}\left[\int_0^{\infty}\mathbf{1}_K\left(X(s)\right)ds>kt_0\right],\ \kappa<\infty\right]$$
$$=\mathbb{E}_x\left[\mathbb{P}_{X(\kappa)}\left[\int_0^{\infty}\mathbf{1}_K\left(X(s)\right)ds>kt_0\right],\ \int_0^{\infty}\mathbf{1}_K\left(X(s)\right)ds\geqslant t_0\right]$$
$$\leqslant \alpha_1\alpha_k. \tag{9.62}$$

From (9.62) and induction we infer

$$\begin{aligned}&\sup_{x\in K}\mathbb{P}_x\left[\int_0^\infty \mathbf{1}_K\left(X(s)\right)ds \geqslant kt_0\right]\\&\leqslant \alpha_1^k = \left(\sup_{x\in K}\left[\int_0^\infty \mathbf{1}_K\left(X(s)\right)ds \geqslant t_0\right]\right)^k \leqslant (1-q)^k,\end{aligned} \tag{9.63}$$

where in the final step of (9.63) we employed (9.58). If $y \in E$ is arbitrary, then we proceed as follows:

$$\begin{aligned}&\mathbb{P}_y\left[\int_0^\infty \mathbf{1}_K\left(X(s)\right)ds > (k+1)t_0\right]\\&= \mathbb{P}_y\left[\int_\kappa^\infty \mathbf{1}_K\left(X(s)\right)ds > kt_0,\, \kappa < \infty\right]\\&= \mathbb{E}_y\left[\mathbb{P}_{X(\kappa)}\left[\int_0^\infty \mathbf{1}_K\left(X(s)\right)ds > kt_0\right],\, \kappa<\infty\right]\\&\leqslant (1-q)^k\mathbb{P}_y\left[\kappa<\infty\right] \leqslant (1-q)^k.\end{aligned} \tag{9.64}$$

The inequality in (9.64) implies the inequality in (9.60). To show the first part of (9.53) we observe that for $x \in E$ we have

$$\begin{aligned}\int_0^\infty P\left(t,x,K\right)dt &= \mathbb{E}_x\left[\int_0^\infty \mathbf{1}_K\left(X(s)\right)ds\right]\\&\leqslant \sum_{k=1}^\infty kt_0\mathbb{P}_x\left[(k-1)t_0 < \int_0^\infty \mathbf{1}_K\left(X(s)\right)ds \leqslant kt_0\right]\\&\leqslant t_0 + \sum_{k=2}^\infty \mathbb{P}_x\left[\int_0^\infty \mathbf{1}_K\left(X(s)\right)ds > (k-1)t_0\right]\\&\leqslant t_0 + t_0\sum_{k=2}^\infty k(1-q)^{k-2} = t_0\left(1+\frac{1}{q}+\frac{1}{q^2}\right) < \infty.\end{aligned} \tag{9.65}$$

The first part of (9.53) is a consequence of (9.65) indeed.

In fact from (9.65) we also obtain $\mu(K) = 0$ for any finite invariant measure μ. Let μ be an invariant probability measure. That $\mu(K) = 0$ can be seen by the following (standard) arguments:

$$\begin{aligned}\mu(K) &= \frac{1}{T}\int_0^T \mu(K)\,dt = \frac{1}{T}\int_0^T\int_E P\left(t,y,K\right)d\mu(y)dt\\&= \int_E\left(\frac{1}{T}\int_0^T P(t,y,K)\,dt\right)d\mu(y) \leqslant \frac{1}{T}\sup_{x\in E}\int_0^\infty P\left(t,x,K\right)dt\end{aligned}$$

$$\leqslant \frac{t_0}{T}\left(1+\frac{1}{q}+\frac{1}{q^2}\right). \tag{9.66}$$

Since $T > 0$ is arbitrary (9.66) implies $\mu(K) = 0$.

Next assume that the compact subset K has the additional property that $P(1, x_0, K) > 0$. Let K' be an arbitrary compact subset. We want to prove that

$$\mathbb{P}_x\left[\sup\left\{t \geqslant 0 : X(t) \in K'\right\} < \infty\right] = 1. \tag{9.67}$$

Put $h(x) = \int_0^\infty P(t, x, K)\, dt$. Then by (9.65) the function $h \in L^\infty(E, \mathcal{E})$. The function $h(x)$ is also lower semi-continuous, because the functions $x \mapsto P(t, x, K)$, $t > 0$, are continuous. Moreover, it is strictly positive, by the fact that for all $t > 0$ and all $x \in E$, $P(t, x, K) > 0$. Put $H_n = \left\{h > n^{-1}\right\}$. Then there exists $m \in \mathbb{N}$ such that $K' \subset H_m$. Fix $x \in E$, denote by σ the first hitting time of K', and let $\sigma(k)$ be the first hitting time of K' after time k, i.e. $\sigma(k) = k + \sigma \circ \vartheta_k$. Taking into account that $X(\sigma(k)) \in K' \subset H_m$ $\mathbb{P}_x$-almost surely on the event $\{\sigma(k) < \infty\}$ we obtain:

$$\begin{aligned}
\frac{1}{m}\mathbb{P}_x[\sigma(k) < \infty] &\leqslant \mathbb{E}_x[h(X(\sigma(k))), \sigma(k) < \infty] \\
&= \mathbb{E}_x\left[\int_0^\infty P(s, X(\sigma(k)), K)\, ds,\ \sigma(k) < \infty\right] \\
&= \int_0^\infty \mathbb{E}_x\left[\mathbb{P}_{X(\sigma(k))}\left[X(s) \in K\right], \sigma(k) < \infty\right] ds \\
&= \int_0^\infty \mathbb{E}_x\left[\mathbb{P}_x\left[X(s + \sigma(k)) \in K \;\middle|\; \mathcal{F}_{\sigma(k)}\right], \sigma(k) < \infty\right] ds \\
&= \int_0^\infty \mathbb{E}_x\left[\mathbf{1}_K(X(s + \sigma(k))), \sigma(k) < \infty\right] ds \\
&= \mathbb{E}_x\left[\int_{\sigma(k)}^\infty \mathbf{1}_K(X(s))\, ds,\ \sigma(k) < \infty\right]
\end{aligned}$$

($\sigma(k) \geqslant k$ on the event $\{\sigma(k) < \infty\}$)

$$\begin{aligned}
&\leqslant \mathbb{E}_x\left[\int_k^\infty \mathbf{1}_K(X(s))\, ds,\ \sigma(k) < \infty\right] \\
&\leqslant \int_k^\infty P(s, x, K)\, ds.
\end{aligned} \tag{9.68}$$

The sequence of events $\{\sigma(k) < \infty\}$, $k \in \mathbb{N}$, decreases. From (9.68) it follows that its intersection has $\mathbb{P}_x$-measure zero. It follows that its complement has full $\mathbb{P}_x$-measure. This means that for $\mathbb{P}_x$-almost all ω there exists $k \in \mathbb{N}$

such that $\sigma(k)(\omega) = \infty$, and, consequently, (9.67) holds. From (9.67) we also readily infer $\lim_{t\to\infty} P(t,x,K') = 0$, because after the process $X(t)$ has visited K' it only returns there finitely many times $\mathbb{P}_x$-almost surely.

This completes the proof of Theorem 9.5. □

A stopping time of the form $\inf\left\{t > 0 : \int_0^t \mathbf{1}_K (X(s))\, ds > 0\right\}$ is called the penetration time of K: compare with (9.59).

Lemma 9.3. *Let the hypotheses and notations be as in Theorem 9.2. Suppose that there exists a compact subset K which is recurrent. Then all Borel subsets B with the property that $P(t_0, x_0, B) > 0$ for some pair $(t_0, x_0) \in (0,\infty) \times E$ (or, equivalently, $P(t,x,B) > 0$ for all pairs $(t,x) \in (0,\infty) \times E$) are recurrent.*

Proof. Let $B \in \mathcal{E}$ be such that $P(t,x,B) > 0$ for some (all) pairs $(t,x) \in (0,\infty) \times E$. Let τ_B be the (first) hitting time of B: $\tau_B = \inf\{t > 0 : X(t) \in B\}$. We need to show that $\mathbb{P}_x[\tau_B < \infty] = 1$ for all $x \in E$. By our assumptions we have

$$\inf_{x\in K} \mathbb{P}_x[\tau_B \leqslant 1] \geqslant \inf_{x\in K} \mathbb{P}_x[X(1) \in B] = \inf_{x\in K} P(1,x,B) =: q > 0. \tag{9.69}$$

Let τ be the first hitting time of K, and define a sequence of hitting times of K as follows:

$$\tau_1 = \tau, \quad \text{and} \quad \tau_{n+1} = \inf\{t > \tau_n + 2 : X(t) \in K\} = \tau_n + 2 + \tau \circ \vartheta_{\tau_n+2}. \tag{9.70}$$

Then, for any $n \in \mathbb{N}$, $\tau_n < \infty$ and $X(\tau_n) \in K$ $\mathbb{P}_x$-almost surely for all $x \in E$. Put

$$A_n = \{\tau_n \leqslant \tau_n + \tau_B \circ \vartheta_{\tau_n} \leqslant \tau_n + 1\} = \{\tau_B \circ \vartheta_{\tau_n} \leqslant 1,\, \tau_n < \infty\}. \tag{9.71}$$

The events in (9.71) should be compared with similar ones in (9.46). Then $A_n \in \mathcal{F}_{\tau_n+1} \subset \mathcal{F}_{\tau_{n+1}}$, and we have with q as in (9.69)

$$\begin{aligned}\sum_{n=1}^{\infty} \mathbb{P}_x\left[A_n \,\middle|\, \mathcal{F}_{\tau_n}\right] &= \sum_{n=1}^{\infty} \mathbb{P}_x\left[\{\tau_B \circ \vartheta_{\tau_n} \leqslant 1\} \,\middle|\, \mathcal{F}_{\tau_n}\right] \\ &= \sum_{k=1}^{\infty} \mathbb{P}_{X(\tau_n)}[\tau_B \leqslant 1] \geqslant \sum_{k=1}^{\infty} \inf_{y\in K} \mathbb{P}_y[\tau_B \leqslant 1] \\ &\geqslant \sum_{k=1}^{\infty} q = \infty, \quad \mathbb{P}_x\text{-almost surely}\end{aligned} \tag{9.72}$$

for all $x \in E$. Therefore by the generalized Borel-Cantelli lemma (see e.g. [Shiryayev (1984)] Corollary VII 5.2, or see equality (9.49) in Theorem 9.3)

$\mathbb{P}_x$-almost all ω belong to A_n for infinitely many $n \in \mathbb{N}$. However, if $\omega \in A_n$, then $\tau_B(\omega) \leqslant \tau_n + 1 < \infty$.

This concludes the proof of Lemma 9.3. □

The following result is a reformulation of Lemma 9.3 for Markov chains with values in E. Its proof can be copied from the proof of Lemma 9.3.

Lemma 9.4. *Let the notation and hypotheses be as in Proposition 9.5. Suppose that there exists a compact subset K which is recurrent. Then all Borel subsets B with the property that $P(t_0, x_0, B) > 0$ for some pair $(t_0, x_0) \in (0, \infty) \times E$ (or, equivalently, $P(t, x, B) > 0$ for all pairs $(t, x) \in (0, \infty) \times E$) are recurrent.*

For the notion of a Harris recurrent subset, the reader is referred to Definition 9.4. The following result follows merely from the recurrence properties of our Markov process. These recurrence properties were established in Lemma 9.3. The existence of a finite invariant probability measure is not required.

Proposition 9.6. *Let the hypotheses and notation be as in Theorem 9.2 except that the existence of an invariant probability measure is required. Assume that there exists a compact recurrent subset. Then every non-empty open subset U of E is Harris recurrent.*

Proof. Let U be any open subset of E. Suppose $\varnothing \neq U \neq E$. Since our Markov process is recurrent, there exists a pair $(t_0, x_0) \in (0, \infty) \times E$ such that $P(t_0, x_0, U) > 0$. Let the compact subset K of U be such that $P(t_0, x_0, K) \geqslant \frac{1}{2} P(t_0, x_0, U) > 0$. From Lemma 9.3 we infer that the compact subset K is recurrent. Let τ_{U^c} be the hitting time of $E \backslash U$. From (9.20) in Proposition 9.1 we see that there exists $q > 0$ such that $\sup_{y \in K} \mathbb{P}_y \left[\tau_{U^c} \leqslant q\right] < \dfrac{1}{2}$. Then we see $\inf_{y \in K} \mathbb{P}_y \left[\tau_{U^c} > q\right] \geqslant \dfrac{1}{2}$. Let $\tau = \tau_K$ be the first hitting time of K. Then by recurrence $\mathbb{P}_x[\tau < \infty] = 1$, $x \in E$. Instead of τ_{U^c} we write σ. We define the double sequence of hitting times of K and $E \backslash U$:

$$\tau_1 = \tau, \quad \sigma_n = \tau_n + \sigma \circ \vartheta_{\tau_n}, \quad \tau_{n+1} = \sigma_n + \tau \circ \vartheta_{\sigma_n}. \tag{9.73}$$

In addition we introduce the events: $Q_n = \{\sigma_n - \tau_n > q\}$. For every $y \in E$ we have:

$$\sum_{n=1}^{\infty} \mathbb{P}_y \left[Q_n \;\middle|\; \mathcal{F}_{\tau_n}\right] = \sum_{n=1}^{\infty} \mathbb{P}_y \left[\sigma \circ \vartheta_{\tau_n} > q \;\middle|\; \mathcal{F}_{\tau_n}\right]$$

$$= \sum_{n=1}^{\infty} \mathbb{P}_{X(\tau_n)}\left[\sigma > \tau\right] \geqslant \sum_{n=1}^{\infty} \inf_{x\in K} \mathbb{P}_x\left[\sigma > \eta\right]$$
$$\geqslant \sum_{n=1}^{\infty} \frac{1}{2} = \infty, \quad \mathbb{P}_y\text{-almost surely.} \tag{9.74}$$

From (9.74) and the generalized Borel-Cantelli lemma (again see e.g. [Shiryayev (1984)] Corollary VII 5.2 or the equality in (9.49) in Theorem 9.3) we infer

$$\mathbb{P}_y\left[\limsup_{n\to\infty} Q_n\right] = 1, \quad y \in E. \tag{9.75}$$

Since $\mathbf{1}_U\left(X(t)\right) = 1$ on the event $\{\tau_n \leqslant t < \sigma_n\}$, (9.75) implies

$$\mathbb{P}_y\left[\int_0^{\infty} \mathbf{1}_U\left(X(t)\right) dt = \infty\right] = 1, \quad y \in E.$$

In other words the open subset U is Harris recurrent.

This completes the proof of Proposition 9.6. □

Definition 9.6. Let $\left\{\mathcal{F}_{t_2}^{t_1} : 0 \leqslant t_1 \leqslant t_2 < \infty\right\}$ be a collection of σ-fields on Ω such that, for every $t_1 \in [0,\infty)$ fixed, the collection $\left(\mathcal{F}_{t_2}^{t_1}\right)_{t_2\geqslant t_1}$ is a filtration, and such that for every $t_2 \in (0,\infty)$ the collection $\left(\mathcal{F}_{t_2}^{t_1}\right)_{t_1\leqslant t_2}$ is also a filtration. A family of random variables $A\left(t_1,t_2\right) : \Omega \to \mathbb{R}$, $0 \leqslant t_1 \leqslant t_2 < \infty$, is called an additive process relative to the collection $\left\{\mathcal{F}_{t_2}^{t_1} : 0 \leqslant t_1 \leqslant t_2 < \infty\right\}$ if it possesses the following properties:

(1) the equality $A\left(t_1,t_2\right) + A\left(t_2,t_3\right) = A\left(t_1,t_3\right)$ holds for all $0 \leqslant t_1 \leqslant t_2 \leqslant t_3$;
(2) for every $0 \leqslant t_1 \leqslant t_2$ the random variable $A\left(t_1,t_2\right)$ is $\mathcal{F}_{t_2}^{t_1}$-measurable.

In case of a time-homogeneous Markov process, like in Theorem 9.2 an additive process $A\left(t\right) : \Omega \to \mathbb{R}$, $0 \leqslant t < \infty$, is called a time-homogeneous additive process relative to the collection $\{\mathcal{F}_t : 0 \leqslant t < \infty\}$ if it possesses the following properties:

(1) the equality $A\left(s\right) + A\left(t-s\right)\circ\vartheta_s = A\left(t\right)$ holds $\mathbb{P}_x$-almost surely for all $0 \leqslant s \leqslant t$;
(2) for every $t \geqslant 0$ the random variable $A\left(t\right)$ is $\mathcal{F}_t$-measurable.

If in the above definitions the *plus signs* are replaced with *multiplication signs*, then the corresponding processes are called multiplicative and time-homogeneous multiplicative respectively.

Instead of time-homogeneous additive process we usually just say additive process; a similar convention is adopted in case of multiplicative processes.

If $A(t_1, t_2)$ is an additive process, then $\exp(A(t_1, t_2)$ is a multiplicative process.

In fact there is a relationship between these two notions. Let $t \mapsto A(t)$ be an additive process in the time-homogeneous case. Then it can also be considered as an additive process of two variables by writing $A(t_1, t_2) = A(t_2 - t_1) \circ \vartheta_{t_1}$, $0 \leqslant t_1 \leqslant t_2 < \infty$.

Let $f : [0, \infty) \times E \to \mathbb{R}$ be a Borel measurable function with the property that $\int_0^t |f(s, X(s))| \, ds < \infty$, $\mathbb{P}_x$-almost surely for all $x \in E$. Then the process $(t_1, t_2) \mapsto A_f(t_1, t_2) = \int_{t_1}^{t_2} f(\rho, X(\rho)) \, d\rho$ is an additive process. In the time-homogenous case, and if the function f only depends on the state variable, then the process $t \mapsto A_f(t) := \int_0^t f(X(\rho)) \, d\rho$ is a (time-homogenous) additive process. Let $\tau : \Omega \to [0, \infty]$ be a terminal stopping time in the sense that for every pair (t_1, t_2), $0 \leqslant t_1 < t_2 < \infty$, the event $\{t_1 < \tau \leqslant t_2\}$ is $\mathcal{F}_{t_2}^{t_1}$-measurable. Then the process $(t_1, t_2) \mapsto M(t_1, t_2)$, $0 \leqslant t_1 \leqslant t_2 < \infty$, defined by $M(t_1, t_2) = \mathbf{1} - \mathbf{1}_{\{t_1 < \tau \leqslant t_2\}}$ is a multiplicative process. If τ is a time-homogeneous terminal stopping time, then the process $t \mapsto \mathbf{1}_{\{\tau > t\}}$ is a multiplicative process. This fact from the observation that $s + \tau \circ \vartheta_s = \tau$ $\mathbb{P}_x$-almost surely on the event $\{\tau > s\}$: the latter is just the notion of (time-homogeneous) terminal stopping time. Examples of terminal stopping times are first entry and first hitting times of Borel subsets; penetration times are terminal stopping times. If a Markov process like (9.14) in Theorem 9.2 is present, then for $\mathcal{F}_{t_2}^{t_1}$ we may take the universal completion of the right closure of $\sigma(X(s) : t_1 \leqslant s \leqslant t_2)$. An important property which is used here is the fact that the corresponding Markov process has right-continuous paths (or orbits).

Let μ be a Radon measure on E which is σ-finite. In the following proposition we write $\mathbb{E}_\mu[F] = \int_E \mathbb{E}_x[F] \, d\mu(x)$ for any random variable $F : \Omega \to \mathbb{R}$ for which $\mathbb{E}_\mu[|F|] \, d\mu(x) < \infty$, or $F \geqslant 0$.

The existence of a σ-finite Radon measure under the recurrence hypotheses of Theorem 10.12 will be proved in Chapter 10.

Proposition 9.7. *Let the hypotheses and notation be as in Theorem 9.2, except that the invariant measure μ is not necessarily finite, but is allowed to be a σ-finite Radon measure. Let $(A(t))_{t \geqslant 0}$ and $(B(t))_{t \geqslant 0}$ be additive processes such that $\mathbb{E}_\mu[|A(1)|] < \infty$, and $0 < \mathbb{E}_\mu[B(1)] < \infty$. Then the* *equality*

$$\mathbb{P}_x\left[\lim_{t\to\infty}\frac{A(t)}{B(t)}=\frac{\mathbb{E}_\mu\left[A(1)\right]}{\mathbb{E}_\mu\left[B(1)\right]}\right]=1 \tag{9.76}$$

holds for μ-almost all $x \in E$. Moreover, the equality

$$\lim_{t\to\infty}\frac{\mathbb{E}_x\left[A(t)\right]}{\mathbb{E}_x\left[B(t)\right]}=\frac{\mathbb{E}_\mu\left[A(1)\right]}{\mathbb{E}_\mu\left[B(1)\right]} \tag{9.77}$$

holds for μ-almost all $x \in E$.

Remark 9.6. Let $t \mapsto A(t)$ be an additive process, and let μ be an invariant measure. Then the function $t \mapsto \mathbb{E}_\mu\left[A(t)\right]$ is linear in t, and so there exists a constant $k\left(A(\cdot)\right)$ such that $\mathbb{E}_\mu\left[A(t)\right] = tk\left(A(\cdot)\right)$. In other words, the equality in (9.77) implies:

$$\lim_{t\to\infty}\frac{\mathbb{E}_x\left[A(t)\right]}{\mathbb{E}_x\left[B(t)\right]}=\frac{\mathbb{E}_\mu\left[A_1\right]}{\mathbb{E}_\mu\left[B_1\right]}=\frac{k\left(A(\cdot)\right)}{k\left(B(\cdot)\right)}. \tag{9.78}$$

Below the proof of Proposition 9.7 is copied from the proof of Proposition 5.5 in [Seidler (1997)]. Some of the techniques are borrowed from Azema et al [Azéma *et al.* (1967)] section II.2, and the Chacon-Ornstein theorem as exhibited in Krengel [Krengel (1985)]: see Theorem 9.9. In the proof of Proposition 9.7 we need some definitions and terminology which we collect next.

Definition 9.7. Let μ be a σ-finite Borel measure on E. An operator $S : L^1(E,\mu) \to L^1(E,\mu)$ is called a positive operator or positivity preserving operator if $f \geqslant 0$ μ-almost everywhere implies $Sf \geqslant 0$ μ-almost everywhere. It is called a contraction operator if $\int_E |Sf|\,d\mu \leqslant \int_E |f|\,d\mu$ for all $f \in L^1(E,\mu)$. The operator $S^* : L^\infty(E,\mu) \to L^\infty(E,\mu)$ is defined by the equality $\int_E (Sf)\,g\,d\mu = \int_E f\,(S^*g)\,d\mu$ for all $f \in L^1(E,\mu)$ and all $g \in L^\infty(E,\mu)$. Since the measure μ is σ-finite the dual space of $L^1(E,\mu)$ is identified with $L^\infty(E,\mu)$. Notice that S^*g_n decreases to 0, whenever g_n decreases pointwise to 0. A function (or better a class of functions) $h \in L^\infty(E,\mu)$ is called harmonic if $S^*h = h$. A harmonic function is also called an S^*-invariant function or just invariant function: see Definition 9.9. A non-negative function h for which $h \geqslant S^*h$ is called superharmonic. A superharmonic function h is called strictly superharmonic on a subset A of E provided $h > S^*h$ on A. A subset $B \in \mathcal{E}$ is called S-absorbing if $Sf \in L^1(B,\mu)$ for all $f \in L^1(B,\mu)$. The system $((E,\mathcal{E},\mu),S)$ is called a dynamical system if $\int Sf\,d\mu = \int f d\mu$ for all $f \in L^1(E,\mu)$. In the same manner the system $((\Omega,\mathcal{F},\mathcal{P}_\mu),S)$ is called a dynamical system if $\int SZ\,d\mathbb{P}_\mu = \int Z\,d\mathbb{P}_\mu$ for $Z \in L^1(\Omega,\mathcal{F},\mathbb{P}_\mu)$.

The definition of dynamical system will be used for invariant measures for Markov processes with state space E and sample path space Ω: i.e. the process

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in E}, (X(t), t \geqslant 0), (E, \mathcal{E})\}$$

as exhibited in (9.51) in Theorem 9.4 and in Theorem 9.2. In this case $\mathbb{P}_\mu$ is defined by $\mathbb{P}_\mu[A] = \int \mathbb{P}_x[A]\, d\mu(x)$, $A \in \mathcal{F}$.

The following decomposition theorems, 9.6, 9.7, and 9.8 can be found in [Krengel (1985)] theorems 1.3, 1.5, and 1.6 in Chapter 3. The decomposition of E into a conservative part C and its complement D a dissipative part is called the Hopf decomposition. The results are also applicable for the measure space $(\Omega, \mathcal{F}, \mathbb{P}_\mu)$ instead of $(E, \mathcal{E}, \mu)$ where μ is a σ-invariant Radon measure on $\mathcal{E}$. Since $\mu(C) = \mathbb{P}_\mu[X(0) \in C]$, $C \in \mathcal{E}$, the measure μ is σ-finite if and only if $\mathbb{P}_\mu$ is so. These theorems will be applied with $Sf(x) = P_a f(x) = e^{aL} f(x) = \mathbb{E}_x[f(X(a))]$ with $f \in L^1(E, \mu)$ where μ is an invariant measure for the operator P_a, $\int P_a f\, d\mu = \int f d\mu$, $f \in L^1(E, \mu)$, and the operator $SZ = Z \circ \vartheta_a$, $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}_\mu)$. In both cases $a > 0$. Since our underlying Markov process is recurrent it follows that the conservative subset C coincides with E μ-almost everywhere. For the recurrence properties of our Markov process see Theorem 10.16 and Theorem 10.9 in Chapter 10.

Theorem 9.6. *Let S be a positive contraction on $L^1(E, \mu)$. Then there exists a decomposition of E into disjoint sets C and D which are determined uniquely modulo μ by:*

- *(C1) If h is superharmonic, then $h = S^* h$ on C;*
- *(D1) There exists a bounded superharmonic function h_0 which is strictly superharmonic on D.*

The function h_0 may be constructed in such a way that $\lim_{n\to\infty} (S^*)^n h_0 = 0$ on D, and $h_0 = 0$ on C.

Theorem 9.7. *Let S be a positive contraction on $L^1(E, \mu)$. Let C and D be the subsets as described in Theorem 9.6. Then the decomposition of E into the disjoint sets C and D is also determined uniquely modulo μ by:*

- *(C2) For all $h \geqslant 0$ $h \in L^\infty(E, \mu)$ the sum $\sum_{n=0}^\infty (S^*)^n h = \infty$ on the subset $C \bigcap \{\sum_{n=0}^\infty (S^*)^n h > 0\}$;*
- *(D2) There exists a function $h_D \in L^\infty(E, \mu)$, $h_D \geqslant 0$, for which $\{h_D > 0\} = D$, and $\sum_{n=0}^\infty (S^*)^n h \leqslant \mathbf{1}$.*

Theorem 9.8. *Let S be a positive contraction on $L^1(E,\mu)$. Let C and D be the subsets as described in Theorem 9.6. Then the decomposition of E into the disjoint sets C and D is also determined uniquely modulo μ by:*

(C3) For all functions $f \geqslant 0$, $f \in L^1(E,\mu)$, the sum $\sum_{n=0}^{\infty} S^n f = \infty$ on the subset $C \bigcap \left\{\sum_{n=0}^{\infty} S^n f > 0\right\}$;
(D3) For all functions $f \geqslant 0$, $f \in L^1(E,\mu)$, $\sum_{n=0}^{\infty} S^n f < \infty$ on D.

Definition 9.8. Let $S : L^1(E,\mu) \to L^1(E,\mu)$ be a positive contraction. The decomposition of E into the disjoint union of C and $D = E \backslash D$ as determined by one of the theorems 9.6, 9.7 or 9.8 is called the Hopf decomposition of E relative to S. The subset C is called the conservative part of S, and D is called the dissipative part. The operator S is called conservative if $\mu(E\backslash C) = 0$.

The following theorem is a version of the Chacon-Ornstein theorem. For more details see e.g. [Petersen (1989)], [Krengel (1985)], [Foguel (1980)], and [Neveu (1979)]. Let B be a Borel subset of E and put $H_B f = I_B \sum_{j=0}^{\infty} \left(S I_{E\backslash B}\right)^j f$, where $I_B f = \mathbf{1}_B f$ and $f \in L^1(E,\mu)$. If $B = E$, then $H_B f = f$, $f \in L^1(E,\mu)$.

Theorem 9.9. *Let S be a positive contraction on $L^1(E,\mu)$. Let $0 \leqslant f, g$ be functions in $L^1(E,\mu)$. Then the limit $\lim_{n\to\infty} \dfrac{\sum_{k=0}^{n} S^k f}{\sum_{k=0}^{n} S^k g}$ converges to a finite μ-almost everywhere on the set $\left\{x \in E : \sum_{k=0}^{\infty} S^k g(x) > 0\right\}$. On the conservative part C the limit can be identified with a quotient of the form $\dfrac{Qf}{Qg}$ where Qf is the μ-conditional expectation on the σ-field of S^*-invariant subsets of $H_C f$. On the dissipative part D the limit of this quotient can be identified as a quotient of fixed numbers.*

Let $\mathcal{J}$ be the σ-field of S^*-invariant subsets. Fix $f \in L^1(E,\mu)$. In Theorem 9.9 the function Qf can be identified with the Radon-Nikodym derivative of the measure $H_C f\mu$ restricted to $\mathcal{I}$ with respect to the measure μ also confined to $\mathcal{J}$: $Qf = \dfrac{d\left(H_C f\mu \mid \mathcal{J}\right)}{d\left(\mu \mid \mathcal{J}\right)}$. Here $H_C f\mu$ is the measure which has density $H_C f$ relative to μ. This Radon-Nikodym derivative is often called the μ-conditional expectation of $H_C f$ on $\mathcal{J}$.

The following result can be found in Skorohod: see Theorem 5 and its corollary in Chapter 1, §1, of [Skorokhod (1989)].

Theorem 9.10. *Let μ be a non-zero invariant σ-finite measure on $\mathcal{E}$, and* *put* $\mathbb{P}_\mu[A] = \int \mathbb{P}_x[A]\, d\mu(x)$, $A \in \mathcal{F}$. *Then* $\mathbb{P}_\mu$ *is a* σ*-finite measure on* $\mathcal{F}$. *Put* $\mathcal{I} = \left\{R \in \mathcal{F} : \mathbb{P}_\mu\left[\vartheta_1^{-1} R \triangle R\right] = 0\right\}$. *Assume that all probability measures of the form* $B \mapsto P(1, x, B)$, $x \in E$, *are equivalent. Then the following assertions are true:*

(a) *If* $B \in \mathcal{E}$ *is such that* $P(1, x, B) = 1$ *for* μ*-almost all* $x \in B$, *then either* $\mu(B) = 0$ *or* $\mu(E \backslash B) = 0$.
(b) *Suppose that the random variable* $Y \in L^1(\Omega, \mathcal{F}, \mu)$ *possesses the following property:* $Y = Y \circ \vartheta_1$ $\mathbb{P}_\mu$*-almost everywhere. Then for all* $n \in \mathbb{N}$ *the equality*
$$\mathbb{E}_{X(n)}[Y] = \mathbb{E}_x\left[Y \;\middle|\; \mathcal{F}_n\right] \tag{9.79}$$
holds $\mathbb{P}_x$*-almost surely for* μ*-almost all* $x \in E$. *The equality* $\mathbb{E}_x[Y] = \mathbb{E}_x\left[\mathbb{E}_{X(n)}[Y]\right]$ *holds* μ*-almost everywhere for all* $n \in \mathbb{N}$, *including* $n = 0$. *Moreover, the equality* $Y = \mathbb{E}_{X(0)}[Y]$ *holds* $\mathbb{P}_\mu$*-almost everywhere.*
(c) *Events in* $\mathcal{I}$ *are* $\mathbb{P}_\mu$*-trivial in the sense that either* $\mathbb{P}_\mu[R] = 0$ *or* $\mathbb{P}_\mu[\Omega \backslash R] = 0$.
(d) *Let* $Y \in L^1(\Omega, \mathcal{F}, \mu)$ *be a random variable with the property that* $Y = Y \circ \vartheta_1$ $\mathbb{P}_\mu$*-almost everywhere. Then* Y *is zero* $\mathbb{P}_\mu$*-almost everywhere if* $\mu(E) = \infty$, *and constant* $\mathbb{P}_\mu$*-almost everywhere if* μ *is finite.*

Remark 9.7. Assertion (b) of Theorem 9.10 only uses the invariance of the σ-finite measure μ. The others also use the fact that all measures of the form $B \mapsto P(1, x, B)$, $B \in \mathcal{E}$, $x \in E$, are equivalent.

Proof. Let $(A_n)_{n\in\mathbb{N}}$ be an increasing sequence in $\mathcal{E}$ such that $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$ and $E = \bigcup_{n\in\mathbb{N}} A_n$. Put $\Omega_n = \{X(0) \in A_n\}$. Then $\Omega_n \subset \Omega_{n+1}$, $\Omega = \bigcup_{n\in\mathbb{N}} \Omega_n$, and $\mathbb{P}_\mu[\Omega_n] = \mu(A_n)$. This shows that the measure $\mathbb{P}_\mu$ is σ-finite.

(a). Let $B \in \mathcal{E}$ be such that $P(1, x, B) = 1$ for μ-almost all $x \in B$, and assume that $\mu(B) > 0$. Then $P(1, x, E \backslash B) = 0$ for μ-almost all $x \in B$. Since $\mu(B) > 0$, there exists at least one $x_0 \in B$ such that $P(1, x_0, E \backslash B) = 0$. Since all measures of the form $C \mapsto P(1, x, C)$, $x \in E$, are equivalent, it follows that $P(1, y, E \backslash B) = 0$ for μ-almost all $y \in E$. Consequently, $\mathbb{P}_\mu[E \backslash B] = 0$. This proves Assertion (a).

(b). First observe that by the Markov property, and by the invariance of the measure μ we have

$$\begin{aligned}
&\mathbb{E}_\mu\left[|Y \circ \vartheta_n - Y \circ \vartheta_{n+1}|\right] = \mathbb{E}_\mu\left[\mathbb{E}_{X(n)}\left[|Y - Y \circ \vartheta_1|\right]\right] \\
&= \mathbb{E}_\mu\left[\mathbb{E}_{X(n-1)}\left[|Y - Y \circ \vartheta_1|\right]\right] = \mathbb{E}_\mu\left[\mathbb{E}_{X(0)}\left[|Y - Y \circ \vartheta_1|\right]\right]
\end{aligned}$$

$$= \mathbb{E}_\mu \left[|Y - Y \circ \vartheta_1| \right]. \tag{9.80}$$

From (9.80) we infer by induction that $Y = Y \circ \vartheta_n$ $\mathbb{P}_\mu$-almost everywhere for all $n \in \mathbb{N}$. Let $A \in \mathcal{F}_n$, and consider the (in-)equalities:

$$\begin{aligned} 0 = \mathbb{E}_\mu \left[|Y - Y \circ \vartheta_n| \right] &= \int_E \mathbb{E}_x \left[|Y - Y \circ \vartheta_n| \right] d\mu(x) \\ &\geqslant \int_E \left| \mathbb{E}_x \left[(Y - Y \circ \vartheta_n) \mathbf{1}_A \right] \right| d\mu(x) \\ &= \int_E \left| \mathbb{E}_x \left[Y \mathbf{1}_A \right] - \mathbb{E}_x \left[\mathbb{E}_x \left[Y \circ \vartheta_n \mid \mathcal{F}_n \right] \mathbf{1}_A \right] \right| d\mu(x) \end{aligned}$$

(Markov property)

$$\begin{aligned} &= \int_E \left| \mathbb{E}_x \left[Y \mathbf{1}_A \right] - \mathbb{E}_x \left[\mathbb{E}_{X(n)} \left[Y \right] \mathbf{1}_A \right] \right| d\mu(x) \\ &= \int_E \left| \mathbb{E}_x \left[\left(Y - \mathbb{E}_{X(n)} \left[Y \right] \right) \mathbf{1}_A \right] \right| d\mu(x). \end{aligned} \tag{9.81}$$

From (9.81) we see that $\mathbb{E}_x \left[Y \mid \mathcal{F}_n \right] = \mathbb{E}_{X(n)} \left[Y \right]$ $\mathbb{P}_x$-almost surely for μ-almost all $x \in E$, and $n \in \mathbb{N}$, $n \geqslant 1$. The latter is the same as saying the for μ-almost all $x \in E$ the process $n \mapsto \mathbb{E}_{X(n)} \left[Y \right]$ is a $\mathbb{P}_x$-martingale. It also proves (9.79) in Assertion (b) for $n \in \mathbb{N}$, $n \geqslant 1$. By putting $A = \Omega$ in (9.81) we infer $\mathbb{E}_x \left[Y \right] = \mathbb{E}_x \left[\mathbb{E}_{X(n)} \left[Y \right] \right]$ μ-almost everywhere on E. In order to complete the proof of Assertion (b) we need to show the equality $Y = \mathbb{E}_{X(0)} \left[Y \right]$ $\mathbb{P}_\mu$-almost everywhere. Since the process $n \mapsto \mathbb{E}_{X(n)} \left[Y \right]$ is a $\mathbb{P}_x$-martingale we see that its limit exists $\mathbb{P}_x$-almost surely for μ-almost all $x \in E$. Moreover this limit is $\mathbb{P}_\mu$-almost surely equal to Y. We shall prove that this limit is also equal to $\mathbb{E}_{X(0)} \left[Y \right]$ $\mathbb{P}_\mu$-almost everywhere. Therefore we consider for $-\infty < \alpha < \beta < \infty$ the quantity

$$\mathbb{P}_\mu \left[\alpha < Y < \beta \right] = \lim_{n \to \infty} \mathbb{P}_\mu \left[\alpha < \mathbb{E}_{X(n)} \left[Y \right] < \beta \right]$$

(employ the invariance of μ)

$$\begin{aligned} &= \lim_{n \to \infty} \mathbb{P}_\mu \left[\alpha < \mathbb{E}_{X(0)} \left[Y \right] < \beta \right] \\ &= \mathbb{P}_\mu \left[\alpha < \mathbb{E}_{X(0)} \left[Y \right] < \beta \right]. \end{aligned} \tag{9.82}$$

Since $-\infty < \alpha < \beta < \infty$ are arbitrary, the equalities in (9.82) yield $Y = \mathbb{E}_{X(0)} \left[Y \right]$ $\mathbb{P}_\mu$-almost everywhere. This completes the proof of assertion (b).

(c). Let R be a member of $\mathcal{I}$. Then $\mathbf{1}_R = \mathbf{1}_R \circ \vartheta_1$ $\mathbb{P}_\mu$-almost everywhere. Since μ is an invariant measure we also get $\mathbf{1}_R = \mathbf{1}_R \circ \vartheta_n$ $\mathbb{P}_\mu$-almost everywhere for all $n \in \mathbb{N}$. In addition, an application of assertion (b) yields

$$\mathbf{1}_R = \mathbb{E}_{X(0)} \left[\mathbf{1}_R \right] = \mathbb{P}_{X(0)} \left[R \right] \quad \mathbb{P}_\mu\text{-almost everywhere.} \tag{9.83}$$

Put $B = \{x \in E : \mathbb{P}_x [R] = 1\}$. From (9.83) we see $R = \{X(0) \in B\}$, and hence $\mathbf{1}_R = \mathbf{1}_B (X(0))$, $\mathbb{P}_\mu$-almost everywhere. We also see that $\Omega \backslash R = \{X(0) \in E \backslash B\}$. It follows that $\mu(B) = \mathbb{P}_\mu [R]$ and $\mu (E \backslash B) = \mathbb{P}_\mu [\Omega \backslash R]$. Assume $P_\mu [R] = \mu(B) > 0$. Let $x_0 \in B$ be any point for which $\mathbf{1}_R \circ \vartheta_1 = \mathbf{1}_R$ $\mathbb{P}_{x_0}$-almost surely. Since R belongs to $\mathcal{I}$, and $\mu(B) > 0$, the latter equality holds for μ-almost all $x_0 \in B$. Then

$$\begin{aligned} P(1, x_0, B) &= P(1, x_0, \{x \in E : \mathbb{P}_x [R] = 1\}) \\ &= \mathbb{P}_{x_0} \left[\mathbb{P}_{X(1)} [R] = 1\right] = \mathbb{P}_{x_0} \left[\mathbb{P}_{X(0)} [R] \circ \vartheta_1 = 1\right] \\ &= \mathbb{P}_{x_0} [\mathbf{1}_R \circ \vartheta_1 = 1] = \mathbb{P}_{x_0} [\mathbf{1}_R = 1] = \mathbb{P}_{x_0} [R] = 1 \end{aligned} \tag{9.84}$$

where in the final equality of (9.84) we used the fact that $x_0 \in B$. It follows that for μ-almost all $x_0 \in B$ we have $P(1, x_0, B) = 1$. From assertion (a) we then infer that $\mu (E \backslash B) = 0$. But then $\mathbb{P}_\mu [\Omega \backslash R] = 0$. This shows Assertion (c).

(d). Let $Y \in L^1 (\Omega, \mathcal{F}, \mu)$ be such that $Y = Y \circ \vartheta_1$ $\mathbb{P}_\mu$-almost everywhere. Since $Y \wedge 0 = (Y \circ \vartheta_1) \wedge 0 = (Y \wedge 0) \circ \vartheta_1$ $\mathbb{P}_\mu$-almost everywhere we assume without loss of generality that $Y \geqslant 0$. Let m be the μ-essential supremum of Y. If $m = \infty$, then we consider the $\mathbb{P}_\mu$-invariant event $\{Y > n\}$. Observe that $\mathbb{P}_\mu [Y > n] > 0$, and so by (c) its complement has $\mathbb{P}_\mu$-measure zero. In other words $Y > n$ $\mathbb{P}_\mu$-almost everywhere. Since this is true for all $n \in \mathbb{N}$ we see $Y = \infty$, $\mathbb{P}_\mu$-almost everywhere. Since μ is non-zero and $Y \in L^1 (\Omega, \mathcal{F}, \mathbb{P}_\mu)$ this is a contradiction. So we assume that $m < \infty$. If $\xi < m$ we have $\mathbb{P}_\mu [Y > \xi] > 0$, and hence by (c) and $\mathbb{P}_\mu$-invariance of the event $\{Y > \xi\}$ it follows that $\mathbb{P}_\mu [Y \leqslant \xi] = 0$. Thus we see $Y \geqslant \xi$ $\mathbb{P}_\mu$-almost everywhere on Ω. Since $\xi < m$ is arbitrary we obtain $Y \geqslant m$ $\mathbb{P}_\mu$-almost everywhere on Ω. By definition we have $m \geqslant Y$ $\mathbb{P}_\mu$-almost everywhere on Ω. Consequently $Y = m$ $\mathbb{P}_\mu$-almost everywhere on Ω. If $\mu(E) = \mathbb{P}_\mu [X(0) \in E] = \mathbb{P}_\mu [\Omega] = \infty$, then necessarily $Y = m = 0$ $\mathbb{P}_\mu$-almost everywhere. If $\mu(B) < \infty$ we see that $Y = m$ $\mathbb{P}_\mu$-almost everywhere, where m is a finite constant.

This completes the proof of Theorem 9.10. □

Definition 9.9. Subsets $B \in \mathcal{E}$ with the property that $P(t, x, B) = 1$ for μ-almost all $x \in B$ are called μ-invariant subsets. Subsets $B \in \mathcal{E}$ with the property that $P(t, x, B) = 1$ for all $x \in B$ are called invariant subsets. Events $A \in \mathcal{F}$ with the property that $R = \vartheta_t^{-1} R$ are called invariant events; events with the property that $\mathbb{P}_\mu \left[\vartheta_t^{-1} R \triangle R\right] = 0$ for all $t > 0$ are called $\mathbb{P}_\mu$-invariant events. For the notion of tail σ-fields in $\mathcal{F}$ the reader is referred to Definition 9.3.

Notice that a subset B is μ-invariant if and only if $P(t,x,A\bigcap B) = P(t,x,A)$ for μ-almost all $x \in B$. If P_t denotes the operator $P_t h(x) = \int P(t,x,dy)h(y)$, $h \in L^1(E,\mu)$, then B is μ-invariant if and only if $P_t^*\mathbf{1}_B(x) = 1$, $x \in B$. Moreover, a function $h \in L^\infty(E,\mu)$, $h \geqslant 0$ μ-almost everywhere is harmonic or invariant if $P_t^* h = h$ on $\{h > 0\}$ for all $t > 0$: compare with Definition 9.7.

Proof. [Proof of Proposition 9.7.] We begin by putting

$$M = \left\{ \sum_{j=1}^{\infty} B(1) \circ \vartheta_j = \infty \right\},$$

and note that M is obviously ϑ_1-invariant, so either

$$\mathbb{P}_\mu[M] = 0, \quad \text{or} \quad \mathbb{P}_\mu[\Omega \backslash M] = 0.$$

This fact follows from Theorem 9.10 assertion (c). But $\mathbb{E}_\mu[B(1)] > 0$ implies

$$\mathbb{E}_\mu\left[\sum_{j=1}^{\infty} B(1)\circ\vartheta_j\right] = \mathbb{E}_\mu\left[\sum_{j=1}^{\infty} B(1)\right] = \infty,$$

(the measure $\mathbb{P}_\mu$ is ϑ_1-invariant), so the possibility $\mathbb{P}_\mu[M] = 0$ is excluded. Define a positive contraction $T : L^1(\mathbb{P}_\mu) \to L^1(\mathbb{P}_\mu)$, by $u \mapsto u \circ \vartheta_1$. Let V be a Borel subset such that $\mu(V) < \infty$ and which satisfies

$$\mathbb{P}_x\left[\int_0^\infty \mathbf{1}_V(X(s))\,ds = \infty\right] = 1 \quad \text{for all } x \in E, \tag{9.85}$$

and set $v = \int_0^1 \mathbf{1}_V(X(s))\,ds$. The existence of such a set V is guaranteed by Proposition 9.6 and the fact that the measure μ is a regular Radon measure. Then $v \in L^1(\mathbb{P}_\mu)$ because

$$\mathbb{E}_\mu[v] = \int v d\mathbb{P}_\mu = \int_0^1 \int_E P(s,y,V)\,d\mu(y)ds = \mu(V) < \infty,$$

and hence we have

$$\sum_{j=0}^{\infty} T^j v = \int_0^\infty \mathbf{1}_V(X(s))\,ds = \infty, \quad \mathbb{P}_\mu\text{-almost everywhere.} \tag{9.86}$$

This means that the operator T is conservative (cf. [Krengel (1985)], Theorem 1.6 Chapter 3: see Theorem 9.8 and Definition 9.8) and by the Chacon-Ornstein theorem (see Theorem 9.9) and the Neveu-Chacon identification theorem (see e.g. [Krengel (1985)], theorems 2.7 and 3.4 Chapter 3) we obtain that

$$\lim_{n\to\infty}\frac{\sum_{j=0}^{n}T^{j}A(1)}{\sum_{j=0}^{n}T^{j}B(1)}=\lim_{n\to\infty}\frac{A(n)}{B(n)}=\frac{\mathbb{E}_{\mu}\left[A(1)\right]}{\mathbb{E}_{\mu}\left[B(1)\right]}\quad \mathbb{P}_{\mu}\text{-almost everywhere.} \tag{9.87}$$

Now, exactly the same procedure as in [Azéma *et al.* (1967)] applies, and hence we see that the discrete time result (9.87) implies that $\mathbb{P}_{\mu}\left[\Omega\backslash C\right]=0$, where

$$C=\left\{\lim_{t\to\infty}\frac{A(t)}{B(t)}=\frac{\mathbb{E}_{\mu}\left[A(1)\right]}{\mathbb{E}_{\mu}\left[B(1)\right]}\right\}. \tag{9.88}$$

For details the reader is referred to Proposition 9.9 in section 9.2. So there exists $N\in\mathcal{E}$, $\mu(N)=0$ and $\mathbb{P}_{x}\left[C\right]=1$ for all $x\notin N$. Let $y\in E$ be arbitrary, then

$$\begin{aligned}\mathbb{P}_{y}[C]&=\mathbb{E}_{y}\left[\mathbf{1}_{C}\circ\vartheta_{1}\right]=\mathbb{E}_{y}\left[\mathbb{E}_{X(1)}\left[\mathbf{1}_{C}\right]\right]=\int_{E}\mathbb{P}_{z}[C]P\left(1,y,dz\right)\\&=\int_{E\backslash N}\mathbb{P}_{z}[C]P\left(1,y,dz\right)=1,\end{aligned}$$

since $P\left(1,y,N\right)=0$ by the fact that all measures $B\mapsto P\left(t,y,B\right)$, $B\in\mathcal{E}$, $(t,y)\in(0,\infty)\times E$, are equivalent, and $\int_{E}P\left(1,z,N\right)\,d\mu(z)=\mu(N)=0$.

This proves equality (9.76) in Proposition 9.7.

In order to prove equality (9.77) we introduce a positivity preserving contraction mapping $S:L^{1}\left(E,\mu\right)\to L^{1}\left(E,\mu\right)$ by setting $Sf(x)=\int_{E}f(z)P\left(1,x,dz\right)$, $f\in L^{1}\left(E,\mu\right)$. As in the proof of equality (9.76) let V be a Borel subset of E such that $\mu(V)<\infty$ and such that (9.85) is satisfied. Put $h(x)=\mathbb{E}_{x}\left[v\right]=\int_{0}^{1}P\left(s,x,V\right)\,ds$. Then $h\in L^{1}\left(E,\mu\right)$ and

$$\sum_{n=0}^{\infty}S^{n}h(x)=\int_{0}^{\infty}P\left(s,x,V\right)ds=\infty,\quad x\in E. \tag{9.89}$$

Hence the contraction mapping S is conservative. Let $\mathcal{A}$ be the σ-field of S-absorbing subsets: see Definition 9.7 (cf. [Krengel (1985)], Definition 1.7 Chapter 3). The equivalence of transition probabilities of our Markov process in (9.14) implies easily that μ is trivial on $\mathcal{A}$: i.e. $\mu(A)=0$ or $\mu\left(E\backslash A\right)=0$ for all $A\in\mathcal{A}$. Define the functions f and g by $f(x)=\mathbb{E}_{x}\left[A(1)\right]$, and $g(x)=\mathbb{E}_{x}\left[B(1)\right]$. Then f, $g\in L^{1}\left(E,\mu\right)$, $g\geqslant 0$, and hence again by the Chacon-Ornstein theorem (see Theorem 9.9) we have

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N}S^{n}f}{\sum_{n=0}^{N}S^{n}g}=\frac{\int_{E}f\,d\mu}{\int_{E}g\,d\mu}=\frac{\mathbb{E}_{\mu}\left[A(1)\right]}{\mathbb{E}_{\mu}\left[B(1)\right]}$$

$$\mu\text{-almost everywhere on } \left\{x \in E : \sum_{n=0}^{\infty} S^n g(x) > 0\right\}. \tag{9.90}$$

Notice that

$$Sf(x) = \int_E \mathbb{E}_z [A_1] P(1, x, dz) = \mathbb{E}_x \left[\mathbb{E}_{X(1)} [A(1)]\right] = \mathbb{E}_x [A(1) \circ \vartheta_1]. \tag{9.91}$$

We know that $\mathbb{P}_\mu \left[\sum_{j=0}^{\infty} B(1) \circ \vartheta_j < \infty\right] = 0$, and thus we also have

$$\mathbb{P}_x \left[\sum_{j=0}^{\infty} B(1) \circ \vartheta_j < \infty\right] = 0 \text{ for } \mu\text{-almost all } x \in E.$$

Therefore

$$\sum_{j=0}^{\infty} S^j g(x) = \sum_{j=0}^{\infty} \mathbb{E}_x [B(1) \circ \vartheta_j] = \infty \tag{9.92}$$

for μ-almost all $x \in E$, and hence (9.90) yields that the equality

$$\lim_{n\to\infty} \frac{\mathbb{E}_x [A(n)]}{\mathbb{E}_x [B(n)]} = \frac{\mathbb{E}_\mu [A(1)]}{\mathbb{E}_\mu [B(1)]}$$

holds for μ-almost all $x \in E$. Again, the proof can be completed as in [Azéma *et al.* (1967)]. For details see Proposition 9.9 in section 9.2.

Altogether this completes the proof of Proposition 9.7. □

In Corollary 9.3 we establish the uniqueness of σ-finite invariant measures. However, notice that, up to a multiplicative constant, the equality in (9.77) is a consequence of the uniqueness of σ-finite invariant measures: see the proof of Lemma 9.7.

Corollary 9.3. *Let the assumptions and notation be as in Proposition 9.7. Let μ_1 and μ_2 be two σ-finite non-trivial invariant measures. Then up to a finite strictly positive constant these two measures coincide.*

Proof. Let $(B(t))_{t\geqslant 0}$ be an additive process such that $0 < \mathbb{E}_{\mu_1} [B(1)] < \infty$ and $0 < \mathbb{E}_{\mu_2} [B(1)] < \infty$. Let $f \in L^1 (E, \mu_1) \bigcap L^1 (E, \mu_2)$. From Proposition 9.7 we infer that

$$\frac{\int_E f\, d\mu_1}{\mathbb{E}_{\mu_1} [B(1)]} = \frac{\int_E f\, d\mu_2}{\mathbb{E}_{\mu_2} [B(1)]}$$

and hence

$$\int_E f\, d\mu_2 = \frac{\mathbb{E}_{\mu_2} [B(1)]}{\mathbb{E}_{\mu_1} [B(1)]} \int_E f\, d\mu_1. \tag{9.93}$$

The asserted uniqueness follows from (9.93) and the density of $L^1 (E, \mu_1) \bigcap L^1 (E, \mu_2)$ in either $L^1 (E, \mu_1)$ or $L^1 (E, \mu_2)$. The proof of Corollary 9.3 is now complete. □

Corollary 9.4. *Let the assumptions and notation be as in Proposition 9.7. Then the following assertions are valid:*

(a) The Markov process in (9.14) is μ-Harris recurrent, that is the equality $\mathbb{P}_x\left[\int_0^\infty \mathbf{1}_A(X(s))\,ds = \infty\right] = 1$ holds for all $x \in E$ and for all $A \in \mathcal{E}$ for which $\mu(A) > 0$.

(b) Suppose $\mu(E) = \infty$. Then $\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X(s))\,ds = 0$ $\mathbb{P}_x$-almost surely for all $x \in E$ and all $f \in L^1(E,\mu)$.

(c) Suppose $\mu(E) < \infty$. Then $\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X(s))\,ds = \frac{\int_E f\,d\mu}{\mu(E)}$ $\mathbb{P}_x$-almost surely for all $x \in E$ and all $f \in L^1(E,\mu)$.

Proof. (a). Assume that there exists $z \in E$ and $A \in \mathcal{E}$ with $\mu(A) > 0$ such that $\int_0^\infty \mathbf{1}_A(X(s))\,ds < \infty$ on an event Ω' with $\mathbb{P}_z(\Omega') > 0$. We will arrive at a contradiction. Let $V \in \mathcal{E}$ be such that $\mu(V) < \infty$ and (9.85) are satisfied. Then by assumption

$$\lim_{t\to\infty} \frac{\int_0^t \mathbf{1}_A(X(s))\,ds}{\int_0^t \mathbf{1}_V(X(s))\,ds} = 0 \quad \mathbb{P}_x\text{-almost surely on } \Omega'. \tag{9.94}$$

However, according to (9.76) in Proposition 9.7 the limit in (9.94) should be

$$\begin{aligned}\frac{\mathbb{E}_\mu\left[\int_0^1 \mathbf{1}_A(X(s))\,ds\right]}{\mathbb{E}_\mu\left[\int_0^1 \mathbf{1}_V(X(s))\,ds\right]} &= \frac{\int_0^1\int_E \mathbb{E}_x\left[\mathbf{1}_A(X(s))\right]d\mu(x)\,ds}{\int_0^1\int_E \mathbb{E}_x\left[\mathbf{1}_V(X(s))\right]d\mu(x)\,ds}\\ &= \frac{\int_0^1\int_E P(s,x,A)\,d\mu(x)\,ds}{\int_0^1\int_E P(s,x,V)\,d\mu(x)\,ds} = \frac{\mu(A)}{\mu(V)}.\end{aligned} \tag{9.95}$$

Since $\mu(A) > 0$ and $\mu(V) < \infty$ the equality in (9.95) leads to a contradiction. Hence assertion (a) follows.

(b). Fix $\varepsilon > 0$, $x \in E$, and $f \in L^1(E,\mu)$, $f \geqslant 0$. Since $\mu(E) = \infty$ and μ is σ-finite there exists a subset $B \in \mathcal{E}$ such that $\mu(B) < \infty$, and $\frac{\int_E f\,d\mu}{\mu(B)} < \frac{\varepsilon}{2}$. By (9.76) of Proposition 9.7 there exists a random variable t_ε which is $\mathbb{P}_x$-almost surely finite such that

$$\frac{\int_0^t f(X(s))\,ds}{\int_0^t \mathbf{1}_B(X(s))\,ds} \leqslant \frac{\int_E f\,d\mu}{\mu(B)} + \frac{\varepsilon}{2} \leqslant \frac{\varepsilon}{2} = \varepsilon \quad \text{for all } t \geqslant t_\varepsilon. \tag{9.96}$$

Since

$$\frac{\int_0^t f(X(s))\,ds}{t} \leqslant \frac{\int_0^t f(X(s))\,ds}{\int_0^t \mathbf{1}_B(X(s))\,ds}$$

Assertion (b) follows from (9.96).

(c). This assertion is an immediate consequence of Proposition 9.7.

Altogether this completes the proof of Corollary 9.4. □

In the proof of Proposition 9.8 below we need Theorem 10.2 of Chapter 10. It is taken from Jamison and Orey [Jamison *et al.* (1965)] Theorem 1, and Lemma 3. A result like Lemma 3 can also be found in Meyn and Tweedie [Meyn and Tweedie (1993b)] Theorem 18.1.2. The result is called Orey's convergence theorem. Let ν be a measure on $\mathcal{E}$. The measures $P(t)^*\nu$, $t \geqslant 0$, are defined by $B \mapsto \int_E P(t,x,B)\,d\nu(x)$. The following proposition should be compared with Theorem 10.2. For the notion of "Harris recurrence" of Markov chains see Definition 10.2 in Chapter 10. The definitions 9.4 and 9.5 contain the corresponding notions for continuous time Markov processes.

Proposition 9.8. *Let the hypotheses and notation be as in Proposition 9.7. Let μ be a σ-finite invariant measure. Then the Markov chain $(X(n) : n \in \mathbb{N})$ is μ-Harris recurrent, and*

$$\lim_{t\to\infty} \operatorname{Var}\left(P(t)^*\mu_2 - P(t)^*\mu_1\right) = 0 \tag{9.97}$$

for all probability measures μ_1 and μ_2 on $\mathcal{E}$.

Remark 9.8. The proof of Proposition 9.8 yields a slightly stronger result than (9.97). In fact by (9.113) we have

$$\lim_{t\to\infty} \iint\limits_{E\times E} \operatorname{Var}\left(P(t,x,\cdot) - P(t,y,\cdot)\right) d\mu_1(x)\,d\mu_2(y) = 0. \tag{9.98}$$

It is clear that the result in (9.98) is stronger than (9.97). Moreover, the function $t \mapsto \iint_{E\times E} \operatorname{Var}\left(P(t,x,\cdot) - P(t,y,\cdot)\right) d\mu_1(x)\,d\mu_2(y)$ decreases, so that (9.98) follows once we know it for any sequence $(t_n : n \in \mathbb{N})$ which increases to ∞. Put

$$(\alpha R(\alpha))^n\,\mathbf{1}_B(x) = \alpha \int_0^\infty \frac{(\alpha t)^{n-1}}{(n-1)!} e^{-\alpha t} P(t,x,B)\,dt = \mathbb{P}_x \otimes \pi_0\left[X(T_n) \in B\right]$$

where $\alpha > 0$, $n \in \mathbb{N}$, and $x \in E$. Here the process $(T_n : n \in \mathbb{N})$ consists of the jump process of a Poisson process

$$\left\{(\Lambda, \mathcal{G}, \pi_t)_{t\geqslant 0}, (N(t), t \geqslant 0), \left(\vartheta_t^P : t \geqslant 0\right), [0,\infty)\right\}$$

which has intensity α_0, and which is independent of the strong Markov process

$$\left\{(\Omega, \mathcal{F}, \mathbb{P}_x)_{x\in E}, (X(t), t \geqslant 0), (\vartheta_t, t \geqslant 0), (E, \mathcal{E})\right\}.$$

For more details see (10.140), (10.143), and Lemma (10.42) in Chapter 10. Again let μ_1 and μ_2 probability measures on $\mathcal{E}$. Fix $\alpha_0 > 0$. Then, under the conditions of Proposition 9.8 we have

$$\begin{aligned}
&\lim_{\alpha\downarrow 0} \iint_{E\times E} \operatorname{Var}\left(\alpha R(\alpha)\mathbf{1}_{(\cdot)}(x) - \alpha R(\alpha)\mathbf{1}_{(\cdot)}(y)\right) d\mu_1(x)\, d\mu_2(y) \\
&= \lim_{n\to\infty} \iint_{E\times E} \operatorname{Var}\left(\left(\alpha_0 R\left(\alpha_0\right)\right)^n \mathbf{1}_{(\cdot)}(x) - \left(\alpha_0 R\left(\alpha_0\right)\right)^n \mathbf{1}_{(\cdot)}(y)\right) d\mu_1(x)\, d\mu_2(y) \\
&= 0. \qquad (9.99)
\end{aligned}$$

Proof. [Proof of Proposition 9.8.] The proof follows the lines of Duflo et al [Duflo and Revuz (1969)], which reduces the proof to the corresponding result for discrete time Markov chains: see Jamison and Orey [Jamison and Orey (1967)]. In the formal sense in [Duflo and Revuz (1969)] the authors only consider a locally compact state space, but changing to a Polish space does not affect their proof. Nevertheless we will repeat the arguments.

First notice that the process $(X(n) : n \in \mathbb{N})$ is a Markov chain with transition probability function $(x, B) \mapsto P(1, x, B)$, $(x, B) \in E \times \mathcal{E}$. Since all these measures are equivalent the chain $(X(n) : n \in \mathbb{N})$ is aperiodic: see Proposition 10.1 in Chapter 10 and the comments preceding it. We will check that it is Harris recurrent. Let μ be the invariant measure, and choose an arbitrary $B \in \mathcal{E}$ for which $0 < \mu(B) < \infty$, and put

$$R = \left\{ \sum_{n=1}^{\infty} \mathbf{1}_B\left(X(n)\right) = \infty \right\}. \qquad (9.100)$$

Then $\vartheta_1^{-1} R \subset R$, i.e. the event R is ϑ_1-invariant. Hence we either have $\mathbb{P}_\mu\left[R\right] = 0$ or $\mathbb{P}_\mu\left[\Omega\backslash R\right] = 0$: see Theorem 9.10 assertion (c). Then the mapping $T : L^1\left(\Omega, \mathcal{F}, \mathbb{P}_\mu\right) \to L^1\left(\Omega, \mathcal{F}, \mathbb{P}_\mu\right)$ defined by $Tu = u \circ \vartheta_1$ is a conservative positive contraction, and $\mathbf{1}_B\left(X(1)\right) \in L^1\left(\Omega, \mathcal{F}, \mathbb{P}_\mu\right)$. Hence

$$\sum_{n=0}^{\infty} T^n \mathbf{1}_B\left(X(1)\right) = \sum_{n=1}^{\infty} \mathbf{1}_B\left(X(n)\right) \in \{0, \infty\} \quad \mathbb{P}_\mu\text{-almost everywhere.} \qquad (9.101)$$

If $\mathbb{P}_\mu\left[R\right] = 0$, then

$$0 = \mathbb{E}_\mu\left[\sum_{n=1}^{\infty} \mathbf{1}_B\left(X(n)\right)\right] = \sum_{n=1}^{\infty} \int_E P\left(n, x, B\right) d\mu(x) = \sum_{n=1}^{\infty} \mu(B). \qquad (9.102)$$

Since $\mu(B) > 0$ the equality in (9.102) is a contradiction. It follows that $\mathbb{P}_\mu\left[\Omega\backslash R\right] = 0$, and hence there exists a subset $N \in \mathcal{E}$ such that $\mu(N) = 0$

and $\mathbb{P}_y[\Omega\backslash R] = 0$ for all $y \in E\backslash N$. So that for $y \in E\backslash N$ we have $\mathbb{P}_y[R] = 1$. Since $\mu(N) = 0$, and μ is invariant we see that $\int_E P(1,z,N)\,d\mu(z) = \mu(N) = 0$, and hence $(1,z,N) = 0$ for μ-almost all $z \in E$. Since μ is non-trivial, this implies that $P(1,z,N) = 0$ for at least one $z \in E$. Since all the measures $B \mapsto P(1,z,B)$, $z \in E$, are equivalent, we see that $P(1,z,N) = 0$ for all $z \in E$. So that furthermore, for $x \in E$ arbitrary, we infer

$$\mathbb{P}_x[R] = \mathbb{E}_x[\mathbf{1}_R \circ \vartheta_1] = \mathbb{E}_x\left[\mathbb{E}_x\left[\mathbf{1}_R \circ \vartheta_1 \,\middle|\, \mathcal{F}_1\right]\right]$$

(Markov property)

$$= \mathbb{E}_x\left[\mathbb{E}_{X(1)}[\mathbf{1}_R]\right] = \int_E \mathbb{P}_y[R]\,P(1,x,dy)$$

(employ $P(1,x,N) = 0$)

$$= \int_{E\backslash N} \mathbb{P}_y[R]\,P(1,x,dy)$$

(for $y \in E\backslash N$ the equality $\mathbb{P}_y[R] = 1$ holds)

$$= \int_{E\backslash N} P(1,x,dy) = P(1,x,E\backslash N) = P(1,x,E) = 1. \tag{9.103}$$

From (9.103) we get $\mathbb{P}_x[R] = 1$ for all $x \in E$. Consequently, the Markov chain $(X(n) : n \in \mathbb{N})$ is Harris recurrent and aperiodic: see Definition 10.2 and Proposition 10.1 in Chapter 10 and the comments preceding it. From Theorem 10.2, which is Orey's convergence theorem, we obtain

$$\lim_{n\to\infty} \operatorname{Var}\left(P(n,x,\cdot) - P(n,y,\cdot)\right) = 0 \quad \text{for all } x,\ y \in E. \tag{9.104}$$

Next our aim is to establish the triviality of the tail σ-field of the Markov process $(X(t) : t \geqslant 0)$. For the notion of tail σ field see Definition 9.3. Let $A \in \mathcal{I}$, the tail σ-field. Then for every $t \geqslant 0$ there exists a tail event $A_t \in \mathcal{I}$ such that $\mathbf{1}_A = \mathbf{1}_{A_t} \circ \vartheta_t$ (see Definition 9.3). So for $x \in E$ we have

$$\begin{aligned}\mathbb{P}_x[A] &= \mathbb{E}_x[\mathbf{1}_A] = \mathbb{E}_x[\mathbf{1}_{A_t} \circ \vartheta_t] = \mathbb{E}_x\left[\mathbb{E}_x\left[\mathbf{1}_{A_t} \circ \vartheta_t \,\middle|\, \mathcal{F}_t\right]\right] \\ &= \mathbb{E}_x\left[\mathbb{E}_{X(t)}[\mathbf{1}_{A_t} \circ \vartheta_t]\right] = \int_E \mathbb{P}_z[A_t]\,P(t,x,dz)\,. \end{aligned} \tag{9.105}$$

By taking $t = n \to \infty$ in (9.104) and employing (9.105) we see that $\mathbb{P}_x[A] = \mathbb{P}_y[A]$, x, $y \in E$, and $A \in \mathcal{I}$. Since $A_t \in \mathcal{I}$ we see that the function $x \mapsto \mathbb{P}_x[A_t]$ is constant. From (9.105) it follows that this constant equals the constant function $x \mapsto \mathbb{P}_x[A]$. By the martingale convergence theorem we see

$$\mathbb{P}_x[A] = \mathbb{P}_{X(n)}[A] = \mathbb{P}_x\left[A \,\middle|\, \mathcal{F}_n\right] \overset{n\to\infty}{\longrightarrow} \mathbf{1}_A \quad \mathbb{P}_x\text{-almost surely.} \tag{9.106}$$

Equality (9.106) implies that either $\mathbb{P}_x[A] = 1$ for all $x \in E$ or that $\mathbb{P}_x[A] = 0$ for all $x \in E$. This proves the triviality of the tail σ-field $\mathcal{I}$. In order to complete the proof of Proposition 9.8 we proceed as in the proof of Theorem II.4 of [Duflo and Revuz (1969)], who follow Blackwell and Freedman [Blackwell and Freedman (1964)] Theorem 2.

Put $\mathcal{F}^t = \vartheta^{-1}\mathcal{F}$. Then the arguments of Duflo and Revuz read as follows. First, let $B \in \mathcal{F}$ and m a probability measure on $\mathcal{E}$. Then we have

$$\begin{aligned}\mathbb{P}_m\left[A\bigcap B\right] - \mathbb{P}_m[A]\,\mathbb{P}_m[B] &= \int_A (\mathbf{1}_B - \mathbb{P}_m[B])\, d\mathbb{P}_m \\ &= \int_A \left(\mathbb{P}_m\left[\mathbf{1}_B \mid \mathcal{F}^t\right] - \mathbb{P}_m[B]\right) d\mathbb{P}_m,\end{aligned}$$

and hence

$$\begin{aligned}\sup_{A\in\mathcal{F}^t}\left|\mathbb{P}_m\left[A\bigcap B\right] - \mathbb{P}_m[A]\,\mathbb{P}_m[B]\right| &\leqslant \sup_{A\in\mathcal{F}^t}\int_A \left|\mathbb{P}_m\left[B \mid \mathcal{F}^t\right] - \mathbb{P}_m[B]\right| d\mathbb{P}_m \\ &= \int_\Omega \left|\mathbb{P}_m\left[B \mid \mathcal{F}^t\right] - \mathbb{P}_m[B]\right| d\mathbb{P}_m. \qquad (9.107)\end{aligned}$$

By the backward martingale convergence theorem (see e.g. [Doob (1953)] Theorem 4.2) the limit

$$\lim_{n\to\infty}\mathbb{P}_m\left[B \mid \mathcal{F}^{t_n}\right] = \lim_{n\to\infty}\int_E \mathbb{P}_x\left[B \mid \mathcal{F}^{t_n}\right] dm(x) \qquad (9.108)$$

exists $\mathbb{P}_m$-almost surely and in $L^1(\Omega,\mathcal{F},\mathbb{P}_m)$ for all sequences $(t_n)_{n\in\mathbb{N}}$ which increase to ∞. The limit in (9.108) is measurable relative to the tail σ-field $\mathcal{I}$. Since the tail σ-field is trivial, this means that the limit $\lim_{t\to\infty}\mathbb{P}_m\left[B \mid \mathcal{F}^t\right] = \mathbb{P}_m[B]$, $\mathbb{P}_m$-almost surely. From (9.107) we see that

$$\lim_{t\to\infty}\sup_{A\in\mathcal{F}^t}\left|\mathbb{P}_m\left[A\bigcap B\right] - \mathbb{P}_m[A]\,\mathbb{P}_m[B]\right| = 0. \qquad (9.109)$$

Let $(x,y) \in E\times E$, and $A_0 \in \mathcal{E}$. We apply (9.109) with $m = \frac{1}{2}(\delta_x + \delta_y)$, $A = \{X(t) \in A_0\}$, and $B = \{X(0) = x\}$ or $B = \{X(0) = y\}$. Then we obtain

$$\lim_{t\to\infty}\sup_{A_0\in\mathcal{E}}|P(t,x,A_0) - P(t,y,A_0)| = 0. \qquad (9.110)$$

Since

$$\operatorname{Var}(P(t,x,\cdot) - P(t,y,\cdot)) \leqslant 2\sup_{A\in\mathcal{E}}|P(t,x,A) - P(t,y,A)| \qquad (9.111)$$

equality (9.110) implies

$$\lim_{t\to\infty}\operatorname{Var}(P(t,x,\cdot) - P(t,y,\cdot)) = 0. \qquad (9.112)$$

Next let μ_1 and μ_2 be two probability measures on E. Then

$$
\begin{aligned}
&\operatorname{Var}\left(\int_E P(t,x,\cdot)\,d\mu_1(x) - \int_E P(t,y,\cdot)\,d\mu_2(x)\right) \\
&= \operatorname{Var}\left(\iint\limits_{E\times E} \left(P(t,x,\cdot) - P(t,y,\cdot)\right) d\mu_1(x)\,d\mu_2(y)\right) \\
&\leqslant \iint\limits_{E\times E} \operatorname{Var}\left(P(t,x,\cdot) - P(t,y,\cdot)\right) d\mu_1(x)\,d\mu_2(y), \qquad (9.113)
\end{aligned}
$$

and hence by equality (9.110) and inequality (9.113) we obtain

$$
\lim_{t\to\infty} \operatorname{Var}\left(\int_E P(t,x,\cdot)\,d\mu_1(x) - \int_E P(t,y,\cdot)\,d\mu_2(x)\right) = 0. \qquad (9.114)
$$

Since equality (9.114) is equivalent to (9.97) this completes the proof of Proposition 9.8. □

The following theorem is another version of Theorem 9.5.

Theorem 9.11. *Let the Markov process have right-continuous sample paths, be strong Feller, and irreducible. Let A be a recurrent compact subset of the state space E, and $K \subset E$ any compact subset. Then there exists an closed neighborhood K_ε with K in its interior such that for $h > 0$*

$$
\sup_{x\in E} \int_0^\infty \mathbb{P}_x\left[X(t) \in K_\varepsilon,\ h + \tau_A \circ \vartheta_h > t\right] dt < \infty. \qquad (9.115)
$$

For the notion of strong Feller property see Definitions 2.5 and 2.16.

Proof. Without loss of generality may and shall assume that $K \supset A$. Otherwise replace K by $K \bigcup A$. From the arguments following (10.222) in the proof of Theorem 10.12 we see that there exists $\varepsilon > 0$ such that

$$
\sup_{y\in E} \mathbb{E}_y\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(\rho)\right) d\rho\right] < \infty \qquad (9.116)
$$

where K_ε is an ε-neighborhood of K. This completes the proof of Theorem 9.11. □

By definition we have

$$
R_A(0)f(x) = \int_0^\infty \mathbb{E}_x\left[F\left(X(\rho)\right),\ \tau_A > \rho\right] d\rho = \mathbb{E}_x\left[\int_0^{\tau_A} f\left(X(\rho)\right) d\rho\right] \qquad (9.117)
$$

for those Borel measurable function f for which the integrals in (9.117) exist. As a corollary to 9.11 we have the following result. The proof follows by observing that $\tau_A \leqslant h+\tau_A\circ\vartheta_h$ and the definition of $R_A(0)f$: see (9.117).

Corollary 9.5. *Let the hypotheses and notation be as in Theorem 9.11. In addition, let K be a compact subset of E and $h > 0$. Then there exists a bounded function $f \in C_b(E)$, $\mathbf{1}_K \leqslant f \leqslant \mathbf{1}$, such that*

$$\begin{aligned}\sup_{y\in E} R_A(0)f(y) &= \sup_{y\in E}\mathbb{E}_y\left[\int_0^{\tau_A} f\left(X(\rho)\right)\,d\rho\right]\\ &\leqslant \sup_{y\in E}\mathbb{E}_y\left[\int_0^{h+\tau_A\circ\vartheta_h} f\left(X(\rho)\right)\,d\rho\right] < \infty. \end{aligned}\tag{9.118}$$

Let f be as in (9.118). Then there exists a constant C_f such that for all $g \in C_b(E)$ the following inequality holds:

$$\sup_{y\in E} R_A(0)\left(|g|\,f\right)(y) \leqslant C_f\,\|g\|_\infty\,. \tag{9.119}$$

Proof. Let K_ε be as in Theorem 9.11 and choose $f \in C_b(E)$ in such a way that $\mathbf{1}_K \leqslant f \leqslant \mathbf{1}_{K_\varepsilon}$. Then f satisfies (9.118), and (9.119) is satisfied with C_f given by the right-hand side of (9.118). Altogether this completes the proof of Corollary 9.5. □

9.2 Some ergodic theorems

In this section we prove some results which are relevant to finish the arguments in the proof of Proposition 9.7. In particular we want to prove that $\mathbb{P}_\mu\left[\Omega\backslash C\right] = 0$, where the invariant subset C is given in (9.88), i.e.

$$C = \left\{\lim_{t\to\infty}\frac{A(t)}{B(t)} = \frac{\mathbb{E}_\mu\left[A(1)\right]}{\mathbb{E}_\mu\left[B(1)\right]}\right\}. \tag{9.120}$$

We also want to prove that

$$\lim_{t\to\infty}\frac{\mathbb{E}_x\left[A(t)\right]}{\mathbb{E}_x\left[B(t)\right]} = \frac{\mathbb{E}_\mu\left[A(1)\right]}{\mathbb{E}_\mu\left[B(1)\right]} \tag{9.121}$$

holds for μ-almost all $x \in E$. Both proofs can be found in [Azéma *et al.* (1967)]. However, since we want to make the present book elf-contained we will give an independent proof. The proofs will be based on the theorem of Chacon-Ornstein for discrete dynamical systems: see Theorem 9.9.

Lemma 9.5. *Let μ be a σ-finite invariant measure, and let $Z : \Omega \to \mathbb{R}$ be a bounded random variable such that $Z\circ\vartheta_a = Z$ $\mathbb{P}_\mu$-almost surely for all $a > 0$. Then the variable Z is constant $\mathbb{P}_\mu$-almost surely.*

Proof. Let ν be a probability measure which is equivalent with μ. Then the process

$$\mathbb{E}_{X(t)}\left[Z\right] = \mathbb{E}_{\nu}\left[Z \circ \vartheta_t \mid \mathcal{F}_t\right] = \mathbb{E}_{\nu}\left[Z \mid \mathcal{F}_t\right] \tag{9.122}$$

is a martingale with $Z = \lim_{t\to\infty} \mathbb{E}_{\nu}\left[Z \mid \mathcal{F}_t\right] = \lim_{t\to\infty} \mathbb{E}_{X(t)}\left[Z\right]$. Then we introduce for given $k \in \mathbb{R}$ the subset $F_k = \{y \in E : \mathbb{E}_y[Z] \geqslant k\}$. Then there are two possibilities either $\nu(F_k) > 0$ or $\nu(E\backslash F_k) = 1$. If $\nu(F_k) > 0$, then we have $\mathbb{P}_{\nu}\left[\int_0^{\infty} \mathbf{1}_{F_k}(X(s))\, ds = \infty\right] = 1$, and hence $\limsup_{t\to\infty} \mathbb{E}_{X(t)}\left[Z\right] \geqslant k$ $\mathbb{P}_{\nu}$-almost surely. Consequently, $Z = \lim_{t\to\infty} \mathbb{E}_{X(t)}[Z] \geqslant k$ $\mathbb{P}_{\nu}$-almost surely. In the other case, $\nu(E\backslash F_k) = 1$, we will get $Z \leqslant k$. It follows that Z is a constant $\mathbb{P}_{\mu}$-almost surely.

This completes the proof of Lemma 9.5. □

Lemma 9.6. *Let the hypotheses and notation be as in Proposition 9.7. Let C be as in (9.120). Suppose that the limit $\lim_{t\to\infty} \dfrac{A(t)}{B(t)}$ exists $\mathbb{P}_{\mu}$-almost surely. Then $\mathbb{P}_x\left[C\right] = 1$ for μ-almost all $x \in E$.*

Proof. We consider the dynamical system

$$\{(\Omega, \mathcal{F}, \mathbb{P}_{\mu}) : (X(t), t \geqslant 0)\}$$

together with the countable dynamical subsystems (skeletons)

$$\{(\Omega, \mathcal{F}, \mathbb{P}_{\mu}) : (X(na), n \in \mathbb{N})\}, \quad a > 0.$$

By assumption, the limit $Z := \lim_{t\to\infty} \dfrac{A(t)}{B(t)}$ exists $\mathbb{P}_{\mu}$-almost surely. Then we have $Z \circ \vartheta_a = Z$ $\mathbb{P}_{\mu}$-almost surely for all $a > 0$. By Lemma 9.5 we see that $Z = C$ $\mathbb{P}_{\mu}$-almost surely, where C is a real constant. Denote $\mathcal{I}_a$ the σ-field invariant corresponding to the operator $T_a : Z \mapsto Z \circ \vartheta_a$, $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}_{\mu})$. Then by the Chacon-Ornstein theorem (Theorem 9.9) we know that

$$C_{A,B} = \lim_{t\to\infty} \frac{A(t)}{B(t)} = \lim_{n\to\infty} \frac{A(na)}{B(na)} = \frac{\mathbb{E}_{\mu}\left[A(a) \mid \mathcal{I}_a\right]}{\mathbb{E}_{\mu}\left[B(a) \mid \mathcal{I}_a\right]} \tag{9.123}$$

where $\mathbb{E}_{\mu}\left[A(a) \mid \mathcal{I}_a\right]$ denotes conditional expectation on the σ-field $\mathcal{I}_a$ relative to the measure $\mathbb{P}_{\mu}$ which is not necessarily a probability measure. Nevertheless the notion "conditional expectation" relative to such measures also makes sense. From (9.123) we deduce:

$$\mathbb{E}_{\mu}\left[A(a) \mid \mathcal{I}_a\right] = C_{A,B}\mathbb{E}_{\mu}\left[B(a) \mid \mathcal{I}_a\right]. \tag{9.124}$$

For the Chacon-Ornstein theorem and the Neveu-Chacon identification theorem see e.g. [Krengel (1985)], theorems 2.7 and 3.4 in Chapter 3, [Petersen

(1989)] Theorem 8.1, [Foguel (1980)] §1.3, and [Neveu (1979)]. So that by integrating the left-hand side and the right-hand side of (9.124) relative to the measure $\mathbb{P}_\mu$ we obtain:

$$\mathbb{E}_\mu\left[A(a)\right] = C_{A,B}\mathbb{E}_\mu\left[B(a)\right]. \tag{9.125}$$

Since the expression in the right-hand side of (9.125) does not depend on a we get $\mathbb{E}_\mu\left[A(1)\right] = C_{A,B}\mathbb{E}_\mu\left[B(1)\right]$. Let C be the event in (9.120). Then $\mathbf{1}_C \circ \vartheta_a = \mathbf{1}_C$ $\mathbb{P}_\mu$-almost surely for all $a > 0$. By Lemma 9.5 it follows that $\mathbf{1}_C = \mathbf{1}$ $\mathbb{P}_\mu$-almost surely. Since, for $a \geqslant 0$,

$$\int \left(1 - \mathbb{E}_y\left[\mathbf{1}_C \circ \vartheta_a\right]\right) d\mu = \mathbb{E}_\mu\left[\mathbf{1} - \mathbf{1}_C \circ \vartheta_a\right] = \mathbb{E}_\mu\left[\mathbf{1} - \mathbf{1}_C\right] = 0, \tag{9.126}$$

we get $\mathbb{P}_y[C] = 1$ for μ-almost all $y \in E$.

This completes the proof of Lemma 9.6. □

Let $a > 0$. In the proof of the following lemma and of equality (9.121) we need the following invariant σ-field on E:

$$\begin{aligned}\mathcal{J}_a &= \Big\{B \in \mathcal{E} : P\left(a, x, A \bigcap B\right) = P(a,x,A)\,\mathbf{1}_B(x) \\ &\qquad \text{for all } A \in \mathcal{E} \text{ and for } \mu\text{-almost all } x \in E\Big\} \\ &= \left\{B \in \mathcal{E} : P(a,x,B) = \mathbf{1}_B(x), \quad \text{for } \mu\text{-almost all } x \in E\right\}. \end{aligned} \tag{9.127}$$

The definition of $\mathcal{J}_a$ should be compared with the notion of μ-invariant subset in Definition 9.9. An application of the Chacon-Ornstein theorem to the dynamical system $\left\{\left((E,\mathcal{E},\mu), P_a\right)\right\}$, with $P_a f(x) = e^{aL}f(x) = \mathbb{E}_x\left[f\left(X(a)\right)\right] = \int f(y)P(a,x,dy)$, $f \in L^1(E,\mu)$, is that for all $a > 0$

$$\lim_{n\to\infty} \frac{\mathbb{E}_x\left[A(na)\right]}{\mathbb{E}_x\left[B(na)\right]} = \frac{Q_a\mathbb{E}_{(\cdot)}\left[A(a)\right](x)}{Q_a\mathbb{E}_{(\cdot)}\left[B(a)\right](x)} \tag{9.128}$$

for μ-almost all $x \in E$. For the Chacon-Ornstein theorem and the Neveu-Chacon identification theorem see e.g. [Krengel (1985)], theorems 2.7 and 3.4 in Chapter 3. Here Q_a is the μ-conditioning operator on the σ-field $\mathcal{J}_a$. In other words: if $h \in L^1(E,\mu)$, then $Q_a h$ is $\mathcal{J}_a$-measurable and $\int Q_a h(x) f(x)\, d\mu(x) = \int h(x) f(x)\, d\mu(x)$ for all bounded functions f which are $\mathcal{J}_a$ measurable. Of course, the measure μ is the invariant measure for the semigroup $\left\{e^{tL} : t \geqslant 0\right\}$. The operator Q_a has the following invariance property:

$$\int P_a f(x) Q_a h(x)\, d\mu(x) = \int f(x) Q_a h(x)\, d\mu(x), \;\; h \in L^1(E,\mu), \; f \in L^\infty(E,\mu). \tag{9.129}$$

Here $P_a f(x) = e^{aL} f(x) = \int P(a,x,dy)f(y)$, $f \in L^\infty(E,\mu)$. The equality in (9.129) can also be written as $P_a^* Q_a h = Q_a h$, $h \in L^1(E,\mu)$. It follows that if $Q_a h(x) = h(x)$ on the subset $\{h \neq 0\}$, then the measure $B \mapsto \int_B h(x) d\mu(x)$ is a P_a-invariant measure.

Lemma 9.7. *Let the hypotheses and notation be as in Proposition 9.7. Suppose that the limit* $\lim_{t\to\infty} \frac{\mathbb{E}_x[A(t)]}{\mathbb{E}_x[B(t)]}$ *exists for μ-almost all $x \in E$. Then* $\lim_{t\to\infty} \frac{\mathbb{E}_x[A(t)]}{\mathbb{E}_x[B(t)]} = \frac{\mathbb{E}_\mu[A(1)]}{\mathbb{E}_\mu[B(1)]}$ *for μ-almost all $x \in E$.*

Proof. Let $h \geqslant 0$ be a function which is bounded and which is measurable with respect to $\mathcal{J}_a$ for all $a > 0$. Then for $f \in L^1(E,\mu)$ we have by invariance of the function h

$$\int e^{aL} f(x)h(x)\, d\mu(x) = \int e^{aL} (fh)(x)\, d\mu(x) = \int f(x)h(x)\, d\mu(x). \quad (9.130)$$

Since the equality in (9.130) holds for all $a > 0$ and all $f \in L^1(E,\mu)$ we infer that the measure $h\mu$ is also an invariant measure. By uniqueness it follows that the function h is a constant μ-almost everywhere. Since for all $a > 0$ we have equality of the following limits

$$H_{A,B}(x) := \lim_{t\to\infty} \frac{\mathbb{E}_x[A(t)]}{\mathbb{E}_x[B(t)]} = \lim_{n\to\infty} \frac{\mathbb{E}_x[A(na)]}{\mathbb{E}_x[B(na)]} \quad (9.131)$$

the Chacon-Ornstein theorem implies that the function $x \mapsto H_{A,B}(x)$ in (9.131) is $\mathcal{J}_a$-measurable for all $a > 0$: see e.g. [Krengel (1985)], theorems 2.7 and 3.4 in Chapter 3. Since such functions are μ-almost everywhere constant, we infer that the function $H_{A,B}$ is μ-almost everywhere a constant $C_{A,B}$. So we see

$$Q_a \mathbb{E}_{(\cdot)}[A(a)](x) = C_{A,B} Q_a \mathbb{E}_{(\cdot)}[B(a)](x), \quad \text{for } \mu\text{-almost all } x \in E. \quad (9.132)$$

Since the constant function $\mathbf{1}$ is $\mathcal{J}_a$-measurable the equality in (9.132) yields:

$$\int \mathbb{E}_x[A(a)]\, d\mu(x) = \int Q_a \mathbb{E}_{(\cdot)}[A(a)](x)\, d\mu(x)$$
$$= C_{A,B} \int Q_a \mathbb{E}_{(\cdot)}[B(a)](x)\, d\mu(x) = C_{A,B} \int \mathbb{E}_x[B(a)](x)\, d\mu(x). \quad (9.133)$$

Since the quotient $\frac{\int \mathbb{E}_x[A(a)]\, d\mu(x)}{\int \mathbb{E}_x[B(a)]\, d\mu(x)}$ does not depend on $a > 0$ we obtain

$$H_{A,B} = C_{A,B} = \frac{\int \mathbb{E}_x[A(1)]\, d\mu(x)}{\int \mathbb{E}_x[B(1)]\, d\mu(x)}. \quad (9.134)$$

The equality in (9.134) completes the proof of Lemma 9.7. □

Proposition 9.9. *Let the hypotheses and notation be as in Proposition 9.7. Let C be the event in (9.120). Then the event $\mathbb{P}_x[C] = 1$ for μ-almost all $x \in E$. The equality in (9.121) holds for μ-almost all $x \in E$.*

Proof. Since we already proved the lemmas 9.5, 9.6 and 9.7 we only need to show that the following limits exist:

(1) in order to see that $\mathbb{P}_x[C] = 1$ for μ-almost all $x \in E$, with C as in (9.120) it is required that the limit $\lim_{t\to\infty} \frac{A(t)}{B(t)}$ exists $\mathbb{P}_\mu$-almost surely.
(2) in order that the equality in (9.121) holds we need the existence of the limit: $\lim_{t\to\infty} \frac{\mathbb{E}_x[A(t)]}{\mathbb{E}_x[B(t)]}$ for μ-almost all $x \in E$.

Let the additive process $B(t) \geqslant 0$ be such that $0 < \mathbb{P}_\mu[B(a)] < \infty$. Then

$$\begin{aligned}\mathbb{E}_x[B(\infty)] &= \sum_{n=0}^{\infty} \mathbb{E}_x[B((n+1)a) - B(na)] \\ &= \sum_{n=0}^{\infty} \mathbb{E}_x[B(a)\circ\vartheta_{na}] = \sum_{n=0}^{\infty} e^{naL}\mathbb{E}_{(\cdot)}[B(a)](x), \end{aligned} \tag{9.135}$$

and hence

$$\int \mathbb{E}_x[B(\infty)]\,d\mu(x) = \sum_{n=0}^{\infty}\int \mathbb{E}_x[B(a)]\,d\mu(x) = \infty. \tag{9.136}$$

From the recurrence property of the process X, the hypothesis that all measures of the form $B \mapsto P(t,x,B)$ are equivalent, and the equality in (9.136) we infer that $B(\infty) = \infty$ $\mathbb{P}_x$-almost surely for μ-almost all $x \in E$, and that $B(\infty) = \infty$ $\mathbb{P}_\mu$-almost surely: see the proof of Proposition 9.7, and see Theorem 9.8. Since $B(\infty) = \infty$ $\mathbb{P}_x$-almost surely for μ-almost all $x \in E$ and $\mathbb{P}_\mu$-almost surely, in both cases it is easy to see that the existence of these limits is guaranteed as soon as we know the existence of these limits by taking $B(t)$ of the form $B(t) = \int_0^t \mathbf{1}_F(X(s))\,ds$ where F is a Borel subset with $0 < \mu(F) < \infty$. Let $a > 0$ and $x \in E$, and put $P_a f(x) = e^{aL}f(x) = \mathbb{E}_x[f(X(a))] = e^{aL}f(x) = \int f(y)P(a,x,dy)$, f bounded and Borel measurable. In addition put $T_a Z = Z\circ\vartheta_a$, where Z is a random variable which is $\mathcal{F}$-measurable. We write $Z_a = \int_0^a \mathbf{1}_F(X(s))\,ds$ and $f_a(x) = \mathbb{E}_x\left[\int_0^a \mathbf{1}_F(X(s))\,ds\right] = \mathbb{E}_x[Z_a]$. Since the process X is recurrent we know that (see the proof of Proposition 9.7)

$$\sum_{n=0}^{\infty} T_a^n Z_a = \int_0^{\infty} \mathbf{1}_F(X(s))\,ds = \infty, \quad \mathbb{P}_\mu\text{-almost surely, and} \tag{9.137}$$

$$\sum_{n=0}^{\infty} P_a^n f_a(x) = \mathbb{E}_x \left[\int_0^{\infty} \mathbf{1}_F \left(X(s) \right) \, ds \right] = \infty \quad \text{for } \mu\text{-almost all } x \in E. \tag{9.138}$$

From the Chacon-Ornstein theorem it then follows that the limits in (1) and (2) above exist as long as we take $t = na$, $a > 0$, and let $n \in \mathbb{N}$ tend to ∞. But then these limits also exist when we let t tend to ∞.

These observations complete the proof of Proposition 9.9. □

9.3 Spectral gap

Next we return to problems and results concerning spectral gaps and related topics. This section is concluded with a proof of Theorem 9.1. We start with an introductory remark.

Remark 9.9. Of course, the estimate in (9.8) in Corollary 9.1 gives an interesting lower bound for gap(L) only in case $\lambda_{\min}(a) > 0$; we always have $\lambda_{\min}(a) \geqslant 0$. Condition (9.5) and the finiteness of $\overline{a}$ in Theorem 9.1 can be replaced by a Γ_2-condition, without violating the conclusion in (9.6). In fact a condition of the form $\Gamma_2\left(\overline{f}, f\right) \geqslant \gamma \Gamma_1\left(\overline{f}, f\right)$, $f \in \mathcal{A}$, yields a stronger result: see Theorem 9.18 and Example 9.1, Proposition 9.18 and the formulas (9.269) and (9.270). It is also noticed that in the presence of an operator L as described in (9.1), and the corresponding squared gradient operator

$$\Gamma_1\left(f, g\right)(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{i,j}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \tag{9.139}$$

the standard Euclidean distance is not the necessarily the "natural" distance for problems related to the presence of a spectral gap. In fact the more adapted distance d_L or d_{Γ_1} is probably given by the following formula:

$$d_L(x, y) = \sup\left\{ |f(x) - f(y)| : \Gamma_1\left(\overline{f}, f\right) \leqslant 1, \ f \in D\left(\Gamma_1\right) \right\}. \tag{9.140}$$

One of the tools used in estimates related to coupling methods is finding the correct metric on $\mathbb{R}^d \times \mathbb{R}^d$ which serves as a "prototype" estimate. The reader should compare this observation with comments and techniques used by Chen and Wang in e.g. [Chen and Wang (1997, 2000, 2003)]. In Lemma 9.8 below the standard Euclidean distance is used (like in [Chen and Wang (1997)]). In fact, it could be that it would be more appropriate to use the distance presented in (9.140). As remarked earlier, this technique might lead to geometric considerations related to Γ_2-calculus.

Remark 9.10. As noticed in the preface of this book recent applications of the Γ_2-condition to problems related to transportation costs can be found in recent work by [Gozlan (2008)], which in turn is related to [Gozlan (2007)] and [Gozlan and Léonard (2007)]. Gozlan also introduces so-called local logarithmic Sobolev inequalities. In addition he establishes a link with the large deviation principle; more particularly, he expresses the rate function in terms of the relative entropy or the mutual information $H(\mu \mid \nu) = -\int_E d\mu \log \frac{d\mu}{d\nu}$, also called the conditional Shannon information in the discrete setting, between probability measures μ and ν on E. Another name for this quantity is Kullback-Leibler distance; for more properties of this "distance" see e.g. [Kullback (1997)]. For a general theory concerning optimal transport see [Villani (2003, 2009)]. A new variational method of finding the rate function for the large deviation principle is used in [Budhiraja *et al.* (2008)].

The proof of Theorem 9.1 will be based on coupling arguments. In the present situation we will consider unique week solutions to the following stochastic differential equation in $\mathbb{R}^d \times \mathbb{R}^d$:

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} X(s) \\ Y(s) \end{pmatrix} + \int_0^t \begin{pmatrix} \sigma(s, X(s)) \\ \sigma(\rho, Y(\rho)) \end{pmatrix} dW(\rho) + \int_s^t \begin{pmatrix} b(\rho, X(\rho)) \\ b(\rho, Y(\rho)) \end{pmatrix} d\rho. \tag{9.141}$$

Of course this equation is a natural analog of an equation of the form

$$X(t) = X(s) + \int_s^t \sigma(\rho, X(\rho))\, dW(\rho) + \int_0^t b(\rho, X(\rho))\, d\rho. \tag{9.142}$$

In equation (9.141) we assume that the column vector $\begin{pmatrix} X(s) \\ Y(s) \end{pmatrix}$ can be prescribed, and in (9.142) we may prescribe $X(s)$. Let us introduce the coupling operator $\widetilde{L}$ as follows:

$$\begin{aligned} \widetilde{L}_s f(x,y) = {} & \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(s,x,x) \frac{\partial^2 f(x,y)}{\partial x_i \partial x_j} + \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(s,x,y) \frac{\partial^2 f(x,y)}{\partial x_i \partial y_j} \\ & + \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(s,y,x) \frac{\partial^2 f(x,y)}{\partial y_i \partial x_j} + \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(s,y,y) \frac{\partial^2 f(x,y)}{\partial y_i \partial y_j} \\ & + \sum_{i=1}^d b_i(s,x) \frac{\partial f(x,y)}{\partial x_i} + \sum_{i=1}^d b_i(s,y) \frac{\partial f(x,y)}{\partial y_i}. \end{aligned} \tag{9.143}$$

Here the matrix $a(s,x,y) = (a_{i,j}(s,x,y))_{i,j=1}^{d}$ is given by

$$a_{i,j}(s,x,y) = (\sigma(s,x)\sigma(s,y)^*)_{i,j} = \sum_{k=1}^{d} \sigma_{i,k}(s,x)\sigma_{j,k}(s,y).$$

It follows that the diffusion matrix $\widetilde{a}(s,x,y)$ of the operator $\widetilde{L}_s$ and the drift vector $\widetilde{b}(s,x,y)$ are given by respectively:

$$\begin{aligned}\widetilde{a}(s,x,y) &= \begin{pmatrix} \sigma(s,x)\sigma(s,x)^* & \sigma(s,x)\sigma(s,y)^* \\ \sigma(s,y)\sigma(s,x)^* & \sigma(s,y)\sigma(s,y)^* \end{pmatrix} \\ &= \begin{pmatrix} \sigma(s,x) & 0 \\ 0 & \sigma(s,y) \end{pmatrix} \begin{pmatrix} I_d & I_d \\ I_d & I_d \end{pmatrix} \begin{pmatrix} \sigma(s,x)^* & 0 \\ 0 & \sigma(s,y)^* \end{pmatrix}, \end{aligned} \tag{9.144}$$

and

$$\widetilde{b}(s,x,y) = \begin{pmatrix} b(s,x) \\ b(s,y) \end{pmatrix}. \tag{9.145}$$

Here I_d is the $d \times d$ identity matrix. Notice that

$$\begin{pmatrix} I_d & I_d \\ I_d & I_d \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \alpha & \beta \end{pmatrix} \begin{pmatrix} \alpha^* & \alpha^* \\ \beta^* & \beta^* \end{pmatrix},$$

where the $d \times d$ matrices α and β are chosen in such a way that $\alpha\alpha^* + \beta\beta^* = I_d$. The stochastic differential equation in (9.141) corresponds to the choice $\alpha = I_d$ (and $\beta = 0$). We also assume that the corresponding martingale problem is well-posed. In the present context the corresponding martingale problem reads as follows. For every pair $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, and $s \geqslant 0$, find a probability measure $\mathbb{P}_{s,x,y}$ on $C_b\left(\mathbb{R}^d \times \mathbb{R}^d\right)$ which makes the process

$$f\left(t, X(t), Y(t)\right) - f\left(s, X(s), Y(s)\right) - \int_s^t \widetilde{L} f\left(\rho, X(\rho), Y(\rho)\right) d\rho \tag{9.146}$$

a $\mathbb{P}_{s,x,y}$-martingale with respect to the filtration determined by Brownian motion $\{W(s) : s \geqslant 0\}$. Moreover, we want the probability measure $\mathbb{P}_{s,x,y}$ to be such that $\mathbb{P}_{s,x,y}\left[X(s) = x,\, Y(s) = y\right] = 1$. For more details on the martingale problem the reader is referred to e.g. Theorems 2.11 and 2.12. Saying that the martingale problem is well posed for the operator is equivalent to saying that the stochastic differential equation in (9.141) has unique weak solutions. If the coefficients $\sigma(s,x)$ and $b(s,x)$ are such that the equation in (9.141) has unique strong solutions, then it possesses unique weak solutions, and hence the martingale problem is well-posed for $\widetilde{L}$. For more details the reader is referred to §1.1 in Chapter 10. Let the pair $\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}$

be a unique weak solution solution to the coupled stochastic differential equation (9.141) starting at time s in $\begin{pmatrix} X(s) \\ Y(s) \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^d \times \mathrm{R}^d$. Then we define the stopping time τ by

$$\tau = \inf\{t > 0 : X(t) = Y(t)\},$$

if there exists $t \in (0,\infty)$ such that $X(t) = Y(t)$. If no such finite t exists, then we write $\tau = \infty$. The following theorem can be found in [Chen and Wang (1997)] as Theorem 3.1. Their proof uses an approximation argument. As Chen and Wang indicate, it is also a consequence of theorems 6.1.3, 8.1.3 (and 10.1.1) in [Stroock and Varadhan (2006)]. The reader should compare the result in Theorem 9.12 with Theorem 1.3.

Theorem 9.12. *Suppose that the martingale is well posed for the operator L, or what is equivalent, suppose that the pair $(\sigma(s,x), b(s,x))$ possesses unique weak solutions. Let $\mathbb{P}_{s,x,y}$ be the unique solution to the martingale problem staring at the pair (x,y). Then $X(t) = Y(t)$ $\mathbb{P}_{s,x,y}$-almost surely on the event $\{\tau \leqslant t\}$.*

The proof Theorem 9.12 will be given after Remark 9.11 below.

The following definition is taken from [Stroock and Varadhan (2006)] Chapter 8. The connection with the well-posedness of the martingale problem will be explained in §1.1 in Chapter 10. In particular we have that the martingale problem is well-posed for the operator L if the pair $(\sigma(t,y), b(t,y))$, $t \geqslant 0$, $y \in \mathbb{R}^d$, satisfies Itô's uniqueness condition from any point $(s,x) \in [0,\infty) \times \mathbb{E}^d$. In fact this is the theorem of Watanabe and Yamada [Watanabe and Yamada (1971)].

Definition 9.10. Let $(s,x) \in [0,\infty) \times \mathbb{R}^d$. The pair $(\sigma(t,y), b(t,y))$, $t \geqslant s$, $y \in \mathbb{R}^d$, is said to possess at most one weak solution from (s,x), if and only if for every probability space $(\Omega, \mathcal{F}, \mathbb{P})$, every non-decreasing family $\{\mathcal{F}_t : t \geqslant 0\}$ of sub-σ-fields of $\mathcal{F}$, and every triple $\beta : [0,\infty) \times \Omega \to \mathbb{R}^d$, $\xi : [0,\infty) \times \Omega \to \mathbb{R}^d$, and $\eta : [0,\infty) \times \Omega \to \mathbb{R}^d$ such that $(\Omega, \mathcal{F}_t, \mathbb{P}; \beta(t))$ is a d-dimensional Brownian motion, and the equations

$$\xi(t) = x + \int_s^{s\vee t} \sigma(\rho, \xi(\rho))\, d\beta(\rho) + \int_s^{s\vee t} b(\rho, \xi(\rho))\, d\rho, \quad t \geqslant 0,$$

and

$$\eta(t) = x + \int_s^{s\vee t} \sigma(\rho, \eta(\rho))\, d\beta(\rho) + \int_s^{s\vee t} b(\rho, \eta(\rho))\, d\rho, \quad t \geqslant 0,$$

hold $\mathbb{P}$-almost surely, then $\xi(t) = \eta(t)$ $\mathbb{P}$-almost surely. Instead of possessing a "unique weak solution from (s,x)", it is also customary to say that for the pair $(\sigma(s,x), b(s,x))$ the Itô's uniqueness condition is satisfied from (s,x) or after s starting from x.

The following definition specializes Theorem 2.11 and 2.12 to the case of the differential operator $L = (L(t);\, t \geqslant 0\}$ as exhibited in (9.1).

Definition 9.11. Let the operator L be given by (9.1), and let

$$\Omega = \left(\mathbb{R}^d\right)^{[0,\infty)}, \quad \text{and} \quad X(t)(\omega) = X(t,\omega) = \omega(t), \quad \omega \in \Omega, \ t \geqslant 0.$$

Put $\mathcal{F}_t^s = \sigma\left(X(\rho) : s \leqslant \rho \leqslant t\right)$, $0 \leqslant s \leqslant t < \infty$, and $\mathcal{F} = \sigma\left(X(s) : s \geqslant 0\right)$. The martingale problem is said to be well-posed for the operator L starting from $(s,x) \in [0,\infty) \times \mathbb{R}^d$ if there exists a unique probability measure $\mathbb{P}$ on $\mathcal{P}$ with the following properties:

(a) $\mathbb{P}\left[X(t) = x : 0 \leqslant t \leqslant s\right] = 1$.
(b) For every $f \in \bigcap_{s>0} D\left(L(s)\right) \bigcap C_0\left(\mathbb{R}^d\right)$ the process

$$t \mapsto f(X(t)) - f(X(s)) - \int_s^t L(s) f\left(X(s)\right) ds$$

is a $\mathbb{P}$ martingale with respect to the filtration $\left(\mathcal{F}_t^s\right)_{t \geqslant s}$.

Let $\left(\Omega, \mathcal{F}_t^s, \mathbb{P}\right)_{t \geqslant s}$ be a filtered probability space, and let $(t,\omega) \mapsto X(t,\omega)$ be a progressively measurable process. There are several equivalent formulations for the process X possessing properties (a) and (b) on some probability space. The reader is referred to e.g. Theorem 4.2.1 in [Stroock and Varadhan (2006)]. We begin by defining a progressively measurable process.

Definition 9.12. Let $\left(\Omega, \mathcal{F}_t^s\right)_{t \geqslant s}$ be a filtered space, and let $(E, \mathcal{E})$ be a measurable space. Let $X : [s,\infty) \times \Omega \to E$ be a processes (or just a function). The process X is called progressively measurable if for every t_1, t_2, with $s \leqslant t_1 < t_2 < \infty$, the function $X : [t_1, t_2] \times \Omega \to E$ is $\mathcal{B}_{[t_1,t_2]} \times \mathcal{F}_{t_2}$-$\mathcal{E}$-measurable. The symbol $\mathcal{B}_{[t_1,t_2]}$ stand for the Borel field of the interval $[t_1, t_2]$. If E is a topological space with Borel field $\mathcal{E}$, and if X is right-continuous, then X is progressively measurable relative to the filtration $\left(\Omega, \mathcal{F}_{t+}^s\right)_{t \geqslant s}$. Here $\mathcal{F}_{t+}^s = \bigcap_{\rho > t} \mathcal{F}_\rho^s$.

Definition 9.13. Let $\left(\Omega, \mathcal{F}_t^s, \mathbb{P}\right)_{t \geqslant s}$ be a filtered probability space and let the progressively measurable process $X(t)$ have the properties in (a) and (b) of Definition 9.11 relative to the present filtered probability space. Then

$X(t)$ is called an Itô process on $(\Omega, \mathcal{F}_t^s, \mathbb{P})_{t \geqslant s}$ with covariance matrix $a(t,x)$ and drift vector $b(t,x)$, $(t,x) \in [0,\infty) \times \mathbb{R}^d$.

In fact the same definition can be used if the coefficients $a(t)$ and $b(t)$ are processes which are progressively measurable.

The following theorem says that an Itô process after a stopping time is again an Itô process. It is the same as Theorem 6.1.3 in [Stroock and Varadhan (2006)]: S_d stands for the symmetric $d \times d$ matrices with real entries.

Theorem 9.13. *Let $(\Omega, \mathcal{F}_t^s, \mathbb{P})$ be a filtered probability space, and let $a : [s,\infty) \times \Omega \to S_d$, and $b[s,\infty) \to \mathbb{R}^d$ be bounded progressively measurable functions. Moreover, let $X : [s,\infty) \times \Omega \to \mathbb{R}^d$ be an Itô process with covariance a and drift b, and let $\tau : \Omega \to [s,\infty)$ be an $(\mathcal{F}_t^s)_{t \geqslant s}$-stopping time. Suppose that the process $t \mapsto X(t)$ is right-continuous and $\mathbb{P}$-almost surely continuous. Let $\omega \mapsto \mathbb{Q}_\omega$ be regular conditional probability distribution corresponding to the conditional probability: $A \mapsto \mathbb{P}\left[A \,\middle|\, \mathcal{F}_\tau^s\right]$. Then there exists a $\mathbb{P}$-null N set such that $t \mapsto X(t)$ is an Itô process on $[\tau(\omega),\infty)$ relative to Q_ω, $\omega \notin N$.*

For a proof of Theorem 9.13 we refer the reader to [Stroock and Varadhan (2006)]. The function $(\omega, A) \mapsto \mathbb{Q}_\omega(A)$, $\omega \in \Omega$, $A \in \mathcal{F}^s = \sigma\left(X(\rho) : \rho \geqslant s\right)$, possesses the following properties:

(a) For every $B \in \mathcal{F}^s$ the function $\omega \mapsto \mathbb{Q}_\omega\left[B\right]$ is $\mathcal{F}_\tau^s$-measurable;
(b) For every $A \in \mathcal{F}_\tau^s$ and $B \in \mathcal{F}^s$ the following equality holds:
$$\mathbb{P}\left[A \bigcap B\right] = \int_A \mathbb{Q}_\omega[B]\, d\mathbb{P}(\omega); \tag{9.147}$$
(c) There exists a $\mathbb{P}$-negligible event N such that $\mathbb{Q}_\omega\left[A(\omega)\right] = 1$ for all for $\omega \notin N$.

In item (c) we write $A(\omega) = \bigcap \{A : A \ni \omega, A \in \mathcal{F}_\tau^s\}$, $\omega \in \Omega$. Property (c) expresses the regularity of the conditional probability $\mathbb{Q}_\omega$. Property (b) is a quantitative property pertaining to the definition of conditional expectation, and (a) is a qualitative property defining conditional expectation.

The following theorem appears as Theorem 8.1.3 in [Stroock and Varadhan (2006)].

Theorem 9.14. *Let a and b be bounded Borel measurable functions with attain values in S_d and $\mathbb{R}^d$ respectively. Define the matrix functions $\widetilde{a}$ and $\widetilde{b}$ as in (9.144) and (9.145). Then the coefficients σ and b satisfy Itô's* *uniqueness conditions starting from* (s, y) *if and only if any solution* $\widetilde{P}$ *to the martingale problem relative to the operator* $\widetilde{L}$ *from* (s, y, y) *has the property that* $\widetilde{P}[X(t) = Y(t), t \geqslant s] = 1$. *Here, the processes* $X(t)$ *and* $Y(t)$ *attain their values in* $\mathbb{R}^d$ *and are such that for all* $f \in C_0^2(\mathbb{R}^d \times \mathbb{R}^d)$ *the process*

$$t \mapsto f(X(t), Y(t)) - f(X(s), Y(s)) - \int_s^t L(\rho) f(X(\rho), Y(\rho))\, d\rho, \quad t \geqslant s,$$

is a $\widetilde{P}$*-martingale after* s *relative to filtration determined by the* σ*-fields* $\mathcal{F}_t^s = \sigma((x(\rho), X(\rho)) : \rho \in [s, t])$.

Remark 9.11. In both theorems 9.13 and 9.14 the bounded progressively measurable processes $t \mapsto a(t)$ and $t \mapsto b(t)$ may be replaced with locally bounded Borel measurable functions from $[0, \infty) \times \mathbb{R}^d$ to S_d and $\mathbb{R}^d$ respectively. Of course the processes $a(t)$ and $b(t)$ have to read as $a(t, X(t))$ and $b(t, X(t))$ respectively. This is a consequence of Theorem 10.1.1 in [Stroock and Varadhan (2006)].

Proof. [Proof of Theorem 9.12.] The result in Theorem 9.12 is a consequence of Theorem 9.13 in conjunction with Theorem 9.14. In fact Theorem 9.13 reduces the stopping time τ to a fixed time of the form $\tau(\omega)$, where $\omega \in \Omega$ is fixed. Since at time $\tau(\omega)$, $X(\tau(\omega)) = Y(\tau(\omega))$ Theorem 9.14 shows that the coupling is successful (i.e. $X(t)(\omega) = Y(t)(\tau)$ Q_ω-almost surely for $\mathbb{P}_{s,x,y}$-almost all ω) in case the pair $(\sigma(t, x), b(t, x))$ consists of bounded functions and admits unique weak solutions. It then follows that $X(t) = Y(t)$ $\mathbb{P}_{s,x,y}$-almost surely on the event $\{\tau \leqslant t\}$. In formulas the arguments read as follows. From Theorem 9.13 we have

$$\begin{aligned} &\mathbb{P}\left[X(t) = Y(t) \mid \mathcal{F}_\tau^s\right] \mathbf{1}_{\{\tau \leqslant t,\, X(s)=x,\, Y(s)=y\}} \\ &= \mathbb{P}_{\tau, X(\tau), Y(\tau)}\left[X(t) = Y(t)\right] \mathbf{1}_{\{\tau \leqslant t,\, X(s)=x,\, Y(s)=y\}}. \end{aligned} \tag{9.148}$$

From Theorem 9.14 and (9.148) we get

$$\begin{aligned} &\mathbb{P}_{s,x,y}\left[X(t) = Y(t),\, \tau \leqslant t\right] \\ &= \mathbb{P}\left[X(t) = Y(t), X(s) = x,\, Y(s) = y,\, \tau \leqslant t\right] \\ &= \mathbb{E}\left[\mathbb{P}\left[X(t) = Y(t) \mid \mathcal{F}_\tau^s\right],\, X(s) = x,\, Y(s) = y,\, \tau \leqslant t\right] \\ &= \mathbb{E}\left[\mathbb{P}_{\tau, X(\tau), Y(\tau)}\left[X(t) = Y(t)\right],\, X(s) = x,\, Y(s) = y,\, \tau \leqslant t\right] \\ &= \mathbb{E}\left[\mathbf{1},\, X(s) = x,\, Y(s) = y,\, \tau \leqslant t\right] = \mathbb{P}\left[X(s) = x,\, Y(s) = y,\, \tau \leqslant t\right] \\ &= \mathbb{P}_{s,x,y}\left[\tau \leqslant t\right]. \end{aligned} \tag{9.149}$$

From (9.149) we infer that $X(t) = Y(t)$ $\mathbb{P}_{s,x,y}$-almost surely on the event $\{\tau \leqslant t\}$.

Remark 9.11 takes care of locally bounded coefficients. This finishes the proof of Theorem 9.12. □

In the following lemma we suppose that the operator L is time-independent.

Lemma 9.8. *Let the process* $\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}$ *be a coupling of the L-diffusion process. If there exists $\gamma \in \mathbb{R}$ such that*

$$\mathbb{E}_{x,y}\left[|X(t) - Y(t)|^2\right] \leqslant |x-y|^2 e^{-\gamma t} \tag{9.150}$$

for all $t \geqslant 0$ and all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, then

$$\left|\nabla e^{tL} f\right|^2 \leqslant e^{-\gamma t} e^{tL} |\nabla f|^2, \tag{9.151}$$

for all functions $f \in C^1\left(\mathbb{R}^d\right)$ with a bounded gradient which is uniformly continuous.

Proof. Let τ be the coupling time of the processes $X(t)$ and $Y(t)$ solving the coupled stochastic differential equation (9.141), and let $f \in C_b\left(\mathbb{R}^d\right)$ have a uniformly bounded gradient ∇f. Then by inequality (9.150) we have

$$\begin{aligned}
&\frac{\left|e^{tL} f(x) - e^{tL} f(y)\right|^2}{|x-y|^2} = \left|\mathbb{E}_{x,y}\left[\frac{f(X(t)) - f(Y(t))}{|X(t) - Y(t)|} \frac{|X(t) - Y(t)|}{|x-y|}, \ \tau > t\right]\right|^2 \\
&\leqslant \mathbb{E}_{x,y}\left[\frac{|f(X(t)) - f(Y(t))|^2}{|X(t) - Y(t)|^2}, \ \tau > t\right] \mathbb{E}_{x,y}\left[\frac{|X(t) - Y(t)|^2}{|x-y|^2}, \tau > t\right] \\
&\leqslant e^{-\gamma t} \mathbb{E}_{x,y}\left[\frac{|f(X(t)) - f(Y(t))|^2}{|X(t) - Y(t)|^2}, \ \tau > t\right] \\
&= e^{-\gamma t} \mathbb{E}_{x,y}\left[\left|\int_0^1 \left\langle \nabla f\left((1-s)Y(t) + sX(t)\right), \frac{X(t) - Y(t)}{|X(t) - Y(t)|}\right\rangle ds\right|^2, \ \tau > t\right] \\
&\leqslant e^{-\gamma t} \mathbb{E}_{x,y}\left[\int_0^1 \left|\nabla f\left((1-s)Y(t) + sX(t)\right)\right|^2 ds, \ \tau > t\right].
\end{aligned} \tag{9.152}$$

Next fix $\varepsilon > 0$, and choose $\delta > 0$ in such a way that $|y - z| \leqslant \delta$ implies $|\nabla f(y)|^2 \leqslant |\nabla f(z)|^2 + \varepsilon^2$. Then from (9.152) we obtain:

$$\frac{\left|e^{tL} f(x) - e^{tL} f(y)\right|^2}{|x-y|^2}$$

$$\begin{aligned}
&\leqslant e^{-\gamma t}\mathbb{E}_{x,y}\left[\int_0^1 \left|\nabla f\left((1-s)Y(t)+sX(t)\right)\right|^2\,ds,\ |Y(t)-X(t)|\leqslant\delta\right]\\
&\quad+e^{-\gamma t}\mathbb{E}_{x,y}\left[\int_0^1 \left|\nabla f\left((1-s)Y(t)+sX(t)\right)\right|^2\,ds,\ |Y(t)-X(t)|>\delta\right]\\
&\leqslant e^{-\gamma t}\mathbb{E}_{x,y}\left[\left|\nabla f\left(X(t)\right)\right|^2,\ |Y(t)-X(t)|\leqslant\delta\right]\\
&\quad+e^{-\gamma t}\left\|\nabla f\right\|_\infty^2\mathbb{P}_{x,y}\left[|Y(t)-X(t)|>\delta\right]+e^{-\gamma t}\varepsilon^2\\
&\leqslant e^{-\gamma t}\mathbb{E}_{x,y}\left[\left|\nabla f\left(X(t)\right)\right|^2,\ |Y(t)-X(t)|\leqslant\delta\right]\\
&\quad+e^{-\gamma t}\frac{1}{\delta^2}\left\|\nabla f\right\|_\infty^2\mathbb{E}_{x,y}\left[\left|Y(t)-X(t)\right|^2\right]+e^{-\gamma t}\varepsilon^2
\end{aligned}$$

(use (9.150))

$$\leqslant e^{-\gamma t}\mathbb{E}_{x,y}\left[\left|\nabla f\left(X(t)\right)\right|^2\right]+e^{-2\gamma t}\frac{1}{\delta^2}\left\|\nabla f\right\|_\infty^2|y-x|^2+e^{-\gamma t}\varepsilon^2. \tag{9.153}$$

In (9.153) we let y tend to x to obtain:

$$\left|\nabla e^{tL}f\right|^2(x)\leqslant e^{-\gamma t}\mathbb{E}_{x,x}\left[\left|\nabla f\left(X(t)\right)\right|^2\right]+e^{-2\gamma t}\varepsilon^2. \tag{9.154}$$

Since $\mathbb{E}_{x,x}\left[g\left(X(t)\right)\right]=e^{-tL}g(x)$, $g\in C_b\left(\mathbb{R}^d\right)$, and $\varepsilon>0$ is arbitrary the conclusion in Lemma 9.8 follows from (9.154). □

We conclude this section with a proof of Theorem 9.1.

Proof. [Proof of Theorem 9.1.] Put $h(x,y)=|x-y|^2$. From the representation in (9.143) of the operator $\widetilde{L}$, which now does not depend on t, we see that

$$\widetilde{L}h(x,y)=\operatorname{Tr}\left(\sigma(x)-\sigma(y)\right)\left(\sigma(x)-\sigma(y)\right)^*+2\left\langle b(x)-b(y),x-y\right\rangle. \tag{9.155}$$

From Itô's formula and (9.141) we get

$$\begin{aligned}
&|X(t)-Y(t)|^2=|X(0)-Y(0)|^2\\
&\quad+\sum_{k,\ell=1}^d\int_0^t\left(X_k(s)-Y_k(s)\right)\left(\sigma_{k,\ell}\left(X(s)\right)-\sigma_{k,\ell}\left(Y(s)\right)\right)dW_k(s)\\
&\quad+2\sum_{k=1}^d\int_0^t\left(X_k(s)-Y_k(s)\right)\left(b_k\left(X(s)\right)-b_k\left(Y(s)\right)\right)ds\\
&\quad+\sum_{i=1}^d\sum_{k=1}^d\int_0^t\left(\sigma_{i,k}\left(X(s)\right)-\sigma_{i,k}\left(Y(s)\right)\right)\left(\sigma_{i,k}\left(X(s)\right)-\sigma_{j,k}\left(Y(s)\right)\right)ds\\
&=|X(0)-Y(0)|^2
\end{aligned}$$

$$+ \sum_{k,\ell=1}^{d} \int_0^t (X_k(s) - Y_k(s)) \left(\sigma_{k,\ell}(X(s)) - \sigma_{k,\ell}(Y(s))\right) dW_k(s)$$

$$+ 2\sum_{k=1}^{d} \int_0^t \langle X(s) - Y(s), b(X(s)) - b(Y(s))\rangle \, ds$$

$$+ \int_0^t \operatorname{Tr}\left(\left(\sigma(X(s)) - \sigma(Y(s))\right)\left(\sigma(X(s)) - \sigma(Y(s))\right)^*\right) ds. \tag{9.156}$$

Put $\varphi(t) = \mathbb{E}_{x,y}\left[|X(t) - Y(t)|^2\right]$. Then (9.156) and the definition of γ in (9.5) we see that $\varphi'(t) \leqslant -\gamma\varphi(t)$. It follows that $\varphi(t) \leqslant \varphi(0)e^{-\gamma t}$. From Lemma 9.8, in particular from (9.151) we see that

$$\left|\nabla e^{tL} f\right|^2 \leqslant e^{-\gamma t} e^{tL} |\nabla f|^2, \tag{9.157}$$

for all functions $f \in C_b\left(\mathbb{R}^d\right)$ with bounded uniformly continuous gradient. Let $f \in C_b\left(\mathbb{R}^d\right)$ be such a function. Then from (9.157) we infer

$$\begin{aligned}
e^{tL} |f| f^2 - \left|e^{tL} f\right|^2 &= \int_0^t \frac{\partial}{\partial s} e^{sL} \left|e^{(t-s)L} f\right|^2 ds \\
&= \int_0^t \frac{\partial}{\partial s} e^{sL} \left\langle a \nabla e^{(t-s)L}\overline{f}, \nabla e^{(t-s)L} f\right\rangle ds \qquad (9.158)\\
&\leqslant \overline{a} \int_0^t \frac{\partial}{\partial s} e^{sL} \left\langle \nabla e^{(t-s)L} f, \nabla e^{(t-s)L} f\right\rangle ds \\
&= \overline{a} \int_0^t \frac{\partial}{\partial s} e^{sL} \left|\nabla e^{(t-s)L} f, \nabla\right|^2 ds \\
&\leqslant \overline{a} \int_0^t e^{-(t-s)\gamma} e^{sL} e^{(t-s)L} |\nabla f|^2 ds \\
&= \overline{a} \frac{1 - e^{-\gamma t}}{\gamma} e^{tL} |\nabla f|^2. \qquad (9.159)
\end{aligned}$$

The inequality in (9.159) completes the proof of Theorem 9.1. □

9.4 Some related stability results

Let L be the generator of a $\mathcal{T}_\beta$-diffusion which, by definition, is a time-homogeneous Markov process

$$\left\{\left(\Omega, \mathcal{F}_t^0, \mathbb{P}_x\right), (X(t), t \geqslant 0), (\vartheta_t, t \geqslant 0), (E, \mathcal{E})\right\}.$$

Let Γ_1 be the corresponding squared gradient operator. With $f \in D(L)$ we associate the martingale $t \mapsto M_f(t)$ defined by

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t Lf(X(s))\, ds.$$

For more details on the squared gradient operators see e.g. [Bakry (1994)] and [Bakry (2006)]. Then for f, $g \in D(L)$ we have

$$\langle M_f, M_g\rangle (t) = \int_0^t \Gamma_1 (f,g) (X(s))\, ds. \tag{9.160}$$

Denote by $\left\{e^{tL} : t \geqslant 0\right\}$ the semigroup generated by L.

Theorem 9.15. *Let $f \in D\left(L^2\right)$. Then the following identities hold for ρ, $t \geqslant 0$ and $x \in E$:*

$$\begin{aligned} & f\left(X(\rho+t)\right) - \mathbb{E}_{X(\rho)}\left[f\left(X(t)\right)\right] \qquad (9.161)\\ &= M_f(\rho+t) - M_f(\rho) + \int_\rho^{\rho+t} \left\{M_{e^{(\rho+t-\sigma)L}Lf}(\sigma) - M_{e^{(\rho+t-\sigma)L}Lf}(\rho)\right\} d\sigma \\ &= M_f(\rho+t) - M_f(\rho) + \int_0^{t} \left\{M_{e^{(t-\sigma)L}Lf}(\rho+\sigma) - M_{e^{(t-\sigma)L}Lf}(\rho)\right\} d\sigma, \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E}_x\left[\left|f\left(X(\rho+t)\right) - \mathbb{E}_{X(\rho)}\left[f\left(X(t)\right)\right]\right|^2\right] \\ &= \mathbb{E}_x\left[\left|f\left(X(\rho+t)\right)\right|^2\right] - \left|\mathbb{E}_x\left[f\left(X(\rho+t)\right)\right]\right|^2 \\ &= e^{(\rho+t)L}\left|f\right|^2(x) - e^{\rho L}\left|e^{tL}f(x)\right|^2 \\ &= \mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{e^{(t-\sigma)L}f}, e^{(t-\sigma)L}f\right)\left(X(\rho+\sigma)\right)\, d\sigma\right] \\ &= \int_0^t e^{(\rho+\sigma)L}\Gamma_1\left(\overline{e^{(t-\sigma)L}f}, e^{(t-\sigma)L}f\right)(x)\, d\sigma. \end{aligned} \tag{9.162}$$

Remark 9.12. In (9.161) we need the fact that $f \in D\left(L^2\right)$. In (9.162) the hypotheses $f \in D(L)$ suffices.

Proof. First we prove the equality in (9.161). Therefore we write:

$$\begin{aligned} & M_f(\rho+t) - M_f(\rho) + \int_\rho^{\rho+t} \left\{M_{e^{(\rho+t-\sigma)L}Lf}(\sigma) - M_{e^{(\rho+t-\sigma)L}Lf}(\rho)\right\} d\sigma \\ &= M_f(\rho+t) - M_f(\rho) \\ &\quad + \int_\rho^{\rho+t} \left\{e^{(\rho+t-\sigma)L}Lf\left(X(\sigma)\right) - e^{(\rho+t-\sigma)L}Lf\left(X(\rho)\right)\right. \\ &\qquad \left. - \int_\rho^{\sigma} e^{(\rho+t-\sigma)L}L^2 f\left(X\left(\sigma_1\right)\right)\, d\sigma_1\right\} d\sigma \\ &= M_f(\rho+t) - M_f(\rho) + \int_\rho^{\rho+t} e^{(\rho+t-\sigma)L}Lf\left(X(\sigma)\right)\, d\sigma \end{aligned}$$

$$
\begin{aligned}
&\quad - \int_{\rho}^{\rho+t} e^{(\rho+t-\sigma)L} Lf\left(X(\rho)\right)\, d\sigma \\
&\quad - \int_{\rho}^{\rho+t} \int_{\sigma_1}^{\rho+t} e^{(\rho+t-\sigma)L} L^2 f\left(X\left(\sigma_1\right)\right)\, d\sigma\, d\sigma_1 \\
&= M_f(\rho+t) - M_f(\rho) + \int_{\rho}^{\rho+t} e^{(\rho+t-\sigma)L} Lf\left(X(\sigma)\right)\, d\sigma \\
&\quad + \int_{\rho}^{\rho+t} \frac{\partial}{\partial\sigma} e^{(\rho+t-\sigma)L} f\left(X(\rho)\right)\, d\sigma \\
&\quad + \int_{\rho}^{\rho+t} \int_{\sigma_1}^{\rho+t} \frac{\partial}{\partial\sigma} e^{(\rho+t-\sigma)L} Lf\left(X\left(\sigma_1\right)\right) d\sigma\, d\sigma_1 \\
&= M_f(\rho+t) - M_f(\rho) + \int_{\rho}^{\rho+t} e^{(\rho+t-\sigma)L} Lf\left(X(\sigma)\right)\, d\sigma \\
&\quad + f\left(X(\rho)\right) - e^{tL} f\left(X(\rho)\right) \\
&\quad + \int_{\rho}^{\rho+t} \Big(Lf\left(X\left(\sigma_1\right)\right) - e^{(\rho+t-\sigma_1)L} Lf\left(X\left(\sigma_1\right)\right)\Big)\, d\sigma_1 \\
&= f\left(X(\rho+t) - f\left(X(\rho)\right) - \int_{\rho}^{\rho+t} Lf\left(X(s)\right) ds \right. \\
&\quad + \int_{\rho}^{\rho+t} e^{(\rho+t-\sigma)L} Lf\left(X(\sigma)\right)\, d\sigma + f\left(X(\rho)\right) - \mathbb{E}_{X(\rho)}\left[f\left(X(t)\right)\right] \\
&\quad + \int_{\rho}^{\rho+t} \Big(Lf\left(X\left(\sigma_1\right)\right) - e^{(\rho+t-\sigma_1)L} Lf\left(X\left(\sigma_1\right)\right)\Big)\, d\sigma_1 \\
&= f\left(X(\rho+t)\right) - \mathbb{E}_{X(\rho)}\left[f\left(X(t)\right)\right].
\end{aligned}
\tag{9.163}
$$

The equality in (9.161) is the same as the one in (9.163). The proof of (9.162) is much more difficult. We will employ the equalities in (9.160) and (9.161) to obtain it. From (9.161) we get

$$
\begin{aligned}
&\left| f\left(X(\rho+t)\right) - \mathbb{E}_{X(\rho)}\left[f\left(X(t)\right)\right]\right|^2 \\
&= \left| M_f(\rho+t) - M_f(\rho) + \int_0^t \left\{ M_{e^{(t-\sigma)L}Lf}(\rho+\sigma) - M_{e^{(t-\sigma)L}Lf}(\rho)\right\}\, d\sigma \right|^2 \\
&= \left| M_f(\rho+t) - M_f(\rho)\right|^2 \\
&+ 2\Re\left(\overline{M_f(\rho+t) - M_f(\rho)} \int_0^t \left\{ M_{e^{(t-\sigma)L}Lf}(\rho+\sigma) - M_{e^{(t-\sigma)L}Lf}(\rho)\right\} d\sigma \right) \\
&\quad + \left| \int_0^t \left\{ M_{e^{(t-\sigma)L}Lf}(\rho+\sigma) - M_{e^{(t-\sigma)L}Lf}(\rho)\right\}\, d\sigma \right|^2 .
\end{aligned}
\tag{9.164}
$$

For brevity we write

$$\triangle_\rho M_g(\sigma) = M_g(\rho+\sigma) - M_g(\rho), \quad g \in D(L). \tag{9.165}$$

Next we use the fact that processes of the form $\sigma \mapsto \triangle_\rho M_g(\sigma)$, $g \in D(L)$, are $\mathbb{P}_x$-martingales with respect to the filtration $\left\{\mathcal{F}^\rho_{\rho+\sigma} : \sigma > 0\right\}$. From this together with (9.160) and (9.164) we obtain, using the notation in (9.165):

$$\begin{aligned}
&\mathbb{E}_x\left[\left|f(X(\rho+t)) - \mathbb{E}_{X(\rho)}\left[f(X(t))\right]\right|^2\right] \\
&= \mathbb{E}_x\left[\left|\triangle_\rho M_f(t)\right|^2\right] + 2\Re\mathbb{E}_x\left[\int_0^t \overline{\triangle_\rho M_f(t)}\left\{\triangle_\rho M_{e^{(t-\sigma)L}Lf}(\sigma)\right\} d\sigma\right] \\
&\quad + \mathbb{E}_x\left[\int_0^t\int_0^t \overline{\triangle_\rho M_{e^{(t-\rho_1)L}Lf}(\rho_1)}\left\{\triangle_\rho M_{e^{(t-\rho_2)L}Lf}(\rho_2)\right\} d\rho_1\, d\rho_2\right] \\
&= \mathbb{E}_x\left[\left|\triangle_\rho M_f(t)\right|^2\right] + 2\Re\left(\mathbb{E}_x\left[\int_0^t \overline{\triangle_\rho M_f(\sigma)}\triangle_\rho M_{e^{(t-\sigma)L}Lf}(\sigma)\, d\sigma\right]\right) \\
&\quad + \mathbb{E}_x\left[\int_0^t\int_0^t \overline{\triangle_\rho M_{e^{(t-\rho_1)L}Lf}(\rho_1\wedge\rho_2)}\triangle_\rho M_{e^{(t-\rho_2)L}Lf}(\rho_1\wedge\rho_2)\, d\rho_1\, d\rho_2\right]
\end{aligned}$$

(employ (9.160) several times)

$$\begin{aligned}
&= \mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{f}, f\right)(X(\rho+\sigma))\, d\sigma\right] \\
&\quad + 2\Re\left(\mathbb{E}_x\left[\int_0^t\int_0^\sigma \Gamma_1\left(\overline{f}, e^{(t-\sigma)L}Lf\right)(X(\rho+\sigma_1))\, d\sigma_1\, d\sigma\right]\right) \\
&+ \mathbb{E}_x\left[\int_0^t\int_0^t\int_0^{\rho_1\wedge\rho_2} \Gamma_1\left(\overline{e^{(t-\rho_1)L}Lf}, e^{(t-\rho_2)L}Lf\right)(X(\rho+\sigma))\, d\sigma\, d\rho_1\, d\rho_2\right] \\
&= \mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{f}, f\right)(X(\rho+\sigma))\, d\sigma\right] \\
&\quad + 2\Re\left(\mathbb{E}_x\left[\int_0^t\int_{\sigma_1}^t \Gamma_1\left(\overline{f}, e^{(t-\sigma)L}Lf\right)(X(\rho+\sigma_1))\, d\sigma\, d\sigma_1\right]\right) \\
&\quad + \mathbb{E}_x\left[\int_0^t\int_\sigma^t\int_\sigma^t \Gamma_1\left(\overline{e^{(t-\rho_1)L}Lf}, e^{(t-\rho_2)L}Lf\right)(X(\rho+\sigma))\, d\rho_1\, d\rho_2\, d\sigma\right]
\end{aligned}$$

(the operator Γ_1 is bilinear)

$$\begin{aligned}
&= \mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{f}, f\right)(X(\rho+\sigma))\, d\sigma\right] \\
&\quad + 2\Re\left(\mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{f}, \int_{\sigma_1}^t e^{(t-\sigma)L}Lf\, d\sigma\right)(X(\rho+\sigma_1))\, d\sigma_1\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{\int_\sigma^t e^{(t-\rho_1)L}Lf\,d\rho_1}, \int_\sigma^t e^{(t-\rho_2)L}Lf\,d\rho_2\right)(X(\rho+\sigma))\,d\sigma\right]\\
&=\mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{f},f\right)(X(\rho+\sigma))\,d\sigma\right]\\
&\quad+2\Re\left(\mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{f}, e^{(t-\sigma_1)L}f-f\right)(X(\rho+\sigma_1))\,d\sigma_1\right]\right)\\
&\quad+\mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{e^{(t-\sigma)L}f-f}, e^{(t-\sigma)L}f-f\right)(X(\rho+\sigma))\,d\sigma\right]\\
&=\mathbb{E}_x\left[\int_0^t \Gamma_1\left(\overline{e^{(t-\sigma)L}f}, e^{(t-\sigma)L}f\right)(X(\rho+\sigma))\,d\sigma\right]. \qquad (9.166)
\end{aligned}
$$

The equality in (9.166) yields (9.162) for $f \in D\left(L^2\right)$. Since $D\left(L^2\right)$ is $\mathcal{T}_\beta$-dense in $D(L)$ we infer (9.162) for $f \in D(L)$.

An easier proof of equality (9.162) reads as follows. We calculate:

$$
\begin{aligned}
&\frac{\partial}{\partial\sigma}\left\{e^{(\rho+\sigma)L}\left|e^{(t-\sigma)L}f\right|^2\right\}\\
&=e^{(\rho+\sigma)L}L\left|e^{(t-\sigma)L}f\right|^2\\
&\quad-e^{(\rho+\sigma)L}\left\{\overline{Le^{(t-\sigma)L}f}e^{(t-\sigma)L}f+\overline{e^{(t-\sigma)L}f}Le^{(t-\sigma)L}f\right\}\\
&=e^{(\rho+\sigma)L}\Gamma_1\left(\overline{e^{(t-\sigma)L}f}, e^{(t-\sigma)L}f\right). \qquad (9.167)
\end{aligned}
$$

In (9.167) we used the identity

$$
L\left(\overline{f}g\right)=\overline{Lf}g+\overline{f}Lg+\Gamma_1\left(\overline{f},g\right). \qquad (9.168)
$$

The equality in (9.168) is true for f, $g \in D(L)$ such that $\overline{f}g \in D(L)$. From (9.167) we obtain:

$$
\begin{aligned}
&\int_0^t e^{(\rho+\sigma)L}\Gamma_1\left(\overline{e^{(t-\sigma)L}f}, e^{(t-\sigma)L}f\right)d\sigma=\int_0^t\frac{\partial}{\partial\sigma}\left\{e^{(\rho+\sigma)L}\left|e^{(t-\sigma)L}f\right|^2\right\}d\sigma\\
&=e^{(\rho+t)L}|f|^2-e^{\rho L}\left|e^{tL}f\right|^2. \qquad (9.169)
\end{aligned}
$$

The equality in (9.169) implies (9.162), and completes the proof of Theorem 9.12. □

Remark 9.13. Suppose that there exist constants $c > 0$ and $\gamma \in \mathbb{R}$ such that

$$
\Gamma_1\left(e^{\rho L}\overline{f}, e^{\rho L}f\right) \leqslant ce^{-\gamma\rho}e^{\rho L}\Gamma_1\left(\overline{f},f\right). \qquad (9.170)
$$

From (9.169) and (9.170) with $\rho = 0$ we obtain:

$$\begin{aligned} e^{tL}\left|f\right|^2 - \left|e^{tL}f\right|^2 &= \int_0^t e^{\rho L}\Gamma_1\left(e^{(t-\rho)L}\overline{f}, e^{(t-\rho)L}f\right)d\rho \\ &\leqslant \int_0^t e^{\rho L}\left\{ce^{(t-\rho)\gamma}e^{(t-\rho)L}\Gamma_1\left(\overline{f},f\right)\right\}d\rho \\ &= \frac{c}{\gamma}\left(1-e^{-t\gamma}\right)e^{tL}\Gamma_1\left(\overline{f},f\right). \end{aligned} \tag{9.171}$$

The inequality in (9.171) is the same as inequality (4.14) in Theorem 4.13 in [Chen and Wang (1997)]. If μ is an invariant probability for the operator L, then (9.171) implies

$$\begin{aligned} &\lim_{t\to\infty}\int e^{tL}\left|f - e^{tL}f(x)\right|^2(x)d\mu(x) \\ &= \int\left|f\right|^2 d\mu - \lim_{t\to\infty}\int\left|e^{tL}f\right|^2 d\mu \leqslant \frac{c}{\gamma}\int\Gamma_1\left(\overline{f},f\right)d\mu, \end{aligned} \tag{9.172}$$

provided $\gamma > 0$. From (9.172) it follows that the L^2-spectral gap of L is bounded from below by γ/c. The inequality in (9.172) can be considered as a spectral gap or Poincaré inequality: compare with Definition 9.15 below.

The following definition is to be compared with the Definitions 8.4 and 8.5. This definition is also closely related to Fang's spectral gap theorem in [Fang (1993)]: see Theorem 5.4 in [Driver (1995; Last revised: January 29, 2003)] as well. The latter reference also contains some results on the relationship between Fang's spectral gap theorem and the logarithmic Sobolev inequality: see Section 5.4 in [Driver (1995)].

Definition 9.14. Let μ be the unique invariant measure of the generator of a diffusion L with associated squared gradient operator Γ_1. Then the $L^2(\mu)$-spectral gap of the operator L is defined by the equality

$$2\text{gap}\,(L) = \inf\left\{\int\Gamma_1\left(\overline{f},f\right)d\mu : f \in D(L), \int f d\mu = 0, \int\left|f\right|^2 d\mu = 1\right\}. \tag{9.173}$$

Proposition 9.10. *Let the measure μ and* $\text{gap}(L) > 0$ *be as in Definition 9.14. Then $\gamma \in (0,\infty)$ satisfies $\gamma \leqslant 2 \times \text{gap}(L)$ if and only if the following inequality holds for all $t > 0$ and for all $f \in C_b(E)$:*

$$\int\left|e^{tL}f\right|^2 d\mu - \left|\int e^{tL}f d\mu\right|^2 \leqslant e^{-t\gamma}\left(\int\left|f\right|^2 d\mu - \left|\int f d\mu\right|^2\right). \tag{9.174}$$

The inequality in (9.174) holds for all $t > 0$ and $f \in C_b(E)$ if and if the inequality

$$\int \Gamma_1\left(\overline{f}, f\right) d\mu \geqslant \gamma \left(\int |f|^2 \, d\mu - \left| \int f d\mu \right|^2 \right) \tag{9.175}$$

holds for all $f \in D(L)$.

Definition 9.15. An inequality of the form (9.175) is called a Poincaré or a spectral gap inequality of $L^2(E,\mu)$-type.

Notice that by invariance of the measure μ we have $\int e^{tL} f d\mu = \int f d\mu$, and $\int Lf d\mu = 0$, $f \in D(L)$. Also notice that, since μ is a probability measure, the decomposition $f = f - \int f d\mu + \int f d\mu$ splits the function f in two orthogonal functions (one of them being the constant $\int f d\mu$) in the space $L^2(E,\mu)$. Hence we have

$$\begin{aligned} \inf_{\alpha\in\mathbb{C}} \int |f-\alpha|^2 \, d\mu &= \int \left| f - \int f d\mu \right|^2 d\mu = \int |f|^2 \, d\mu - \left| \int f d\mu \right|^2 \\ &= \frac{1}{2} \iint |f(x) - f(y)|^2 \, d\mu(x) \, d\mu(y). \end{aligned} \tag{9.176}$$

Remark 9.14. If the probability measure μ is invariant under the semi-group generated by L, then

$$\begin{aligned} \int \Gamma_1\left(\overline{f}, g\right) d\mu &= \int L\left(\overline{f} g\right) d\mu - \int \left(L\overline{f}\right) g \, d\mu - \int \overline{f} \, (Lg) \, d\mu \\ &= -\int \left(L\overline{f}\right) g \, d\mu - \int \overline{f} \, (Lg) \, d\mu = -\int \left(L + L^*\right) \overline{f} \cdot g \, d\mu \end{aligned} \tag{9.177}$$

where L^* is the adjoint of the operator L in the space $L^2(E,\mu)$. From (9.177) we infer

$$\begin{aligned} 2\mathrm{gap}(L) &= \inf \left\{ \int \Gamma_1\left(\overline{f}, f\right) d\mu : \int f d\mu = 0, \int |f|^2 \, d\mu = 1 \right\} \\ &= \inf \left\{ -\int \left(\left(L + L^*\right) \overline{f} \right) \cdot f \, d\mu : \int f d\mu = 0, \int |f|^2 \, d\mu = 1 \right\}, \end{aligned}$$

and hence the number $2\mathrm{gap}(L)$ is the bottom of the spectrum of the operator $-\left(L + L^*\right)$ in the space $\left\{ f - \int f d\mu : f \in L^2(E,\mu) \right\}$ which is the orthogonal complement of the subspace consisting of the constant functions in $L^2(E,\mu)$. In particular, if $L = L^*$, then $\mathrm{gap}(L)$ is the gap in the spectrum of $-L$ between 0 and $[\mathrm{gap}(L), \infty) \bigcap \sigma_\mu(L)$. Here $\sigma_\mu(L)$ denotes the spectrum of L as an operator in the space $L^2(E,\mu)$. In fact it would have been better to write $\mathrm{gap}\left(L + L^*\right)$ instead of $2\mathrm{gap}(L)$. Of course 0 is an eigenvalue of L and the constant functions are the corresponding eigenvectors.

Proof. [Proof of Proposition 9.10.] If $\gamma > 0$ is such that (9.174) is satisfied for all $t > 0$ and for all $f \in C_b(E)$. Then we subtract $\int |f|^2 \, d\mu - \left|\int f d\mu\right|^2$ from both sides of (9.174) and divide by $t > 0$ to obtain:

$$\frac{1}{t}\left(\int \left|e^{tL}f\right|^2 d\mu - \int |f|^2 \, d\mu\right) \leqslant \frac{e^{-\gamma t}-1}{t}\left(\int |f|^2 \, d\mu - \left|\int f d\mu\right|^2\right). \tag{9.178}$$

In (9.178) we let $t \downarrow 0$ to obtain:

$$\int \left(L\overline{f} \cdot f + \overline{f} \cdot Lf\right) d\mu \leqslant -\gamma \left(\int |f|^2 \, d\mu - \left|\int f d\mu\right|^2\right), \tag{9.179}$$

or what amounts to the same:

$$\int \left(L\,|f|^2 - \Gamma_1\left(\overline{f}, f\right)\right) d\mu \leqslant -\gamma \left(\int |f|^2 \, d\mu - \left|\int f d\mu\right|^2\right). \tag{9.180}$$

Since by invariance $\int L\,|f|^2 \, d\mu = 0$ from (9.180) we infer (9.175) and hence $\gamma \leqslant \text{gap}(L)$.

For the converse statement we consider, for $f \in D(L)$ and $\gamma > 0$ such that $\gamma \leqslant \text{gap}(L)$, the function

$$\begin{aligned}\varphi(t) = \int \left|e^{tL}f - \int e^{tL} f d\mu\right|^2 d\mu &= \int \left|e^{tL}f\right|^2 d\mu - \left|\int e^{tL} f d\mu\right|^2 \\ &= \int \left|e^{tL}f\right|^2 d\mu - \left|\int f d\mu\right|^2. \end{aligned} \tag{9.181}$$

Then from (9.175) we infer

$$\begin{aligned}\varphi'(t) &= \int \left(L\left(e^{tL}\overline{f}\right) e^{tL}f + \left(e^{tL}\overline{f}\right) Le^{tL}f\right) d\mu \\ &= \int L\left(\left|e^{tL}f\right|^2\right) d\mu - \int \Gamma_1\left(\overline{e^{tL}f}, e^{tL}f\right) \, d\mu \\ &= -\int \Gamma_1\left(\overline{e^{tL}f}, e^{tL}f\right) \, d\mu \\ &\leqslant -\gamma\left(\int \left|e^{tL}f\right|^2 d\mu - \left|\int e^{tL} f d\mu\right|^2\right) = -\gamma\varphi(t), \end{aligned} \tag{9.182}$$

and hence $\varphi(t) \leqslant e^{-\gamma t}\varphi(0)$, which is the same as (9.174).

Since, it is easy to see that $\gamma \leqslant \text{gap}\,(L)$ if and only if inequality (9.175) holds for all $f \in D(L)$, this proves Proposition 9.10. □

Definition 9.16. Let μ be an invariant probability measure and let $f \geqslant 0$ be a Borel measurable function which is not μ-almost everywhere zero. Then the entropy of f with respect to μ is defined by

$$\operatorname{Ent}(f) = \int f \log f \, d\mu - \int f d\mu \log \int f d\mu = \int f \log \frac{f}{\int f d\mu} d\mu. \tag{9.183}$$

Definition 9.17. Let μ be an probability measure. A logarithmic Sobolev inequality takes the form

$$\operatorname{Ent}\left(|f|^2\right) \leqslant A \int |f|^2 \, d\mu + \frac{1}{\lambda} \int \Gamma_1\left(\overline{f}, f\right) d\mu \tag{9.184}$$

for all f in a large enough subalgebra $\mathcal{A}$ of $C_b(E)$. Here $A \geqslant 0$ and $\lambda > 0$ are constants. If the constant A can be chosen to be $A = 0$, then (9.184) is called a tight logarithmic Sobolev inequality.

Here $\operatorname{Ent}\left(|f|^2\right)$ is defined in Definition 9.16. The following proposition gives a relationship between tight logarithmic Sobolev inequalities and the Poincaré inequality: see Definition 9.15 and inequality (9.175).

Proposition 9.11. *Suppose that L satisfies a logarithmic Sobolev inequality with constants A and $\lambda > 0$, and suppose that L satisfies a Poincaré inequality with a constant $\gamma > 0$. Then L satisfies a tight logarithmic Sobolev inequality.*

In the proof we use an inequality which we owe to Rothaus. It is given in Proposition 9.19 below as inequality (9.274).

Proof. Let $f \in \mathcal{A}$ and put $\widehat{f} = f - \int f d\mu$. A combination of inequality (9.274) and Poincaré's inequality yields:

$$\begin{aligned}\operatorname{Ent}\left(|f|^2\right) &\leqslant 2 \int \left|\widehat{f}\right|^2 d\mu + \operatorname{Ent}\left(\left|\widehat{f}\right|^2\right) \\ &\leqslant (2 + A) \int \left|\widehat{f}\right|^2 d\mu + \frac{1}{\lambda} \int \Gamma_1\left(\overline{f}, f\right) d\mu\end{aligned}$$

(invoke Poincaré's inequality with constant $\gamma > 0$)

$$\leqslant \left(\frac{2 + A}{\gamma} + \frac{1}{\lambda}\right) \int \left|\widehat{f}\right|^2 d\mu. \tag{9.185}$$

The inequality in (9.185) is a tight logarithmic Sobolev inequality. This proves Proposition 9.11. □

Definition 9.18. Let μ be an probability measure. A Sobolev inequality of order $p > 2$ has the form

$$\left(\int |f|^p \, d\mu\right)^{2/p} \leqslant A \int |f|^2 \, d\mu + \frac{1}{\lambda} \int \Gamma_1 \left(\overline{f}, f\right) d\mu \tag{9.186}$$

for all f in a large enough subalgebra $\mathcal{A}$ of $C_b(E)$. Here, as in Definition 9.17, $A \geqslant 0$ and $\lambda > 0$ are constants.

In the following proposition we see that a tight logarithmic Sobolev inequality implies the Poincaré inequality.

Proposition 9.12. *Suppose that in (9.184) the constant $A = 0$. Then the inequality in (9.175) is satisfied with $\gamma = 2\lambda$, and hence $\lambda \leqslant \operatorname{gap}(L)$.*

Proof. Insert $f = 1 + \varepsilon g$, $\varepsilon > 0$, $g \in C_b(E, \mathbb{R})$, into (9.184), and divide by ε^2, to obtain

$$\begin{aligned} 0 &\geqslant \lambda \int \frac{(1+\varepsilon g)^2}{\varepsilon^2} \log \frac{(1+\varepsilon g)^2}{\int (1+\varepsilon g)^2 d\mu} d\mu - \int \Gamma_1 (g,g) \, d\mu \\ &= \lambda \int \frac{(1+\varepsilon g)^2}{\varepsilon^2} \log \frac{1 + 2\varepsilon g + \varepsilon^2 g^2}{1 + 2\varepsilon \int g d\mu + \varepsilon^2 \int g^2 d\mu} d\mu - \int \Gamma_1 (g,g) \, d\mu \end{aligned}$$

$\left(\log(1+x) = x - \frac{1}{2}x^2 + \mathcal{O}\left(x^3\right) \text{ for } x \to 0\right)$

$$\begin{aligned} &= \lambda \int \frac{(1+\varepsilon g)^2}{\varepsilon} \left\{ 2g + \varepsilon g^2 - \frac{\varepsilon}{2} \left(2g + \varepsilon g^2\right)^2 - 2 \int g \, d\mu - \varepsilon \int g^2 d\mu \right. \\ &\qquad \left. + \frac{\varepsilon}{2} \left(2 \int g \, d\mu + \varepsilon \int g^2 d\mu\right)^2 \right\} d\mu - \int \Gamma_1 (g,g) \, d\mu + \mathcal{O}(\varepsilon) \\ &= 2\lambda \left(\int g^2 d\mu - \left(\int g \, d\mu\right)^2\right) - \int \Gamma_1 (g,g) \, d\mu + \mathcal{O}(\varepsilon). \end{aligned} \tag{9.187}$$

From (9.187) we infer

$$2\lambda \left(\int g^2 d\mu - \left(\int g \, d\mu\right)^2\right) \leqslant \int \Gamma_1 (g,g) \, d\mu. \tag{9.188}$$

From (9.188) and the bi-linearity of Γ_1 it follows that

$$2\lambda \left(\int |g|^2 \, d\mu - \left|\int g \, d\mu\right|^2\right) \leqslant \int \Gamma_1 (\overline{g}, g) \, d\mu, \tag{9.189}$$

for $g \in D(L)$ which take complex values.

By employing Proposition 9.10 this proves Proposition 9.12. □

A combination of the propositions 9.11 and 9.12 yields the following corollary.

Corollary 9.6. *Suppose that the operator L satisfies a logarithmic Sobolev inequality. Then L satisfies a tight logarithmic Sobolev inequality if and only if it satisfies a Poincaré inequality.*

In the following proposition we see that a Sobolev inequality combined with a Poincaré inequality yields a Sobolev inequality with a constant $A = 1$. In the proof we employ inequality (9.273) in Proposition 9.19 below.

Proposition 9.13. *Suppose that the operator L (or in fact the corresponding squared gradient operator Γ_1) satisfies a Sobolev inequality of order $p > 2$ with constants A and λ: see inequality (9.186) in Definition 9.18. In addition suppose that L satisfies a Poincaré inequality of the form (9.175) with constant $\gamma > 0$. Then L satisfies a Sobolev inequality of order $p > 2$ with constants $A = 1$ and λ_0 satisfying $\dfrac{1}{\lambda_0} = (p-1)\left(\dfrac{A}{\gamma} + \dfrac{1}{\lambda}\right)$.*

Proof. Let $f \in \mathcal{A}$. An appeal to inequality (9.273) in Proposition 9.19 yields the following inequalities:

$$\begin{aligned}\left(\int |f|^p \, d\mu\right)^{2/p} &\leqslant \left|\int f \, d\mu\right|^2 + (p-1)\left(\int \left|f - \int f d\mu\right|^p d\mu\right)^{2/p} \\ &\leqslant \left|\int f \, d\mu\right|^2 + (p-1)A\int \left|f - \int f d\mu\right|^2 d\mu \\ &\quad + \frac{p-1}{\lambda}\int \Gamma_1\left(\overline{f - \int f d\mu}, f - \int f d\mu\right) d\mu \\ &\leqslant \left|\int f \, d\mu\right|^2 + (p-1)\left(\frac{A}{\gamma} + \frac{1}{\lambda}\right)\int \Gamma_1\left(\overline{f}, f\right) d\mu. \qquad (9.190)\end{aligned}$$

The claim in Proposition 9.13 follows from (9.190). □

The following theorem says that the entropy defined in terms of an invariant probability measure has exponential decay for $t \to \infty$ provided that L satisfies a tight logarithmic Sobolev inequality.

Theorem 9.16. *Let $\lambda > 0$. The following assertions are equivalent.*

(i) For all functions $f \in \mathcal{A}$ the following inequality holds:

$$\lambda \mathrm{Ent}\left(|f|^2\right) \leqslant \int \Gamma_1\left(\overline{f}, f\right) d\mu. \qquad (9.191)$$

(ii) For all functions $f \in \mathcal{A}$ the following inequality holds:

$$\operatorname{Ent}\left(e^{tL}|f|^2\right) \leqslant e^{-2\lambda t}\operatorname{Ent}\left(|f|^2\right). \tag{9.192}$$

The proof is based on the equalities (see (9.196) below):

$$\begin{aligned}\frac{d}{dt}\operatorname{Ent}\left(e^{tL}|f|^2\right) &= -\frac{1}{2}\int\frac{\Gamma_1\left(e^{tL}|f|^2, e^{tL}|f|^2\right)}{e^{tL}|f|^2}d\mu \\ &= -2\int\Gamma_1\left(\left(e^{tL}|f|^2\right)^{1/2}, \left(e^{tL}|f|^2\right)^{1/2}\right)d\mu.\end{aligned}$$

Proof. [Proof of Theorem 9.16.] Let $f \in \mathcal{A}$. We calculate

$$\frac{d}{dt}\operatorname{Ent}\left(e^{tL}|f|^2\right) = \frac{d}{dt}\left(\int e^{tL}|f|^2\log\left(\frac{e^{tL}|f|^2}{\int e^{tL}|f|^2 d\mu}\right)d\mu\right)$$

(μ is L-invariant)

$$\begin{aligned}&= \frac{d}{dt}\left(\int e^{tL}|f|^2\log\left(e^{tL}|f|^2\right)d\mu\right) \\ &= \int\left(Le^{tL}|f|^2\right)\log\left(e^{tL}|f|^2\right)d\mu + \int Le^{tL}|f|^2 d\mu \\ &= \int\left(Le^{tL}|f|^2\right)\log\left(e^{tL}|f|^2\right)d\mu. \end{aligned}\tag{9.193}$$

Put $h = e^{tL}|f|^2$. We will rewrite the expression in (9.193 as follows. First we notice the equality:

$$(Lf_1)f_2 = L(f_1f_2) - f_1Lf_2 - \Gamma_1(f_1, f_2) \tag{9.194}$$

for appropriately chosen f_1 and f_2. Hence, by L-invariance of μ we have

$$\int(Lf_1)f_2d\mu = -\int f_1Lf_2d\mu - \int\Gamma_1(f_1, f_2)d\mu. \tag{9.195}$$

From (9.193) and (9.195) with $f_1 = e^{tL}|f|^2 = h$, and $f_2 = \log\left(e^{tL}|f|^2\right) = \log h$ we get, by using transformation properties of the squared gradient operator Γ_1,

$$\begin{aligned}\frac{d}{dt}\operatorname{Ent}\left(e^{tL}|f|^2\right) &= -\int hL\log h\,d\mu - \int\Gamma_1(h, \log h)\,d\mu \\ &= -\int h\frac{Lh}{h}d\mu + \frac{1}{2}\int h\frac{\Gamma_1(h,h)}{h^2}d\mu - \int\frac{\Gamma_1(h,h)}{h}d\mu \\ &= -\frac{1}{2}\int\frac{\Gamma_1(h,h)}{h}d\mu = -2\int\Gamma_1\left(h^{1/2}, h^{1/2}\right)d\mu\end{aligned}$$

$$= -2\int \Gamma_1\left(\left(e^{tL}|f|^2\right)^{1/2}, \left(e^{tL}|f|^2\right)^{1/2}\right) d\mu. \tag{9.196}$$

If assertion (i) is true, then (9.196) implies:

$$\frac{d}{dt}\mathrm{Ent}\left(e^{tL}|f|^2\right) \leqslant -2\lambda \mathrm{Ent}\left(e^{tL}|f|^2\right),$$

and consequently $\mathrm{Ent}\left(e^{tL}|f|^2\right) \leqslant e^{-2\lambda t}\mathrm{Ent}\left(|f|^2\right)$ which is assertion (ii). Conversely, if (ii) holds true, then we have

$$\frac{\mathrm{Ent}\left(e^{tL}|f|^2\right) - \mathrm{Ent}\left(|f|^2\right)}{t} \leqslant \frac{e^{-2\lambda t}-1}{t}\mathrm{Ent}\left(|f|^2\right). \tag{9.197}$$

In (9.197) we let $t \downarrow 0$ and we use (9.196) to obtain

$$\lambda \mathrm{Ent}\left(|f|^2\right) \leqslant \int \Gamma_1\left(|f|, |f|\right) d\mu \leqslant \int \Gamma_1\left(\overline{f}, f\right) d\mu. \tag{9.198}$$

The proof of the inequality $\Gamma_1(|f|,|f|) \leqslant \Gamma_1\left(\overline{f}, f\right)$ is given in the proof of Lemma 9.11: see (9.248), (9.249), and (9.250). From (9.198) assertion (i) follows.

This completes the proof of Theorem 9.16. □

Proposition 9.14. *Fix $A \geqslant 0$ and $\lambda > 0$, and let μ be an invariant probability measure. The following assertions are equivalent:*

(i) For all $f \in \mathcal{A}$ the logarithmic Sobolev inequality in (9.184) holds.
(ii) There exists $p \in (1,\infty)$ such that

$$\begin{aligned}\mathrm{Ent}\left(f^p\right) &\leqslant A\int f^p d\mu + \frac{p^2}{4\lambda}\int f^{p-2}\Gamma_1(f,f)\, d\mu \\ &= A\int f^p d\mu - \frac{p^2}{4\lambda(p-1)}\int f^{p-1}Lf\, d\mu\end{aligned} \tag{9.199}$$

for $f \in \mathcal{A}$, $f \geqslant 0$.
(iii) For all $p \in (1,\infty)$ the inequality

$$\begin{aligned}\mathrm{Ent}\left(f^p\right) &\leqslant A\int f^p d\mu + \frac{p^2}{4\lambda}\int f^{p-2}\Gamma_1(f,f)\, d\mu \\ &= A\int f^p d\mu - \frac{p^2}{4\lambda(p-1)}\int f^{p-1}Lf\, d\mu\end{aligned} \tag{9.200}$$

holds for $f \in \mathcal{A}$, $f \geqslant 0$.

Proof. The proof follows by observing that $\Gamma_1(\varphi(f),\varphi(f)) = (\varphi'(f))^2\,\Gamma_1(f,f)$ for all C^1-functions φ and for all $f \in \mathcal{A}$. The choice $\varphi(f) = f^{p/2}$ shows that (i) implies (ii). The choice $\varphi(f) = f^{q/p}$ shows (ii) implies (iii) with q instead of p. Finally, the choice $p = 2$ shows the implication (iii) $\Longrightarrow$ (i), and completes the proof of Proposition 9.14. □

The following result is taken from [Bakry (2006)].

Theorem 9.17. *Let $A \geqslant 0$ and $\lambda > 0$ be two constants, and let $p \in (0,\infty)$. Let the functions $p(t)$ and $m(t)$ be determined by the equalities:*

$$\frac{p(t)-1}{p-1} = e^{4\lambda t}, \quad \textit{and} \quad m(t) = A\left(\frac{1}{t} - \frac{1}{p(t)}\right). \tag{9.201}$$

Then the following assertions are equivalent:

(i) The logarithmic Sobolev inequality (9.184) is satisfied with constants A and λ;
(ii) For all $t > 0$ and $f \in L^p(E,\mu)$ the following inequality holds

$$\left\| e^{tL} f\right\|_{p(t)} \leqslant e^{m(t)} \left\| f\right\|_p. \tag{9.202}$$

Notice that for $A = 0$ we have $\left\|e^{tL}f\right\|_{p(t)} \leqslant \|f\|_p$, and hence the mapping $f \mapsto e^{tL} f$ is contractive form $L^p(E,\mu)$ to $L^{p(t)}(E,\mu)$.

Proof. (i) $\Longrightarrow$ (ii). Fix $t_0 \in (0,\infty)$. Without loss of generality we may and do assume that

$$\|e^{tL} f\|_{p(t)}^{p(t)} = \int \left(e^{tL} f\right)^{p(t)} d\mu \leqslant 1, \quad 0 \leqslant t \leqslant t_0. \tag{9.203}$$

Otherwise we divide $f \geqslant 0$ by $\sup\left\{\left\|e^{tL}f\right\|_{p(t)} : t \in [0,t_0]\right\}$. Define the function $g(t)$, $t > 0$, by

$$g(t) = \exp\left(-A\left(\frac{1}{p} - \frac{1}{p(t)}\right)\right)\left(\int \left(e^{tL} f\right)^{p(t)} d\mu\right)^{1/p(t)}. \tag{9.204}$$

Then a calculation shows the equality:

$$\begin{aligned}
&g'(t)\exp\left(A\left(\frac{1}{p} - \frac{1}{p(t)}\right)\right)\left(\int \left(e^{tL} f\right)^{p(t)} d\mu\right)^{1-1/p(t)} \\
&= -A\frac{p'(t)}{p(t)^2}\int \left(e^{tL} f\right)^{p(t)} d\mu \\
&\quad + \frac{p'(t)}{p(t)^2}\left\{\mathrm{Ent}\left(\left(e^{tL} f\right)^{p(t)}\right) + \int \left(e^{tL} f\right)^{p(t)} \log\left(\int \left(e^{tL} f\right)^{p(t)} d\mu\right) d\mu\right\}
\end{aligned}$$

$$+\int \left(e^{tL}f\right)^{p(t)-1} Le^{tL}f\,d\mu. \tag{9.205}$$

From assertion (iii) in Proposition 9.14 with $p(t)$ instead of p we see

$$\begin{aligned}\operatorname{Ent}\left(\left(e^{tL}f\right)^{p(t)}\right) &\leqslant A\int \left(e^{tL}f\right)^{p(t)} d\mu - \frac{1}{4\lambda}\frac{p(t)^2}{p(t)-1}\int \left(e^{tL}f\right)^{p(t)-1} Le^{tL}f\,d\mu,\\ &= A\int \left(e^{tL}f\right)^{p(t)} d\mu - \frac{p(t)^2}{p'(t)}\int \left(e^{tL}f\right)^{p(t)-1} Le^{tL}f\,d\mu.\end{aligned} \tag{9.206}$$

Then (9.206) together with (9.205) shows:

$$\begin{aligned}&g'(t)\exp\left(A\left(\frac{1}{p}-\frac{1}{p(t)}\right)\right)\left(\int \left(e^{tL}f\right)^{p(t)} d\mu\right)^{1-1/p(t)}\\ &\leqslant -A\frac{p'(t)}{p(t)^2}\int \left(e^{tL}f\right)^{p(t)} d\mu\\ &\quad + \frac{p'(t)}{p(t)^2}\left\{A\int \left(e^{tL}f\right)^{p(t)} d\mu - \frac{p(t)^2}{p'(t)}\int \left(e^{tL}f\right)^{p(t)-1} Le^{tL}f\,d\mu\right.\\ &\qquad\left. + \int \left(e^{tL}f\right)^{p(t)}\log\left(\int \left(e^{tL}f\right)^{p(t)} d\mu\right) d\mu\right\}\\ &\quad + \int \left(e^{tL}f\right)^{p(t)-1} Le^{tL}f\,d\mu\\ &= \frac{p'(t)}{p(t)^2}\int \left(e^{tL}f\right)^{p(t)}\log\left(\int \left(e^{tL}f\right)^{p(t)} d\mu\right) d\mu \leqslant 0\end{aligned} \tag{9.207}$$

where we used (9.203). From (9.207) it follows that $g'(t) \leqslant 0$, $t \in [0, t_0]$, and hence $g(t) \leqslant g(0)$, which shows inequality (9.202) in assertion (ii). Since $t_0 \in (0,\infty)$ is arbitrary assertion (ii) follows from (i).

(ii) $\Longrightarrow$ (i). It suffices to prove assertion (i) in case $\int f^p d\mu = 1$. Again let the function g be defined in (9.204). Now we use $g'(0)$ to show that (i) is a consequence of (ii). In fact from assertion (ii) we get $g(t) \leqslant g(0)$, $t \geqslant 0$, and hence $g'(0) \leqslant 0$. From (9.205) for $t = 0$ and the fact that $\int f^p d\mu = 1$ we see that inequality (9.199) in assertion (ii) of Proposition 9.14 follows. □

Proposition 9.15. *Suppose that there exist constants $c > 0$ and $\gamma \in \mathbb{R}$, such that*

$$0 \leqslant e^{tL}\left|e^{\rho L}f\right|^2 - \left|e^{(\rho+t)L}f\right|^2 \leqslant ce^{-\rho\gamma}e^{\rho L}\left\{e^{tL}\left|f\right|^2 - \left|e^{tL}f\right|^2\right\} \tag{9.208}$$

for all $f \in C_b(E)$, ρ, $t \geqslant 0$. Then the inequality

$$\Gamma_1\left(e^{\rho L}\overline{f}, e^{\rho L}f\right) \leqslant ce^{-\rho\gamma}e^{\rho L}\Gamma_1\left(\overline{f}, f\right) \tag{9.209}$$

holds for all $\rho \geqslant 0$ and all $f \in C_b(E)$. Consequently, the inequality in (9.171) holds. Conversely, if (9.209) holds, then the inequality in (9.208) holds as well. Consequently, if (9.208) or (9.209) is valid, then the inequality in (9.171) holds.

Proof. Let $\rho \geqslant 0$, and $0 \leqslant t \leqslant s$. Then by (9.208) with $e^{(s-t)L}f$ instead of f we get

$$\begin{aligned} & e^{tL}\left|e^{(\rho+s-t)L}f\right|^2 - \left|e^{(\rho+s)L}f\right|^2 \\ & \leqslant ce^{-\gamma\rho}e^{(\rho+t)L}\left|e^{(s-t)L}f\right|^2 - ce^{-\gamma\rho}e^{\rho L}\left|e^{sL}f\right|^2 \\ & = ce^{-\gamma\rho}e^{\rho L}\left(e^{tL}\left|e^{(s-t)L}f\right|^2 - \left|e^{(\rho+s)L}f\right|^2\right). \end{aligned} \tag{9.210}$$

We divide the terms in (9.210) by $t > 0$ and let t tend to zero to obtain:

$$\begin{aligned} & \Gamma_1\left(e^{(s+\rho)L}\overline{f}, e^{(s+\rho)L}f\right) \\ & = L\left|e^{(s+\rho)L}f\right|^2 - \overline{Le^{(\rho+s)L}f}\, e^{(\rho+s)L}f - \overline{e^{(\rho+s)L}f}\, Le^{(\rho+s)L}f \\ & \leqslant ce^{-\gamma\rho}e^{\rho L}\left(L\left|e^{sL}f\right|^2 - \overline{Le^{sL}f}\, e^{sL}f - \overline{e^{sL}f}\, Le^{sL}f\right) \\ & = ce^{-\gamma\rho}e^{\rho L}\Gamma_1\left(e^{sL}\overline{f}, e^{sL}f\right). \end{aligned} \tag{9.211}$$

In order to obtain (9.211) we again employed (9.168). In (9.211) we let s tend to zero to get:

$$\begin{aligned} \Gamma_1\left(e^{\rho L}\overline{f}, e^{\rho L}f\right) & = L\left|e^{\rho L}f\right|^2 - \overline{Le^{\rho L}f}\, e^{\rho L}f - \overline{f}\, Le^{\rho L}f \\ & \leqslant ce^{-\gamma\rho}e^{\rho L}\left(L\left|f\right|^2 - \overline{Lf}\, f - \overline{f}\, Lf\right) \\ & = ce^{-\gamma\rho}e^{\rho L}\Gamma_1\left(\overline{f}, f\right). \end{aligned} \tag{9.212}$$

Notice that (9.212) coincides with the inequality in (9.209). As in Remark 9.13 we see that (9.212) yields

$$\begin{aligned} e^{tL}\left|f\right|^2 - \left|e^{tL}f\right|^2 & = \int_0^t e^{\rho L}\Gamma_1\left(e^{(t-\rho)L}\overline{f}, e^{(t-\rho)L}f\right)\, d\rho \\ & \leqslant \int_0^t e^{\rho L}\left\{ce^{-(t-\rho)\gamma}e^{(t-\rho)L}\Gamma_1\left(\overline{f}, f\right)\right\} d\rho \\ & = \frac{c}{\gamma}\left(1 - e^{-t\gamma}\right)e^{tL}\Gamma_1\left(\overline{f}, f\right). \end{aligned} \tag{9.213}$$

Notice that (9.213) is the same as (9.171).

Next suppose that (9.209) holds. Then by (9.162) with $e^{(t-\sigma)L}f$ instead of f we obtain

$$\begin{aligned}
&e^{tL}\left|e^{\rho L}f\right|^2-\left|e^{tL}e^{\rho L}f\right|^2\\
&=\int_0^t e^{\sigma L}\Gamma_1\left(\overline{e^{(t-\sigma)L}e^{\rho L}f},e^{(t-\sigma)L}e^{\rho L}f\right)d\sigma\\
&\leqslant\int_0^t ce^{-\gamma\rho}e^{\rho L}e^{\sigma L}\Gamma_1\left(\overline{e^{(t-\sigma)L}f},e^{(t-\sigma)L}f\right)d\sigma\\
&=ce^{-\gamma\rho}e^{\rho L}\int_0^t e^{\sigma L}\Gamma_1\left(\overline{e^{(t-\sigma)L}f},e^{(t-\sigma)L}f\right)d\sigma\\
&=ce^{-\gamma\rho}e^{\rho L}\left(e^{tL}\left|f\right|^2-\left|e^{tL}f\right|^2\right).
\end{aligned}\tag{9.214}$$

The inequality in (9.214) is the same as the one in (9.208).

Altogether this proves Proposition 9.15. □

In the following lemma we want to establish conditions in order that the inequality (9.208) or the equivalent one (9.209) is satisfied.

Lemma 9.9. *Let $f\in C_b(E)$, fix $t>0$, and put for $\rho\geqslant 0$*

$$u(\rho)=e^{-\gamma\rho}e^{\rho L}v(0)=e^{-\gamma\rho}e^{\rho L}\left(e^{tL}\left|f\right|^2-\left|e^{tL}f\right|^2\right),\tag{9.215}$$

$$v(\rho)=e^{tL}\left|e^{\rho L}f\right|^2-\left|e^{(\rho+t)L}f\right|^2,\quad \textit{and}$$

$$w(\rho)=e^{tL}\Gamma_1\left(\overline{e^{\rho L}f},e^{\rho L}f\right)-\Gamma_1\left(\overline{e^{(\rho+t)L}f},e^{(\rho+t)L}f\right).\tag{9.216}$$

Suppose $w(\rho)\geqslant\gamma v(\rho)$, $\rho\geqslant 0$. Then $u(\rho)\geqslant v(\rho)$, $\rho\geqslant 0$.

Proof. A calculation shows:

$$u'(\rho)-v'(\rho)=Lu(\rho)-Lv(\rho)+w(\rho)-\gamma u(\rho),\quad \rho\geqslant 0.\tag{9.217}$$

Inserting the inequality $w(\rho)\geqslant\gamma v(\rho)$ in (9.217) shows:

$$u'(\rho)-v'(\rho)\geqslant Lu(\rho)-Lv(\rho)-\gamma\left(u(\rho)-v(\rho)\right),\quad \rho\geqslant 0.\tag{9.218}$$

From (9.218) we see

$$u'(\rho)-v'(\rho)=Lu(\rho)-Lv(\rho)-\gamma\left(u(\rho)-v(\rho)\right)+p(\rho),\quad \rho\geqslant 0,\tag{9.219}$$

where $p(\rho)\geqslant 0$. Then we have

$$u(\rho)-v(\rho)=e^{-\gamma\rho}e^{\rho L}\left(u(0)-v(0)\right)+\int_0^\rho e^{-\gamma(\rho-\sigma)}e^{(\rho-\sigma)L}p(\sigma)d\sigma\geqslant 0.\tag{9.220}$$

This proves Lemma 9.9. □

We observe that $w(\rho) \geqslant \gamma v(\rho)$ for all $\rho \geqslant 0$ if and only if

$$e^{tL}\Gamma_1\left(\overline{g}, g\right) - \Gamma_1\left(\overline{e^{tL}g}, e^{tL}g\right) \geqslant \gamma\left(e^{tL}\left|g\right|^2 - \left|e^{tL}g\right|^2\right) \tag{9.221}$$

for all functions g of the form $g = e^{\rho L}f$, $\rho \geqslant 0$. The following lemma gives conditions in order that the inequality (9.221) is satisfied.

Lemma 9.10. *Suppose that*

$$L\Gamma_1\left(\overline{g}, g\right) - 2\Re\left(\Gamma_1\left(\overline{Lg}, g\right)\right) \geqslant \gamma\Gamma_1\left(\overline{g}, g\right) \tag{9.222}$$

for all functions g of the form $g = e^{\rho L}f$, $\rho \geqslant 0$. Then the inequality in (9.221) is satisfied for such functions.

Proof. [Proof of Lemma 9.10.] We write

$$\begin{aligned}
&e^{tL}\Gamma_1\left(\overline{g}, g\right) - \Gamma_1\left(\overline{e^{tL}g}, e^{tL}g\right) \\
&= \int_0^t e^{\rho L}\left(L\Gamma_1\left(\overline{e^{(t-\rho)L}g}, e^{(t-\rho)L}g\right) - 2\Re\left(\Gamma_1\left(\overline{Le^{(t-\rho)L}g}, e^{(t-\rho)L}g\right)\right)\right) d\rho \\
&\geqslant \gamma\int_0^t e^{\rho L}\Gamma_1\left(\overline{e^{(t-\rho)L}g}, e^{(t-\rho)L}g\right) d\rho = \gamma\left(e^{tL}\left|g\right|^2 - \left|e^{tL}g\right|^2\right).
\end{aligned} \tag{9.223}$$

The inequality in (9.223) proves Lemma 9.10. □

The bilinear mapping $(f, g) \mapsto \Gamma_2(f, g)$, f, $g \in \mathcal{A}$, where

$$\Gamma_2(f, g) = L\Gamma_1(f, g) - \Gamma_1(Lf, g) - \Gamma_1(f, Lg) \tag{9.224}$$

is called the first iterated square gradient operator. The inequality in (9.222) says that

$$\Gamma_2\left(\overline{g}, g\right) \geqslant \gamma\Gamma_1\left(\overline{g}, g\right), \quad g \in \mathcal{A}. \tag{9.225}$$

The following result can also be found as Lemma 1.2 and Lemma 1.3 in [Ledoux (2000)]: proofs go back to [Bakry (1985a,b)]. It shows that there is a close connection between semigroup inequalities and the Γ_2-condition as exhibited in (9.226).

Theorem 9.18. *Let $\gamma \in \mathbb{R}$. The following assertions are equivalent:*

(i) The following inequality holds for all $f \in \mathcal{A}$:

$$\Gamma_2\left(\overline{f}, f\right) - \gamma\Gamma_1\left(\overline{f}, f\right) \geqslant 0. \tag{9.226}$$

(ii) For every $t \geqslant 0$ and $f \in \mathcal{A}$ the following inequality holds:

$$\Gamma_1\left(e^{tL}\overline{f}, e^{tL}f\right) \leqslant e^{-\gamma t}e^{tL}\Gamma_1\left(\overline{f}, f\right). \tag{9.227}$$

(iii) For every $t \geqslant 0$ and $f \in \mathcal{A}$ the following inequality holds:

$$\left(\Gamma_1\left(e^{tL}\overline{f}, e^{tL}f\right)\right)^{1/2} \leqslant e^{-\frac{1}{2}\gamma t} e^{tL}\left(\Gamma_1\left(\overline{f}, f\right)^{1/2}\right). \tag{9.228}$$

(iv) The following inequality holds for all $f \in \mathcal{A}$:

$$\begin{aligned}\Gamma_2\left(\overline{f}, f\right) - \gamma\Gamma_1\left(\overline{f}, f\right) &\geqslant \frac{\Gamma_1\left(\Gamma_1\left(\overline{f}, f\right), \Gamma_1\left(\overline{f}, f\right)\right)}{4\Gamma_1\left(\overline{f}, f\right)} \\ &= \Gamma_1\left(\Gamma_1\left(\overline{f}, f\right)^{1/2}, \Gamma_1\left(\overline{f}, f\right)^{1/2}\right).\end{aligned} \tag{9.229}$$

Notice that the inequality in (9.226) is the same as the one in (9.225). For more details on the iterated square gradient operators see e.g. [Bakry (1985a,b, 1994, 2006, 1991)], [Bakry and Ledoux (2006)], [Ledoux (2000, 2004)], and [Rothaus (1981b,a, 1986)].

Remark 9.15. Let Ψ_1, $\Psi_2 : \mathbb{R}^n \to \mathbb{R}$ be smooth, i.e. $C^{(2)}$-functions, and $F = (f_1, \dots, f_n)$ a vector in $\mathcal{A}^n$. In the proof we employ the following equality:

$$\Gamma_1\left(\Psi_1(F), \Psi_2(F)\right) = \sum_{i,j=1}^{n} X_i^{(1)} X_j^{(2)} \Gamma_1\left(f_i, f_j\right) \tag{9.230}$$

where $X_i^{(k)} = \dfrac{\partial \Psi_k}{\partial x_i}(F)$, $1 \leqslant i \leqslant n$, $k = 1, 2$. With $\Psi_1(g) = \Psi_2(g) = g^2$, $g = \Gamma_1\left(\overline{f}, f\right)^{1/2}$, this shows the equality-sign in (9.229). Compare (9.230) and (9.237) below. In the implication (i) $\Longrightarrow$ (iv) we also need the Hessian of a function f. The Hessian $H(f)$ of f is the bilinear mapping defined in (9.235) below. Its main transformation property is given in (9.236). The equality in (9.230) is a consequence of the equality $\Gamma_1\left(f, g\right) = L(fg) - (Lf)g - f(Lg)$ for appropriately chosen functions f and g together with the transformation property of the operator L: see equality (9.168) and (7.1) with L instead of $-K_0$. In Bakry's terminology the operator L is the generator of a diffusion.

Proof. The implication (iii) $\Longrightarrow$ (ii) follows from the Cauchy-Schwarz inequality in conjunction with (9.228). In fact

$$\left|e^{tL}\left(\overline{g}h\right)\right|^2 \leqslant e^{tL}\left|g\right|^2 \cdot e^{tL}\left|h\right|^2 \leqslant \left\|g\right\|_\infty^2 e^{tL}\left|h\right|^2$$

applied with $g = \mathbf{1}$ and $h = \Gamma_1\left(\overline{f}, f\right)^{1/2}$ shows that (iii) $\Longrightarrow$ (ii).

(ii) $\Longrightarrow$ (i). Subtracting the left-hand side from the right-hand side of (9.227) and dividing by $t > 0$ and letting $t \downarrow 0$ yields:

$$\left(L - \gamma\right)\Gamma_1\left(\overline{f}, f\right) - \Gamma_1\left(L\overline{f}, f\right) - \Gamma_1\left(\overline{f}, Lf\right) \geqslant 0. \tag{9.231}$$

However, the inequality in (9.231) is equivalent to (9.226).

(iv) $\Longrightarrow$ (iii). We fix $f \in \mathcal{A}$ and $t > 0$, and we define the function $\Phi(s)$, $s \in [0, t]$, by

$$\Phi(s) = e^{-\frac{1}{2}\gamma s} e^{sL} \left\{ \left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2} \right\}. \tag{9.232}$$

Then we want to show $\Phi(t) \geqslant \Phi(0)$. Since $\Phi(t) - \Phi(0) = \int_0^t \Phi'(s) ds$, it suffices to prove that $\Phi'(s) \geqslant 0$. Therefore we calculate:

$$\begin{aligned}
&\Phi'(s) \\
&= e^{-\frac{1}{2}\gamma s + sL} \left(L - \frac{1}{2}\gamma \right) \left[\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2} \right] \\
&\quad + e^{-\frac{1}{2}\gamma s + sL} \left[\frac{\partial}{\partial s} \left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2} \right] \\
&= e^{-\frac{1}{2}\gamma s + sL} \left(L - \frac{1}{2}\gamma \right) \left[\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2} \right] \\
&\quad - \frac{1}{2} e^{-\frac{1}{2}\gamma s + sL} \Bigg[\frac{1}{\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2}} \\
&\qquad \left(\Gamma_1 \left(L e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) + \Gamma_1 \left(e^{(t-s)L}\overline{f}, L e^{(t-s)L} f \right) \right) \Bigg] \\
&= e^{-\frac{1}{2}\gamma s + sL} \left(L - \frac{1}{2}\gamma \right) \left[\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2} \right] \\
&\quad - \frac{1}{2} e^{-\frac{1}{2}\gamma s + sL} \Bigg[\frac{1}{\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2}} \\
&\qquad \left(L\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) - \Gamma_2 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right) \Bigg] \\
&= e^{-\frac{1}{2}\gamma s + sL} \left(L - \frac{1}{2}\gamma \right) \left[\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2} \right] \\
&\quad - \frac{1}{2} e^{-\frac{1}{2}\gamma s + sL} \left[\frac{L\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right)}{\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2}} \right] \\
&\qquad + \frac{1}{2} e^{-\frac{1}{2}\gamma s + sL} \left[\frac{\Gamma_2 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right)}{\left(\Gamma_1 \left(e^{(t-s)L}\overline{f}, e^{(t-s)L} f \right) \right)^{1/2}} \right]
\end{aligned}$$

$$\left(L\left(g^2\right) = 2gLg + \Gamma_1\left(g,g\right) \text{ with } g = \left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}\right)$$

$$
\begin{aligned}
&= e^{-\frac{1}{2}\gamma s+sL}\left(L-\frac{1}{2}\gamma\right)\left[\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}\right] \\
&\quad -\frac{1}{2}e^{-\frac{1}{2}\gamma s+sL}\left[2L\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}\right] \\
&\quad -\frac{1}{2}e^{-\frac{1}{2}\gamma s+sL}\left[\frac{\Gamma_1\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)^{1/2}, \Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)^{1/2}\right)}{\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}}\right] \\
&\quad +\frac{1}{2}e^{-\frac{1}{2}\gamma s+sL}\left[\frac{\Gamma_2\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)}{\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}}\right]
\end{aligned}
$$

$$\left(\Gamma_1\left(g^2,g^2\right) = 4g^2\Gamma_1\left(g,g\right) \text{ with } g = \left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}\right)$$

$$
\begin{aligned}
&=-\frac{1}{2}e^{-\frac{1}{2}\gamma s+sL}\left[\frac{\Gamma_1\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right), \Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)}{4\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{3/2}}\right] \\
&\quad +\frac{1}{2}e^{-\frac{1}{2}\gamma s+sL}\left[\frac{\Gamma_2\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)-\gamma\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)}{\left(\Gamma_1\left(e^{(t-s)L}\overline{f}, e^{(t-s)L}f\right)\right)^{1/2}}\right].
\end{aligned}
\tag{9.233}
$$

Put $g = e^{(t-s)L}f$. In order that the expression in (9.233) is positive it suffices to prove the inequality:

$$\Gamma_1\left(\overline{g},g\right)\left(\Gamma_2\left(\overline{g},g\right)-\gamma\Gamma_1\left(\overline{g},g\right)\right) \geqslant \frac{1}{4}\Gamma_1\left(\Gamma_1\left(\overline{g},g\right), \Gamma_1\left(\overline{g},g\right)\right). \tag{9.234}$$

The inequality in (9.234) is a consequence of assertion (iv).

The implication (i) $\Longrightarrow$ (iv) remains to be shown. Here we use the fact that L generates a diffusion. We will start from (9.226), i.e. from

$$\Gamma_2\left(\overline{f},f\right)-\gamma\Gamma_1\left(\overline{f},f\right) \geqslant 0$$

for all $f \in \mathcal{A}$. Without loss of generality we assume that the function f is real-valued. For a given function $f \in \mathcal{A}$ we introduce its Hessian $H(f)$ as the bilinear form:

$$H(f)\left(g,h\right) = \frac{1}{2}\left[\Gamma_1\left(\Gamma_1\left(f,g\right),h\right)+\Gamma_1\left(\Gamma_1\left(f,h\right),g\right)-\Gamma_1\left(f,\Gamma_1\left(g,h\right)\right)\right], \tag{9.235}$$

g, $h \in \mathcal{A}$. Let $\Psi_1, \Psi_2 : \mathbb{R}^n \to \mathbb{R}$ be smooth functions, and let $F = (f_1, \dots, f_n)$ be a vector in $\mathcal{A}^n$. Put $X_i^{(1)} = \dfrac{\partial \Psi_1}{\partial x_i}(F)$ and

$X_{i,j}^{(1)} = \frac{\partial^2 \Psi_1}{\partial x_i \partial x_j}(F)$, $1 \leqslant i, j \leqslant n$. A similar convention is used for Ψ_2 with (2) instead of (1). A cumbersome calculation shows that the first iterated square gradient operator Γ_2 satisfies:

$$\begin{aligned}\Gamma_2\left(\Psi_1(F), \Psi_2(F)\right) = &\sum_{i,j=1}^{n} X_i^{(1)} X_j^{(2)} \Gamma_2\left(f_i, f_j\right) \\ &+ \sum_{i,j,k=1}^{n} \left(X_i^{(2)} X_{j,k}^{(1)} + X_i^{(1)} X_{j,k}^{(2)}\right) H\left(f_i\right)\left(f_j, f_k\right) \\ &+ \sum_{i,j,k,\ell=1}^{n} X_{i,j}^{(1)} X_{k,\ell}^{(2)} \Gamma_1\left(f_i, f_k\right) \Gamma_1\left(f_j, f_\ell\right). \end{aligned} \tag{9.236}$$

For the definition of the iterated squared gradient operator Γ_2 see (9.224). In the calculation to obtain (9.236) a similar but much simpler formula is used:

$$\Gamma_1\left(\Psi_1(F), \Psi_2(F)\right) = \sum_{i,j=1}^{n} X_i^{(1)} X_j^{(2)} \Gamma_1\left(f_i, f_j\right). \tag{9.237}$$

In Remark 9.15 it was shown how the formula in (9.237) can be obtained.

Again, let $\Psi : \mathbb{R}^n \to \mathbb{R}$ be a "smooth" function. If the function $F \mapsto \Psi(F)$ varies among all real polynomials of second order, then the function

$$\left(X_1, \dots, X_n; \left(X_{i,j}\right)_{i,j=1}^{n}\right) \mapsto \Gamma_2\left(\Psi(F), \Psi(F)\right) - \gamma \Gamma_1\left(\Psi(F), \Psi(F)\right) \tag{9.238}$$

is a positive polynomial. We may apply this for $n = 2$, $f_1 = f$, $f_2 = g$, and the function $\Psi(f, g)$ chosen in such a way that $X_2 = X_{1,1} = X_{2,2} = 0$. Then from (9.236), (9.237) and (9.238) we get:

$$\begin{aligned} &X_1^2\left(\Gamma_2(f, f) - \gamma \Gamma_1(f, f)\right) + 4 X_1 X_{1,2} H(f)(f, g) \\ &\quad + 2 X_{1,2}^2 \left(\Gamma_1(f, g)^2 + \Gamma_1(f, f)\Gamma_1(g, g)\right) \geqslant 0 \end{aligned} \tag{9.239}$$

for all X_1 and $X_{1,2} \in \mathbb{R}\backslash\{0\}$. Then we choose $\Psi(f, g)$ in such a way that $X_1\left(\Gamma_2(f, f) - \gamma \Gamma_1(f, f)\right) = -2 X_{1,2} H(f)(f, g)$. From (9.239) we then infer:

$$4\left(H(f)(f, g)\right)^2 \leqslant 2\left(\Gamma_2(f, f) - \gamma \Gamma_1(f, f)\right)\left(\Gamma_1(f, g)^2 + \Gamma_1(f, f)\Gamma_1(g, g)\right). \tag{9.240}$$

Since $2H(f)(f, g) = \Gamma_1\left(\Gamma_1(f, f), g\right)$, (9.240) implies

$$\begin{aligned} &\left(\Gamma_1\left(\Gamma_1(f, f), g\right)\right)^2 \\ &\leqslant 2\left(\Gamma_2(f, f) - \gamma \Gamma_1(f, f)\right)\left(\Gamma_1(f, g)^2 + \Gamma_1(f, f)\Gamma_1(g, g)\right) \end{aligned}$$

(use the inequality $\Gamma_1(f,g)^2 \leqslant \Gamma_1(f,f)\Gamma_1(g,g)$)

$$\leqslant 4\left(\Gamma_2(f,f) - \gamma\Gamma_1(f,f)\right)\Gamma_1(f,f)\Gamma_1(g,g). \tag{9.241}$$

Choosing $g = \Gamma_1(f,f)$, and employing (9.241) entails (9.229) with the real function f instead of a complex function $f \in \mathcal{A}$. By splitting a complex function in its real and imaginary part we see that (9.229) follows for all $f \in \mathcal{A}$.

This completes the proof of the implication (i) $\Longrightarrow$ (iv), and concludes the proof of Theorem 9.18. □

Proposition 9.16. *Suppose that (9.222) is satisfied for all functions $g \in D(L)$. Then the equivalent inequalities (9.208) and (9.209) in Proposition 9.15 are satisfied with $c = 1$. If $\gamma > 0$, then the operator L has a spectral gap $\geqslant \gamma$.*

Proof. [Proof of Proposition 9.16.] If (9.222) is satisfied for all functions $g \in D(L)$, then by Lemma 9.10 the inequality (9.221) is satisfied for all functions $g \in D(L)$. Lemma 9.9 implies that

$$e^{-\gamma\rho}e^{\rho L}\left(e^{tL}|f|^2 - \left|e^{tL}f\right|^2\right) \geqslant e^{tL}\left|e^{\rho L}f\right|^2 - \left|e^{(\rho+t)L}f\right|^2. \tag{9.242}$$

Proposition 9.15 and (9.242) show that the equivalent inequalities (9.208) and (9.209) in Proposition 9.15 are satisfied with $c = 1$. Hence we obtain the inequality in (9.171) with $c = 1$:

$$\begin{aligned} e^{tL}|f|^2 - \left|e^{tL}f\right|^2 &= \int_0^t e^{\rho L}\Gamma_1\left(e^{(t-\rho)L}\overline{f}, e^{(t-\rho)L}f\right) d\rho \\ &\leqslant \int_0^t e^{\rho L}\left\{e^{(t-\rho)\gamma}e^{(t-\rho)L}\Gamma_1\left(\overline{f}, f\right)\right\} d\rho \\ &= \frac{1}{\gamma}\left(1 - e^{-t\gamma}\right)e^{tL}\Gamma_1\left(\overline{f}, f\right). \end{aligned} \tag{9.243}$$

Let μ be an invariant probability measure such that $\lim_{t\to\infty} e^{tL}f(x) = \int f d\mu$ for all $f \in C_b(E)$ and $x \in E$. The existence and uniqueness of such an invariant probability measure is guaranteed by Orey's convergence theorem: see Theorem 10.2, and also (9.104). It is required that the Markov process is Harris recurrent. The monotonicity property in Lemma 10.15 of Chapter 10 implies that this limit exists by letting $t > 0$ tend to ∞ instead of $n \in \mathbb{N}$. Then by integrating (9.243) against μ and taking the limit as $t \to \infty$, we find

$$\gamma\left(\int |f|^2 d\mu - \left|\int f d\mu\right|^2\right) \leqslant \int \Gamma_1\left(\overline{f}, f\right) d\mu. \tag{9.244}$$

From (9.244) and Definition 9.14 the claim in Proposition 9.16 readily follows. □

The inequality (9.252) below is a consequence of equality (9.162) in Theorem 9.15. For convenience we insert a (short) proof here as well. The inequality in (9.253) employs the full power of Theorem 9.18. The proof of Theorem 9.19 requires the following lemma.

Lemma 9.11. *Suppose that the constant* γ *satisfies the inequality in (9.251) in Theorem 9.19 below. Let* $f \in \mathcal{A}$, *and* $s \geqslant 0$. *Then the following inequality holds:*

$$\frac{\Gamma_1\left(e^{sL}\left|f\right|^2, e^{sL}\left|f\right|^2\right)}{e^{sL}\left|f\right|^2} \leqslant e^{-\gamma s} e^{sL}\left\{\frac{\Gamma_1\left(\left|f\right|^2, \left|f\right|^2\right)}{\left|f\right|^2}\right\}. \tag{9.245}$$

In addition, the following inequality holds:

$$\frac{\Gamma_1\left(\left|f\right|^2, \left|f\right|^2\right)}{\left|f\right|^2} \leqslant 4\Gamma_1\left(\overline{f}, f\right). \tag{9.246}$$

If in (9.246) the function f is real-valued, then this inequality is in fact an equality.

Proof. From inequality (9.228) in assertion (iii) of Theorem 9.18 we infer

$$\Gamma_1\left(e^{sL}\left|f\right|^2, e^{sL}\left|f\right|^2\right)^{1/2} \leqslant e^{-\frac{1}{2}\gamma s} e^{sL}\left(\Gamma_1\left(\left|f\right|^2, \left|f\right|^2\right)^{1/2}\right)$$
$$= e^{-\frac{1}{2}\gamma s} e^{sL}\left\{\left|f\right|\left(\frac{\Gamma_1\left(\left|f\right|^2, \left|f\right|^2\right)}{\left|f\right|^2}\right)^{1/2}\right\}$$

(Cauchy-Schwarz inequality)

$$\leqslant e^{-\frac{1}{2}\gamma s}\left(e^{sL}\left|f\right|^2\right)^{1/2}\left(e^{sL}\left\{\frac{\Gamma_1\left(\left|f\right|^2, \left|f\right|^2\right)}{\left|f\right|^2}\right\}\right)^{1/2}. \tag{9.247}$$

The inequality in (9.245) easily follows from (9.247). The equality in (9.246) follows from the transformation rules of the squared gradient operator Γ_1. More precisely, with $f = u + iv$, u, v real and imaginary part of f, we have

$$\Gamma_1\left(\left|f\right|^2, \left|f\right|^2\right) = 4u^2\Gamma_1(u, u) + 8uv\Gamma_1(u, v) + 4v^2\Gamma_1(v, v), \tag{9.248}$$

and

$$4|f|^2\,\Gamma_1\left(\overline{f},f\right) = 4\left(u^2+v^2\right)\left(\Gamma_1(u,u)+\Gamma_1(u,u)\right). \tag{9.249}$$

Since

$$2uv\Gamma_1(u,v) \leqslant 2\left|u\sqrt{\Gamma_1(v,v)}\right|\cdot\left|v\sqrt{\Gamma_1(u,u)}\right| \leqslant u^2\Gamma_1(v,v)+v^2\Gamma_1(u,u), \tag{9.250}$$

the inequality in (9.246) readily follows from (9.248), (9.249) and (9.250).

This completes the proof of Lemma 9.11. □

Theorem 9.19. *Suppose that the constant $\gamma \in \mathbb{R}$ satisfies one of the equivalent conditions in Theorem 9.18 for the operator L: i.e.*

$$\Gamma_2\left(\overline{f},f\right) \geqslant \gamma\Gamma_1\left(\overline{f},f\right) \quad \textit{for all } f \in \mathcal{A}. \tag{9.251}$$

Then the following inequalities hold for $f \in \mathcal{A}$ and $t \geqslant 0$:

$$e^{tL}\left(|f|^2\right) - \left|e^{tL}f\right|^2 \leqslant \frac{1-e^{-\gamma t}}{\gamma}e^{tL}\left(\Gamma_1\left(\overline{f},f\right)\right), \quad \textit{and} \tag{9.252}$$

$$e^{tL}\left(|f|^2\log|f|^2\right) - e^{tL}\left(|f|^2\right)\log\left(e^{tL}\left(|f|^2\right)\right) \leqslant 4\frac{1-e^{-\gamma t}}{\gamma}e^{tL}\left(\Gamma_1\left(\overline{f},f\right)\right). \tag{9.253}$$

The inequality in (9.252) can be called a pointwise Poincaré inequality. It is a consequence of assertion (ii) of Theorem 9.18. The inequality in (9.253) may be called a logarithmic Sobolev inequality. Its proof is based on the assertion (iii) in Theorem 9.18, which is a consequence of assertion (iv). It is clear that assertion (iv) is an improvement of our basic assumption (9.251).

Proof. Let $f \in \mathcal{A}$ and $t > 0$. Then we have

$$\begin{aligned}
e^{tL}|f|^2 - \left|e^{tL}f\right|^2 &= \int_0^t \frac{\partial}{\partial s}\left(e^{sL}\left|e^{(t-s)L}f\right|^2\right)ds \\
&= \int_0^t e^{sL}\left\{L\left|e^{(t-s)L}f\right|^2 - \left(L\overline{e^{(t-s)L}f}\right)e^{(t-s)L}f - \overline{e^{(t-s)L}f}\left(Le^{(t-s)L}f\right)\right\}ds \\
&= \int_0^t e^{sL}\Gamma_1\left(\overline{e^{(t-s)L}f},e^{(t-s)L}f\right)ds.
\end{aligned} \tag{9.254}$$

We employ (9.227) in assertion (ii) of Theorem 9.18 and use the identity in (9.254) to obtain:

$$e^{tL}|f|^2 - \left|e^{tL}f\right|^2 \leqslant \int_0^t e^{-\gamma(t-s)}e^{sL}e^{(t-s)L}\Gamma_1\left(\overline{f},f\right)ds$$

$$= \frac{1 - e^{-\gamma t}}{\gamma} e^{tL} \Gamma_1 \left(\overline{f}, f\right). \tag{9.255}$$

The inequality in (9.255) is the same as the one in (9.252).

The proof of inequality (9.253) is similar, be it (much) more sophisticated. In fact we write:

$$\begin{aligned}
&e^{tL}\left(|f|^2 \log |f|^2\right) - e^{tL}\left(|f|^2\right) \log\left(e^{tL}\left(|f|^2\right)\right) \\
&= \int_0^t \frac{\partial}{\partial s}\left\{e^{sL}\left(\left(e^{(t-s)L}|f|^2\right)\log\left(e^{(t-s)L}|f|^2\right)\right)\right\} ds \\
&= \int_0^t e^{sL}\left\{L\left(\left(e^{(t-s)L}|f|^2\right)\log\left(e^{(t-s)L}|f|^2\right)\right)\right. \\
&\qquad -L\left(e^{(t-s)L}|f|^2\right)\log\left(e^{(t-s)L}|f|^2\right) \\
&\qquad \left. -\left(e^{(t-s)L}|f|^2\right)L\log\left(e^{(t-s)L}|f|^2\right)\right\} ds \\
&= \int_0^t e^{sL}\left\{\Gamma_1\left(e^{(t-s)L}|f|^2, \log\left(e^{(t-s)L}|f|^2\right)\right)\right\} ds \\
&= \int_0^t e^{sL}\left\{\frac{\Gamma_1\left(e^{(t-s)L}|f|^2, e^{(t-s)L}|f|^2\right)}{e^{(t-s)L}|f|^2}\right\} ds.
\end{aligned} \tag{9.256}$$

An appeal to inequality (9.245) in Lemma 9.11 and employing the equality in (9.256) yields:

$$\begin{aligned}
&e^{tL}\left(|f|^2 \log |f|^2\right) - e^{tL}\left(|f|^2\right) \log\left(e^{tL}\left(|f|^2\right)\right) \\
&\leqslant \int_0^t e^{-\gamma(t-s)} e^{sL} e^{(t-s)L}\left\{\frac{\Gamma_1\left(|f|^2, |f|^2\right)}{|f|^2}\right\} ds \\
&= \frac{1-e^{-\gamma t}}{\gamma} e^{tL}\left\{\frac{\Gamma_1\left(|f|^2, |f|^2\right)}{|f|^2}\right\} \\
&\leqslant 4\frac{1-e^{-\gamma t}}{\gamma} e^{tL}\left(\Gamma_1\left(\overline{f}, f\right)\right).
\end{aligned} \tag{9.257}$$

The inequality (9.257) shows (9.253) and completes the proof of Theorem 9.19. □

The following theorem contains some sufficient conditions in order that an operator L possesses a spectral gap in $L^2(E, \mu)$, where μ is an invariant probability measure on the Borel field $\mathcal{E}$ of E.

Theorem 9.20. *Let L be the generator of a diffusion process with transition probability function $P(t,x,\cdot)$, $t \geqslant 0$, $x \in E$. Suppose that the following conditions are satisfied:*

(a) $\Gamma_2\left(\overline{f}, f\right) \geqslant \gamma\Gamma_1\left(\overline{f}, f\right)$ *for all* $f \in \mathcal{A}$.
(b) All probability measures $B \mapsto P(t,x,B)$, $B \in \mathcal{E}$, *with* $(t,x) \in (0,\infty)\times E$ *are equivalent, in the sense that they have the same null-sets.*
(c) The operator L has an invariant probability measure μ.

If in (a) $\gamma > 0$, then the spectral gap of L, $\operatorname{gap}(L)$, *in* $L^2(E,\mu)$ *satisfies:* $\operatorname{gap}(L) \geqslant \gamma$.

Proof. By invoking (9.252) in Theorem 9.19 we have

$$e^{tL}\left(|f|^2\right) - \left|e^{tL}f\right|^2 \leqslant \frac{1-e^{-\gamma t}}{\gamma} e^{tL}\left(\Gamma_1\left(\overline{f}, f\right)\right), \ f \in \mathcal{A}. \tag{9.258}$$

From (9.258), and the invariance of the measure μ we get

$$\int \left(|f|^2\right) d\mu - \int \left|e^{tL}f\right|^2 d\mu \leqslant \frac{1-e^{-\gamma t}}{\gamma} \int \left(\Gamma_1\left(\overline{f}, f\right)\right) d\mu, \ f \in \mathcal{A}. \tag{9.259}$$

Suppose that $\gamma > 0$. The recurrence of the underlying Markov process in conjunction with Orey's convergence theorem (see the arguments in the proof of Proposition 9.16) shows the following inequality by letting t tend to ∞ in (9.259):

$$\int \left(|f|^2\right) d\mu - \left|\int f\, d\mu\right|^2 \leqslant \frac{1}{\gamma} \int \left(\Gamma_1\left(\overline{f}, f\right)\right) d\mu, \ f \in \mathcal{A}. \tag{9.260}$$

The assertion in Theorem 9.20 then follows from (9.260) and the definition of L^2-spectral gap. □

Example 9.1. Next let $E = \mathbb{R}^d$, and L be the differential operator:

$$Lf = \frac{1}{2}\sum_{j,k=1}^{d} a_{j,k}\partial_j\partial_k f + \sum_{j=1}^{d} b_j \partial_j f, \quad f \in C_b^{(2)}\left(\mathbb{R}^d\right), \tag{9.261}$$

where $\partial_j f = \dfrac{\partial f}{\partial x_j}$, $1 \leqslant j \leqslant d$. It is assumed that the coefficients $a_{j,k}$, and b_j, $1 \leqslant j \leqslant d$, $1 \leqslant k \leqslant d$, are space dependent and twice continuously differentiable. Then the corresponding square gradient operator is given by

$$\Gamma_1(f,g) = \sum_{j,k=1}^{d} a_{j,k}\partial_j f \cdot \partial_k g. \tag{9.262}$$

Let f and g be functions in $C^{(3)}\left(\mathbb{R}^d\right)$. We want to simplify an expression of the form

$$L\Gamma_1(f,g)-\Gamma_1(Lf,g)-\Gamma_1(f,Lg). \tag{9.263}$$

Notice that if $f=\overline{g}$, then (9.263) is the same as (9.222). In order to rewrite (9.263) we need the following proposition. This proposition is also valid for general diffusion operators L.

Proposition 9.17. *Let the functions f, g and h belong to $C^{(3)}\left(\mathbb{R}^d\right)$. Then the following identities hold:*

$$\begin{aligned} L(fgh) &= (Lf)\,gh+f\,(Lg)\,h+fg\,(Lh)\\ &\quad +\Gamma_1(f,g)\,h+f\Gamma_1(g,h)+g\Gamma_1(f,h), \quad \textit{and}\\ \Gamma_1(fg,h) &= \Gamma_1(f,g)\,h+g\Gamma_1(f,h). \end{aligned} \tag{9.264}$$

Proposition 9.18. *Let the functions f and g belong to $C^{(3)}\left(\mathbb{R}^d\right)$. Then*

$$\begin{aligned} &L\Gamma_1(f,g)-\Gamma_1(Lf,g)-\Gamma_1(f,Lg)\\ &=\sum_{j,k=1}^d\left\{La_{j,k}-\sum_{n=1}^d a_{n,k}\partial_n b_j-\sum_{n=1}^d a_{n,j}\partial_n b_k\right\}\partial_j f\partial_k g. \end{aligned} \tag{9.265}$$

Proof. [Proof of Proposition 9.18.] First we rewrite

$$\begin{aligned} L\Gamma_1(f,g) &= \sum_{j,k=1}^d L\left(a_{j,k}\partial_j f\cdot\partial_k g\right)\\ &=\sum_{j,k=1}^d\left\{(La_{j,k})\,\partial_j f\cdot\partial_k g+a_{j,k}\,(L\partial_j f)\,\partial_k g+a_{j,k}\partial_j f\,(L\partial_k g)\right.\\ &\quad\left.+\Gamma_1(a_{j,k},\partial_j f)\,\partial_k g+\Gamma_1(a_{j,k},\partial_k g)\,\partial_j f+a_{j,k}\Gamma_1(\partial_j f,\partial_k g)\right\}\\ &=\sum_{j,k=1}^d(La_{j,k})\,\partial_j f\cdot\partial_k g+\sum_{j,k=1}^d a_{j,k}\,(L\partial_j f)\,\partial_k g+\sum_{j,k=1}^d a_{j,k}\partial_j f\,(L\partial_k g)\\ &\quad+\sum_{j,k=1}^d\Gamma_1(a_{j,k},\partial_j f)\,\partial_k g+\sum_{j,k=1}^d\Gamma_1(a_{j,k},\partial_k g)\,\partial_j f\\ &\quad+\sum_{j,k=1}^d a_{j,k}\Gamma_1(\partial_j f,\partial_k g)\\ &=\sum_{j,k=1}^d(La_{j,k})\,\partial_j f\cdot\partial_k g \end{aligned}$$

$$
\begin{aligned}
&+ \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{j,k} a_{n,m} \partial_n \partial_m \partial_j f \cdot \partial_k g + \sum_{j,k=1}^{d} \sum_{n=1}^{d} a_{j,k} b_n \partial_n \partial_j f \cdot \partial_k g \\
&+ \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{j,k} a_{n,m} \partial_j f \cdot \partial_n \partial_m \partial_k g + \sum_{j,k=1}^{d} \sum_{n=1}^{d} a_{j,k} b_n \partial_j f \cdot \partial_n \partial_k g \\
&+ \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{n,m} \partial_n a_{j,k} \cdot \partial_m \partial_j f \cdot \partial_k g \\
&+ \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{n,m} \partial_n a_{j,k} \cdot \partial_m \partial_k g \cdot \partial_j f \\
&+ \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{j,k} a_{n,m} \partial_n \partial_j f \cdot \partial_m \partial_k g. \qquad (9.266)
\end{aligned}
$$

We also rewrite

$$
\begin{aligned}
\Gamma_1 (Lf, g) &= \Gamma_1 \left(\sum_{j,k=1}^{d} a_{j,k} \partial_j \partial_k f + \sum_{j=1}^{d} b_j \partial_j f, g \right) \\
&= \sum_{j,k=1}^{d} \Gamma_1 (a_{j,k} \partial_j \partial_k f, g) + \sum_{j=1}^{d} \Gamma_1 (b_j \partial_j f, g) \\
&= \sum_{j,k=1}^{d} a_{j,k} \Gamma_1 (\partial_j \partial_k f, g) + \sum_{j,k=1}^{d} \Gamma_1 (a_{j,k}, g) \partial_j \partial_k f \\
&\quad + \sum_{j=1}^{d} b_j \Gamma_1 (\partial_j f, g) + \sum_{j=1}^{d} \Gamma_1 (b_j, g) \partial_j f \\
&= \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{j,k} a_{n,m} \partial_n \partial_j \partial_k f \cdot \partial_m g \\
&\quad + \sum_{j,k=1}^{d} \sum_{n,m=1}^{d} a_{n,m} \partial_n a_{j,k} \cdot \partial_j \partial_k f \cdot \partial_m g \\
&\quad + \sum_{j=1}^{d} b_j \sum_{n,m=1}^{d} a_{n,m} \partial_n \partial_j f \cdot \partial_m g \\
&\quad + \sum_{j=1}^{d} \sum_{n,m=1}^{d} a_{n,m} \partial_n b_j \cdot \partial_j f \cdot \partial_m g. \qquad (9.267)
\end{aligned}
$$

By the same token we have

$$
\begin{aligned}
\Gamma_1(f, Lg) &= \Gamma_1\left(f, \sum_{j,k=1}^{d} a_{j,k}\partial_j\partial_k g + \sum_{j=1}^{d} b_j\partial_j g\right) \\
&= \sum_{j,k=1}^{d} a_{j,k}\Gamma_1(f, \partial_j\partial_k g) + \sum_{j,k=1}^{d} \Gamma_1(a_{j,k}, f)\,\partial_j\partial_k g \\
&\quad + \sum_{j=1}^{d} b_j\Gamma_1(f, \partial_j g) + \sum_{j=1}^{d} \Gamma_1(b_j, f)\,\partial_j g \\
&= \sum_{j,k=1}^{d}\sum_{n,m=1}^{d} a_{j,k}a_{n,m}\partial_n f\cdot\partial_m\partial_j\partial_k g \\
&\quad + \sum_{j,k=1}^{d}\sum_{n,m=1}^{d} a_{n,m}\partial_n a_{j,k}\partial_m f\cdot\partial_j\partial_k g \\
&\quad + \sum_{j=1}^{d} b_j\sum_{n,m=1}^{d} a_{n,m}\partial_n f\cdot\partial_m\partial_j g \\
&\quad + \sum_{j=1}^{d}\sum_{n,m=1}^{d} a_{n,m}\partial_n b_j\cdot\partial_m f\cdot\partial_j g. \qquad (9.268)
\end{aligned}
$$

From (9.266), (9.267) and (9.268) we infer:

$$
\begin{aligned}
& L\Gamma_1(f, g) - \Gamma_1(Lf, g) - \Gamma_1(f, Lg) \\
&= \sum_{j,k=1}^{d} (La_{j,k})\,\partial_j f\partial_k g - \sum_{j=1}^{d}\sum_{n,k=1}^{d} a_{n,k}\partial_n b_j\left\{\partial_j f\partial_k g + \partial_k f\partial_j g\right\} \\
&= \sum_{j,k=1}^{d}\left\{La_{j,k} - \sum_{n=1}^{d} a_{n,k}\partial_n b_j - \sum_{n=1}^{d} a_{n,j}\partial_n b_k\right\}\partial_j f\partial_k g \\
&= \sum_{j,k=1}^{d} L_b(A)_{j,k}\,\partial_j f\partial_k g \qquad (9.269)
\end{aligned}
$$

where $L_b(C)$ is a matrix with entries:

$$
L_b(C)_{j,k} = Lc_{j,k} - \sum_{n=1}^{d} c_{n,k}\partial_n b_j - \sum_{n=1}^{d} c_{n,j}\partial_n b_k. \qquad (9.270)
$$

Here C is the matrix with entries $c_{j,k}$ and b stands for the column vector with components b_j. The symbol L_b can be considered as a mapping which assigns to a square matrix consisting of functions again a square matrix consisting of functions. The operator L is the original differential operator given in (8.132).

The proof of Proposition 9.18 is now complete. □

If we want to check an inequality like (9.208) or, what is equivalent, (9.209), then it is probably better to consider the corresponding stochastic differential equations.

Corollary 9.7. *Fix $\gamma \in \mathbb{R}$. Suppose that the inequality*

$$\sum_{j,k=1}^{d}\left\{La_{j,k}-\sum_{n=1}^{d}a_{n,k}\partial_n b_j-\sum_{n=1}^{d}a_{n,j}\partial_n b_k\right\}\overline{\lambda_j}\lambda_k \geqslant \gamma\sum_{j,k=1}^{d}a_{j,k}\overline{\lambda_j}\lambda_k \tag{9.271}$$

holds for all complex vectors $(\lambda_1,\dots,\lambda_d)\in\mathbb{C}^d$. Then

$$\Gamma_1\left(e^{\rho L}\overline{f}, e^{\rho L}f\right)\leqslant e^{-\rho\gamma}e^{\rho L}\Gamma_1\left(\overline{f},f\right) \tag{9.272}$$

for all $f\in D(L)$ and $\rho\geqslant 0$.

Remark 9.16. The inequality in (9.271) says that in matrix sense the following inequalities hold: $L_b(A)\geqslant\gamma A$. Here we used the notation as in (9.270), and A is the symmetric matrix with entries $a_{j,k}$, $1\leqslant j$, $k\leqslant d$.

For some of our applications we will need the following somewhat technical proposition. The result is due to Rothaus (see [Rothaus (1986)]) and the inequality in (9.274) is named after him. A proof of the inequality (9.274) can be found in [Deuschel and Stroock (1989)]. Another proof can be found in [Bakry (1994)]; for completeness we insert an outline of a proof.

Proposition 9.19. *Let μ be a probability measure on the Borel field of E and let $f\in C_b(E)$. Fix $p\geqslant 2$. Then the following inequalities hold:*

$$\left(\int|f|^p\,d\mu\right)^{2/p}\leqslant\left|\int f\,d\mu\right|^2+(p-1)\left(\int\left|f-\int f\,d\mu\right|^p d\mu\right)^{2/p}, \tag{9.273}$$

and

$$\operatorname{Ent}\left(|f|^2\right)\leqslant 2\int\left|f-\int f d\mu\right|^2 d\mu+\operatorname{Ent}\left(\left|f-\int f\,d\mu\right|^2\right). \tag{9.274}$$

Proof. Put $\widehat{f}=f-\int f d\mu$. By homogeneity we assume that f is the form $f=1+tg$ where the function g is such that $\Re\int g\,d\mu=0$ and $\int|g|^2\,d\mu=1$, and where $t\geqslant 0$. Then $f-\int f\,d\mu=tg$, and $\int|1+tg|^2\,d\mu=1+t^2$. Put

$$F_1(t)=\left(\int|1+tg|^p\,d\mu\right)^{2/p}-(p-1)t^2\left(\int|g|^p\,d\mu\right)^{2/p}, \quad\text{and} \tag{9.275}$$

$$F_2(t) = \int |1+tg|^2 \log |1+tg|^2 \, d\mu - \left(1+t^2\right) \log \left(1+t^2\right) \\ - t^2 \int |g|^2 \log |g|^2 \, d\mu. \tag{9.276}$$

We will show $F_1(t) \leqslant 1$ and $F_2(t) \leqslant 2t^2$, $t \geqslant 0$. The inequality in (9.273) is a consequence of $F_1(t) \leqslant 1$, and similarly (9.276) follows from $F_2(t) \leqslant 2t^2$. We will use the following representations:

$$F_k(t) = F_k(0) + tF_k'(0) + \int_0^t (t-s)\, F_k''(s) ds, \quad k = 1, 2. \tag{9.277}$$

Then

$$F_1'(t) = 2\left(\int |1+tg|^p \, d\mu\right)^{\frac{2}{p}-1} \int |1+tg|^{p-2} \left(\Re g + t\,|g|^2\right) d\mu \\ - 2(p-1)t\left(\int |g|^p \, d\mu\right)^{\frac{2}{p}}. \tag{9.278}$$

From (9.278) we infer:

$$\begin{aligned} F_1''(t) &= 2\left(\frac{2}{p}-1\right)\left(\int |1+tg|^p \, d\mu\right)^{\frac{2}{p}-2}\left(\int |1+tg|^{p-2}\left(\Re g + t\,|g|^2\right) d\mu\right)^2 \\ &\quad + 2\left(\int |1+tg|^p \, d\mu\right)^{\frac{2}{p}-1} \int |1+tg|^{p-2} |g|^2 \, d\mu \\ &\quad + 2(p-2)\left(\int |1+tg|^p \, d\mu\right)^{\frac{2}{p}-1} \int |1+tg|^{p-4}\left(\Re g + t\,|g|^2\right)^2 d\mu \\ &\quad - 2(p-1)\left(\int |g|^p \, d\mu\right)^{\frac{2}{p}} \\ &\leqslant 2\left(\frac{2}{p}-1\right)\left(\int |1+tg|^p \, d\mu\right)^{\frac{2}{p}-2}\left(\int |1+tg|^{p-2}\left(\Re g + t\,|g|^2\right) d\mu\right)^2 \\ &\quad + 2(p-1)\left(\int |1+tg|^p \, d\mu\right)^{\frac{2}{p}-1} \int |1+tg|^{p-2} |g|^2 \, d\mu \\ &\quad - 2(p-1)\left(\int |g|^p \, d\mu\right)^{\frac{2}{p}}. \end{aligned} \tag{9.279}$$

In (9.279) we apply Hölder's inequality to obtain:

$$\int |1+tg|^{p-2} |g|^2 \, d\mu \leqslant \left(\int |1+tg|^p\right)^{1-\frac{2}{p}} \left(\int |g|^p \, d\mu\right)^{\frac{2}{p}}. \tag{9.280}$$

As conjugate exponents we used $\frac{p}{p-2}$ and $\frac{p}{2}$. From (9.280) and (9.279) we then infer $F_1''(t) \leqslant 0$. Since $F_1'(0) = 0$ equality (9.277) with $k = 1$ implies $F_1(t) \leqslant F_1(0) = 1$.

Next we calculate the first and second derivative of $t \mapsto F_2(t)$:

$$\begin{aligned} F_2'(t) &= 2\int \left(\Re g + t\,|g|^2\right) \log\left(1 + 2t\Re g + t^2\,|g|^2\right) d\mu \\ &\quad + 2\int \left(\Re g + t\,|g|^2\right) d\mu - 2t\log\left(1+t^2\right) - 2t - 2t\int |g|^2 \log|g|^2\, d\mu \\ &= 2\int \left(\Re g + t\,|g|^2\right) \log\left(1 + 2t\Re g + t^2\,|g|^2\right) d\mu \\ &\quad - 2t\log\left(1+t^2\right) - 2t\int |g|^2 \log|g|^2\, d\mu. \end{aligned} \tag{9.281}$$

Its second derivative is given by

$$\begin{aligned} F_2''(t) &= 2\int |g|^2 \log\frac{1 + 2t\Re g + t^2\,|g|^2}{|g|^2}\, d\mu + 4\int \frac{\left(\Re g + t\,|g|^2\right)^2}{|1+tg|^2} d\mu \\ &\quad - 2\log\left(1+t^2\right) - \frac{4t^2}{1+t^2} \\ &\leqslant 2\int |g|^2 \log\frac{1 + 2t\Re g + t^2\,|g|^2}{|g|^2}\, d\mu + 4\int |g|^2\, d\mu \\ &\quad - 2\log\left(1+t^2\right) - \frac{4t^2}{1+t^2} \\ &= 2\int |g|^2 \log\frac{1 + 2t\Re g + t^2\,|g|^2}{|g|^2}\, d\mu + \frac{4}{1+t^2} - 2\log\left(1+t^2\right). \end{aligned} \tag{9.282}$$

Since the function $x \mapsto \log x$, $x > 0$, is concave, and the measure $B \mapsto \int_B |g|^2\, d\mu$ is a probability measure, Jensen inequality implies $\int |g|^2 \log h\, d\mu \leqslant \log\left(\int |g|^2\, h\, d\mu\right)$. Applying this inequality to $h = |1+tg|^2 = 1 + 2t\Re g + t^2\,|g|^2$ in (9.282) shows

$$F_2''(t) \leqslant 2\log\int \left(1 + 2t\Re g + t^2\,|g|^2\right) d\mu + \frac{4}{1+t^2} - 2\log\left(1+t^2\right) = \frac{4}{1+t^2}. \tag{9.283}$$

Since $F_2(0) = 0 = F_2'(0)$ it follows from the representation in (9.277) that

$$F_2(t) \leqslant \int_0^t (t-s)\frac{4}{1+s^2} ds \leqslant 2t^2. \tag{9.284}$$

From (9.283) the inequality in (9.274) follows. This concludes the proof of Proposition 9.19. □

9.5 Notes

The result in Theorem 9.1 is taken from [Chen and Wang (1997)] Theorem 4.13. In [Chen and Wang (1997)] the authors wonder whether the condition that there exists a constant $\overline{a} > 0$ such that $\langle a(x)\xi, \xi\rangle \leqslant \overline{a}\,|\xi|^2$ for all x, $\xi \in \mathbb{R}^d$ is really necessary to arrive at a Poincaré inequality. This problem is not solved in Theorem 9.19. However, inequality (9.226) gives a condition in terms of the iterated squared gradient operator Γ_2 and Γ_1 which guarantees a pointwise Poincaré type inequality: see inequality (9.252) in Theorem 9.19.

As mentioned earlier in [Bakry (1994)] and [Bakry (2006)] Bakry gives much more information on (iterates) of squared gradient operators. The squared gradient operator was introduced by Roth in [Roth (1976)] as a tool to study Markov processes. For that matter this is still an important tool: see e.g. Carlen and Stroock [Carlen and Stroock (1986)], Qian [Qian (1998)], Aida [Aida (1998)], Mazet [Mazet (2002)], Barlow, Bass and Kumagai [Barlow *et al.* (2005)], and Wang [Wang (2005)]. Of course the main inspirators for promoting and studying the subject of (iterated) squared gradient operators were and still are Émery and Bakry: see e.g. [Bakry (1985a,b, 1991, 1994, 2006); Bakry and Émery (1985)]. For a link with isoperimetric inequalities the reader is referred to [Chavel and Feldman (1991)], and to [Chavel (2001, 2005)]. More information about Sobolev inequalities and log-Sobolev inequalities can be found in [Carlen *et al.* (1987)], and [Davies (1990)] where a connection with heat kernel diagonal bounds is established. Another relevant paper is [Varopoulos (1985)], and the book by Varopoulos et al [Varopoulos *et al.* (1992)]. The latter book contains a wealth of information related to Sobolev inequalities, Poincaré inequalities, isoperimetric inequalities, and Nash inequalities, and their interrelations.

In the abstract of [Ledoux (1992)] Ledoux writes "In the line of investigation of the works by D. Bakry and M. Emery ([Bakry and Émery (1985)]) and O. S. Rothaus ([Rothaus (1981a, 1986)]) we study an integral inequality behind the "Γ_2" criterion of D. Bakry and M. Emery (see previous reference) and its applications to hypercontractivity of diffusion semigroups. With, in particular, a short proof of the hypercontractivity property of the Ornstein-Uhlenbeck semigroup, our exposition unifies in a simple way several previous results, interpolating smoothly from the spectral gap inequalities to logarithmic Sobolev inequalities and even true Sobolev inequalities. We examine simultaneously the extremal functions for hypercontractivity and logarithmic Sobolev inequalities of the Ornstein-Uhlenbeck semigroup and heat semigroup on spheres."

It seems that these phrases are still in place. In fact the techniques of (iterated) squared gradient operators can also be applied in the infinite-dimensional setting: see e.g. [Wang (2005)].

Examples and other results about invariant measures in the infinite-dimensional context can be found in the books by Da Prato and Zabczyk [Da Prato and Zabczyk (1996)], and by Cerrai [Cerrai (2001)], and papers by Seidler [Seidler (1997)], Eckmann and Hairer [Eckmann and Hairer (2001)], Goldys and van Neerven [Goldys and van Neerven (2003)], Goldys and Maslowski [Goldys and Maslowski (2006b)], and Es-Sarhir and Stannat [Es-Sarhir and Stannat (2007)]. In our abstract setting we followed for a great part the paper by Seidler [Seidler (1997)]. In the following Chapter 10 we also employ techniques from Markov chain theory as exhibited by Meyn and Tweedie [Meyn and Tweedie (1993b)]. Of course the general Chacon-Ornstein theorem goes back to Chacon and Ornstein [Chacon and Ornstein (1960)]. Other relevant literature can be found in [Petersen (1989)], [Krengel (1985)], [Foguel (1980)], and [Neveu (1979)]. It is also mentioned that Azema et al [Azéma *et al.* (1967)] made the Chacon-Ornstein theorem corresponding to continuous time Markov processes available to the mathematical public.

A novelty in the present chapter is the fact that the almost separation property of the generator of the Markov process in (9.14) together with a topological irreducibility condition implies that the process admits a compact recurrent subset. This observation follows from a combination of Propositions 9.1, 9.2, and 9.4: in particular see Corollary 9.2 and Theorem 9.4.

This page is intentionally left blank

Chapter 10

Invariant measure

In this final chapter we prove the existence and uniqueness of invariant measures for recurrent time-homogeneous Markov processes. Our uniqueness result relies on Orey's theorem for Markov chains: see Theorem 10.2. The proof of Orey's convergence theorem is based on renewal theory: see Lemma 10.14, and the bivariate linked forward recurrence time chain employed in its proof. Orey's theorem is combined with the presence of a compact recurrent subset to obtain, up to a multiplicative constant, a unique invariant measure; see Theorem 10.12. The equalities in (10.205) and (10.206) play a central role. Proposition 10.8 contains the technical relevant details. In particular the identity in (10.272) is a crucial equality. Under certain conditions this invariant measure is finite: see Corollary 10.5. In Theorem 10.9 we see that for certain conservative strong Feller processes the notions of recurrent and Harris recurrent coincide.

10.1 Markov Chains: Invariant measure

Some of what follows is taken from [Chib (2004)] and [Meyn and Tweedie (1993b)]. One of the motivations to study time-homogeneous Markov chains is the fact that Monte Carlo methods sample a given multivariate distribution π by constructing a suitable Markov chain with the property that its limiting, invariant distribution, is the target distribution π. In most problems of interest, the distribution π is absolutely continuous and, as a result, the theory of MCMC (Markov Chain Monte Carlo) methods is based on that of Markov chains on continuous state spaces outlined, for example, in [Meyn and Tweedie (1993b)] and [Nummelin (1984)]. Reference [Tierney (1994)] is the fundamental reference for drawing the connections between this elaborate Markov chain theory and MCMC methods. Basically, the goal of the analysis is to specify conditions under which the constructed Markov chain converges to the invariant distribution, and conditions under which sample path averages based on the output of the Markov chain satisfy a law of large numbers and a central limit theorem.

10.1.1 *Some definitions and results*

A Markov chain is a sequence of random variables (or state variables) $X = \{X(i) : i \in \mathbb{N}\}$ together with a transition probability function $(x, B) \mapsto P(x, B)$, $x \in E$, $B \in \mathcal{E}$. The evolution of the Markov chain on a space E is governed by the transition kernel

$$\begin{aligned} P(x,B) &= \mathbb{P}\left[X(i+1) \in B \,\middle|\, X(i) = x, \mathcal{F}_{i-1}\right] \\ &= \mathbb{P}\left[X(i+1) \in B \,\middle|\, X(i) = x\right], \quad (x,B) \in E \times \mathcal{E}, \end{aligned} \tag{10.1}$$

where the second line embodies the time-homogeneous Markov property that the distribution of each succeeding state in the sequence, given the current and the past states, depends only on the current state. Note that $\mathcal{F}_{i-1}$ represents the σ-field generated by the variables $\{X(j) : 0 \leqslant j \leqslant i-1\}$. In fact, a complete description of a time-homogeneous Markov chain is given by:

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(i), i \in \mathbb{N}), (\vartheta_i, i \in \mathbb{N}), (E, \mathcal{E})\} \tag{10.2}$$

where

$$\begin{aligned} \mathbb{P}_x\left[X(1) \in B\right] &= \mathbb{P}\left[X(1) \in B \,\middle|\, X(0) = x\right] = \mathbb{P}\left[X(i+1) \in B \,\middle|\, X(i) = x\right] \\ &= P(x,B) = P(1,x,B). \end{aligned} \tag{10.3}$$

The operators ϑ_i, $i \in \mathbb{N}$, are time shift operators: $\vartheta_i \circ \vartheta_j = \vartheta_{i+j}$, i, $j \in \mathbb{N}$. Moreover, $X(i) \circ \vartheta_j = X(i+j)$ $\mathbb{P}_x$-almost surely for all $x \in E$ and all i, $j \in \mathbb{N}$. A convenient way to express the Markov property goes as follows:

$$\mathbb{P}_x\left[X(i+1) \in B \,\middle|\, \mathcal{F}_i\right] = \mathbb{P}_{X(i)}\left[X(1) \in B\right], \quad (x,B) \in E \times \mathcal{E}, \; i \in \mathbb{N}.$$

If in (9.14) we confine the time $[0, \infty)$ to the discrete time $\mathbb{N}$, then we get a Markov chain with a not necessarily discrete state space. The Markov chain obtained from (9.14) is called a skeleton of the time-homogeneous Markov process in continuous time. The transition kernel is thus the distribution of $X(i+1)$ given that $X(i) = x$. The nth step ahead transition kernel is given by

$$P(n,x,B) = P^n(x,B) = \int_E P(x,dy)\, P^{(n-1)}(y,B), \tag{10.4}$$

where

$$B \mapsto P^{(1)}(x, B) = P(x, B) = P(1, x, B), \quad B \in \mathcal{E}, \tag{10.5}$$

is a probability measure on $\mathcal{E}$, the Borel field of the state space E. In fact the Markov property of the time-discrete process in (10.2) is equivalent to the following Chapman-Kolmogorov equation:

$$P(n+m, x, B) = P^{n+m}(x, B) = \int_E P^n(x, dy)\, P^m(y, B), \;\; n, m \in \mathbb{N}, \; x \in E. \tag{10.6}$$

Instead of the skeleton $\{X(i) : i \in \mathbb{N}\}$ we could have taken a skeleton of the form

$$\{X(\delta i) : i \in \mathbb{N}\}, \quad \delta > 0. \tag{10.7}$$

Again we get a Markov chain, and the results of Meyn and Tweedie can be used. However, note that hitting times phrased in terms of a skeleton in general are larger than the original hitting times. On the other hand, in our setup the paths of the Markov process are continuous from the right, and so in principle our Markov process can be approximated by skeletons of the form (10.7).

The goal is to find conditions under which the nth iterate of the transition kernel converges to the invariant measure or distribution π as $n \to \infty$. The invariant distribution is one that satisfies

$$\pi(B) = \int_E P(x, B) d\pi(x). \tag{10.8}$$

The invariance condition states that if $X(i)$ is distributed according to π, then all subsequent elements of the chain are also distributed as π. Markov chain samplers are invariant by construction and therefore the existence of the invariant distribution does not have to be checked.

A Markov chain is reversible, or satisfies the detailed balance condition, if there exists a reference measure m on $\mathcal{E}$ such that the transition function $P(x, B)$ can be written as $P(x, B) = \int_E p(x, y) dm(y)$, where the integral kernel $p(x, y)$ satisfies

$$f(x)p(x, y) = f(y)p(y, x), \tag{10.9}$$

for a Borel measurable function $f(\cdot)$ which is the Radon-Nikodym derivative of some Borel measure $B \mapsto \pi(B)$, $B \in \mathcal{E}$. If this condition holds, it can be shown that π is an invariant measure: see e.g. [Tierney (1994)]. To verify this we evaluate the right hand side of (10.8):

$$\int P(x, B)\, d\pi(x) = \int \left\{ \int_B p(x, y)\, dm(y) \right\} f(x)\, dm(x)$$

$$
\begin{aligned}
&= \int_B \left\{ \int f(x)p(x,y)dm(x) \right\} dm(y) \\
&= \int_B \left\{ \int p(y,x)f(y)dm(x) \right\} dm(y) \\
&= \int_B f(y)dm(y) = \pi(B).
\end{aligned}
$$

A minimal requirement on the Markov chain for it to satisfy a law of large numbers is the requirement of π-irreducibility. This means that the chain is able to visit all sets with strictly positive probability under π from any starting point in E. Formally, a Markov chain is said to be π-irreducible if for every $x \in E$,

$$\pi(A) > 0 \Rightarrow P\left[X(i) \in A \,\middle|\, X(0) = x\right] > 0, \quad \text{for some } i \geqslant 1. \tag{10.10}$$

The property in (10.10) can also be phrased in terms of the hitting time of A: $\tau_A^1 = \min\{m \geqslant 1 : X(m) \in A\}$. If $X(m) \notin A$ for all $m \in \mathbb{N}$, $m \geqslant 1$, then we put $\tau_A^1 = \infty$. An equivalent way to write (10.10) goes as follows: if $A \in \mathcal{A} \in \mathcal{E}$ is such that $\pi(A) > 0$, then $\mathbb{P}_x\left[\tau_A^1 < \infty\right] > 0$ for all $x \in E$. If the space E is connected and the function $p(x,y)$ is positive and continuous, then the Markov chain with transition probability function given by $P(x,B) = \int_B p(x,y)dm(y)$ and the invariant probability measure π is π-irreducible.

In our case another important property of the Markov chain is its aperiodicity, which ensures that the chain does not cycle through a finite number of sets. For topics related to 10.1 see Definitions 10.6 and 10.7.

Definition 10.1. A Markov chain is aperiodic if there exists no partition of $E = (D_0, D_1, \dots, D_{p-1})$ for some $p \geqslant 2$ such that for all $i \in \mathbb{N}$

$$P\left[X(i) \in D_{i \bmod (p)} | X(0) \in D_0\right] = \int_{D_0} \mathbb{P}_x\left[X(i) \in D_{i \bmod (p)}\right] d\mu_0(x) = 1, \tag{10.11}$$

for some initial probability distribution μ_0.

If the probability μ_0 and the partition $(D_0, \dots, D_{p-1})$ did have the property spelled out in (10.11), then there exists a state $x_0 \in D_0$ such that

$$P\left(i, x_0, D_{i \bmod (p)}\right) = \mathbb{P}_{x_0}\left[X(i) \in D_{i \bmod (p)}\right] = 1, \quad \text{for all } i \in \mathbb{N}. \tag{10.12}$$

It follows that not all probability measures $B \mapsto P(i, x_0, B)$, $i \in \mathbb{N}$, $i \geqslant 1$, have the same null-sets. So we have the following result.

Proposition 10.1. *Let the time-homogeneous Markov chain in (10.2) have a transition probability function $P(i,x,B)$, $i \in \mathbb{N}$, $x \in E$, $B \in \mathcal{E}$, where*

$P(x,B) = P(1,x,B)$, and $P(0,x,B) = \mathbf{1}_B(x)$. Suppose that all probability measures $B \mapsto P(i,x,B)$, $i \geqslant 1$, $i \in \mathbb{N}$, $x \in E$, have the same negligible sets. Then the Markov chain in (10.2) is aperiodic.

These definitions allow us to state the following results from [Tierney (1994)] which form the basis for Markov chain Monte Carlo methods, and other asymptotic results. The first of these results gives conditions under which a strong law of large numbers holds and the second gives conditions under which the probability density of the nth iterate of the Markov chain converges to its unique, invariant density.

Theorem 10.1. *Suppose $\{X(i), \mathbb{P}_x\}_{x\in E}$ is a π-irreducible time-homogeneous Markov chain with transition kernel $P(x,B) = P(1,x,B)$ and invariant probability distribution π. Then π is the unique invariant distribution of $P(x,B)$ and for all π-integrable real-valued functions h,*

$$\frac{1}{n}\sum_{i=1}^{n} h(X(i)) \to \int h(x)d\pi(x) \quad \text{as } n \to \infty,\ \mathbb{P}_x\text{-almost surely.} \tag{10.13}$$

If the invariant measure is σ-finite and not finite, then the limit in (10.13) is zero. That is why irreducible Markov chains with a (unique) σ-finite invariant measure, which is not finite, are called Markov chains which are null-recurrent. For ergodicity results in null recurrent Markov chains, like the theorem of Chacon-Ornstein for quotients of time averages as in (10.13), the reader is referred to [Krengel (1985)]: see Theorem 9.9. Recurrent Markov chains with a finite invariant measure are called positive recurrent. There is a close relationship between expectations of (first) return times and invariant measures. In the discrete state space setting we have the following. Put $T_y = \inf\{m \geqslant 1 : X(m) = y\}$, $y \in E$, and write $\mu_{x,y} = \mathbb{E}_x[T_y]$. Then the following equality holds:

$$\pi(y) = \lim_{n\to\infty} P^n(x,\{y\}) = \frac{1}{\mu_{y,y}}. \tag{10.14}$$

The result in (10.14) is called Kac's theorem: see Theorem 10.2.2 in [Meyn and Tweedie (1993b)]. For more details the reader is referred to the literature: [Norris (1998)] and [Karlin and Taylor (1975)]. Some older work can be found in [Orey (1964)], [Kingman and Orey (1964)], and [Jamison *et al.* (1965)]. The following Theorem of Orey, or Orey's convergence theorem can be found in Meyn and Tweedie [Meyn and Tweedie (1993b)] theorem 13.3.3 and 18.1.2. For the claim in (10.15) the positivity of the recurrent Markov chain is not required. It suffices to have a σ-finite invariant measure, which is guaranteed by a result due to Foguel [Foguel (1966)] for irreducible chains with a recurrent compact subset: see Theorem 2.2 in [Seidler (1997)]. The existence of a σ-finite Borel measure is also proved in Chapter 10 of the new version of the book [Meyn and Tweedie (1993b)]. The assertion as written is proved in Duflo and Revuz [Duflo and Revuz (1969)], who use a method developed by Blackwell and Freedman [Blackwell and Freedman (1964)], who in turn rely on a result by Orey [Orey (1959)] which states a result like (10.15) for point measures $\mu_i = \delta_{x_i}$, $i = 1, 2$. The following theorem was used in the proof of Proposition 9.8. In [Kaspi and Mandelbaum (1994)] Theorem 1 and Lemma 1 the authors establish a close relationship between recurrence and Harris recurrence. A similar result for the fine topology was found by Azema et al in [Azéma *et al.* (1965/1966)] Proposition IV 4.

Theorem 10.2. *Suppose that $\{X(n), \mathbb{P}_x\}_{x\in E}$ is an irreducible time-homogeneous aperiodic Markov chain with a transition kernel, denoted by $P(x, B) = P(1, x, B)$, which is Harris recurrent. Then for all probability measures μ_1 and μ_2 on $\mathcal{E}$*

$$\lim_{n\to\infty} \iint \operatorname{Var}\left(P^n(x,\cdot) - P^n(y,\cdot)\right) d\mu_1(x)\, d\mu_2(y) = 0, \tag{10.15}$$

where Var *denotes the total variation norm. If the Markov chain is positive Harris recurrent, then for μ_2 the invariant probability measure π may be chosen. This existence follows from positive recurrence. Then the following equality holds for all probability measures μ_1 on $\mathcal{E}$:*

$$\lim_{n\to\infty} \operatorname{Var}\left(\int P^n(x,\cdot) d\mu_1(x) - \pi(\cdot)\right) = 0. \tag{10.16}$$

Let $B \in \mathcal{E}$. The proof of Theorem 10.2 is based on among other things the decomposition of the event $\{X(n) \in B\}$ over the times of the first and the last entrance time, or entry time to A prior to the time n:

$$\begin{aligned}
&\mathbb{P}_x\left[X(n) \in B\right] = \mathbb{P}_x\left[X(n) \in B,\, \tau_A^1 \geqslant n\right] \\
&\quad + \sum_{j=1}^{n-1} \sum_{k=1}^{j} \mathbb{E}_x\left[\mathbb{E}_{X(k)}\left[\mathbb{P}_{X(j-k)}\left[X(n-j) \in B,\, \tau_A^1 \geqslant n-j\right],\, X(j-k) \in A\right],\right. \\
&\qquad \left.\tau_A^1 \geqslant k,\, X(k) \in A\right] \\
&= \mathbb{P}_x\left[X(n) \in B,\, \tau_A^1 \geqslant n\right] \\
&\quad + \sum_{j=1}^{n-1} \sum_{k=1}^{j} \mathbb{E}_x\left[\mathbb{E}_{X(k)}\left[\mathbb{P}_{X(j-k)}\left[X(n-j) \in B,\, \tau_A^1 \geqslant n-j\right],\, X(j-k) \in A\right],\right. \\
&\qquad \left.\tau_A^1 = k\right],
\end{aligned} \tag{10.17}$$

where in the final step of (10.17) we used the following equality of events: $\{\tau_A^1 \geqslant k,\, X(k) \in A\} = \{\tau_A^1 = k\}$, $k \in \mathbb{N}$, $k \geqslant 1$. The formula in (10.17) can be found in [Meyn and Tweedie (1993b)] formula (13.39). Its proof is an easy consequence of the Markov property. The entrance time $\tau_A^1 = \tau_A^{1,1}$ is defined in (10.25) of Theorem 10.4: $\tau_A^1 = \inf\{n \geqslant 1 : X(n) \in A\}$. In terms of functions the equality in (10.17) reads:

$$\begin{aligned}\mathbb{E}_x\left[f\left(X(n)\right)\right] &= \mathbb{E}_x\left[f\left(X(n)\right),\, \tau_A^1 \geqslant n\right] \\ &+ \sum_{j=1}^{n-1}\sum_{k=1}^{j} \mathbb{E}_x\left[\mathbb{E}_{X(k)}\left[\mathbb{E}_{X(j-k)}\left[f\left(X(n-j)\right),\, \tau_A^1 \geqslant n-j\right],\, X(j-k) \in A\right],\right. \\ &\quad \left.\tau_A^1 = k\right], \quad f \in C_b(E). \end{aligned} \tag{10.18}$$

In §10.3 we provide a proof of Orey's convergence theorem.

Definition 10.2. The Markov chain $(X(n) : n \in \mathbb{N})$ in (10.2) is called positive Harris recurrent if there exists an invariant probability measure on $\mathcal{E}$ relative to X, and if $(x, B) \in E \times \mathcal{E}$ satisfies $P(x, B) > 0$, then

$$\mathbb{P}_x\left[\sum_{n=1}^{\infty} \mathbf{1}_B\left(X(n)\right) = \infty\right] = 1.$$

A further strengthening of the conditions is required to obtain a central limit theorem for sample-path averages. A key requirement is that of an ergodic chain, i.e., a chain that is irreducible, aperiodic and positive Harris-recurrent: for a definition of the latter, see [Tierney (1994)] and [Meyn and Tweedie (1993b)]. In addition, one needs the notion of geometric ergodicity. An ergodic Markov chain with invariant distribution π is geometrically ergodic if there exists a non-negative real-valued Borel function $x \mapsto C(x)$ and a positive constant $r < 1$ such that $\operatorname{Var}\left(P^n(x,\cdot) - \pi(\cdot)\right) \leqslant C(x) r^n$ for all $n \in \mathbb{N}$, and such that $\int C(x)\, d\pi(x) < \infty$. The authors of [Chan and Ledolter (1995)] show that if the Markov chain is ergodic, has invariant probability distribution π, and is geometrically ergodic, then for all $L^2(E,\pi)$-integrable measurable real-valued functions h, and any initial distribution, the distribution of $\sqrt{n}\left(\widehat{h}_n - \int h(x)\,dx\right)$ converges weakly to a normal distribution with mean zero and variance $\sigma_h^2 \geqslant 0$ as $n \to \infty$. Here $\widehat{h}_n = \dfrac{1}{n}\sum_{i=1}^{n} h\left(X(i)\right)$, and $\sigma_h^2 = \operatorname{Var} h\left(X(0)\right) + 2\sum_{k=1}^{\infty} \operatorname{Cov}\left[\{h\left(X(0)\right), h\left(X(k)\right)\}\right]$. The following theorem discusses the problem of the existence of an invariant measure. It is taken from [Meyn and Tweedie (1993b)] Theorem 10.0.1. It is supposed that all measures $B \mapsto P(1,x,B)$, $B \in \mathcal{E}$, $x \in E$, have the same null-sets. Put $\mathcal{E}^+ = \{A \in \mathcal{E} : P(1,x_0,A) > 0\}$. Irreducibility is meant in the sense that $\mathbb{P}_x[\tau_A < \infty] > 0$ for all $x \in E$, and all subsets $A \in \mathcal{E}^+$. In our setting we may assume that irreducibility can be phrased in terms of reachability of any open subset with positive probability from any starting point and in as short time as we please: see Lemma 9.1. It is not clear what the exact analog is of (10.22) below in case we are working with continuous time processes like the one in (9.14). It is quite well possible that in that case Dynkin's formula plays a central role. Let $\{(\Omega,\mathcal{F},\mathbb{P}_x),(X(t),t\geqslant 0),(\vartheta_t,t\geqslant 0),(E,\mathcal{E})\}$ be a time-homogeneous strong Markov process, and let A be a Borel subset of E with hitting time τ_A: $\tau_A = \inf\{s > 0 : X(s) \in A\}$. For $\lambda > 0$ we have Dynkin's formula:

$$\begin{aligned}\int_0^\infty e^{-\lambda s}\mathbb{E}_x[f(X(s))]\,ds - \int_0^\infty e^{-\lambda s}\mathbb{E}_x[f(X(s)),\,\tau_A > s]\,ds \\ = \mathbb{E}_x\left[e^{-\lambda\tau_A}\mathbb{E}_{X(\tau_A)}\left[\int_0^\infty e^{-\lambda s} f(X(s))\,ds\right]\right].\end{aligned} \tag{10.19}$$

If we use the resolvent notation:

$$R(\lambda)f(x) = \int_0^\infty e^{-\lambda s}e^{sL}f(x)ds, \quad \text{and} \quad R_A(\lambda)f(x) = \int_0^\infty e^{-\lambda s}e^{sL_A}f(x)ds, \tag{10.20}$$

then the equality in (10.19) can be rewritten as:

$$R(\lambda)f(x) - R_A(\lambda)f(x) = \mathbb{E}_x\left[e^{-\lambda\tau_A}R(\lambda)f(X(\tau_A))\right]. \tag{10.21}$$

The semigroup $\left\{e^{sL_A} : s \geqslant 0\right\}$ is defined by

$$e^{sL_A}f(x) = \mathbb{E}_x[f(X(s)) : \tau_A > s],\ f \in L^\infty(E,\mathcal{E}),\ s \geqslant 0,\ x \in E.$$

This semigroup need not be strongly continuous. It lives on $A^c = E\backslash A$. It is quite well possible that equality (10.272) below in Proposition 10.8 is the correct analog of (10.22). The following result appears as Theorem 10.0.1 in [Meyn and Tweedie (1993b)]: see Theorem 10.4.4 and Theorem l0.4.9 l.c. as well. The result refines Theorem 1 in [Harris (1956)].

Theorem 10.3. *Let the time-homogeneous Markov chain X with a Polish state space E be m-irreducible in the sense that for all $A \in \mathcal{E}$ for which $m(A) > 0$ and all $x \in E$ there exists $n \in \mathbb{N}$ such that $\mathbb{P}_x[X(n) \in A] > 0$. Suppose that the process X be recurrent relative to the measure m. Then it admits, up to multiplicative constants, a unique σ-finite invariant measure*

π. Let $A \in \mathcal{E}$ be such that $\mathbb{P}_x\left[\tau_A < \infty\right] = 1$ for π-almost all $x \in E$. The measure π satisfies:

$$\begin{aligned}\pi(B) &= \int_A \mathbb{E}_x\left[\sum_{i=1}^{\tau_A^1} \mathbf{1}_B\left(X(i)\right)\right] d\pi(x) \\ &= \int_A \mathbb{E}_x\left[\sum_{i=0}^{\tau_A^1-1} \mathbf{1}_B\left(X(i)\right)\right] d\pi(x), \quad B \in \mathcal{E}. \end{aligned} \tag{10.22}$$

This measure π is such that $\pi[B] = 0$ if and only if $m(B) = 0$. The invariant measure is finite (rather than merely σ-finite), if there exists a compact subset C such that $\sup_{x \in C} \mathbb{E}_x\left[\tau_C^1\right] < \infty$. Moreover,

$$\pi(E) = \int_A \mathbb{E}_x\left[\tau_A^0\right] d\pi(x) = \int_C \mathbb{E}_x\left[\tau_C^1\right] d\pi(x).$$

In (10.22) τ_A^1 stands for the first hitting time of the Borel subset A:

$$\tau_A^1 = \min\left\{m \geqslant 1 : X(m) \in A\right\} = 1 + \tau_A^0 \circ \vartheta_1$$

where another stopping time τ_A^0 also plays a relevant role:

$$\tau_A^0 = \min\left\{m \geqslant 0 : X(m) \in A,\ m \text{ non-negative integer}\right\}.$$

In [Meyn and Tweedie (1993b)] Meyn and Tweedie discuss "petite" and "small" sets. Theorem 10.3 follows from a combination of the following theorems and propositions in [Meyn and Tweedie (1993b)]: Theorem 10.0.1 (in which ψ-irreducibility and "petite sets" play a crucial role), a rephrasing of assertion (i) of Proposition 5.2.4 in terms of petite sets, which is in fact the same as assertion (i) of Proposition 5.5.4, and assertion (ii) in Theorem 6.2.5 (which states that in a topological Markov chain all compact subsets are "petite"). Meyn and Tweedie use the following terminology. Let $B \mapsto \varphi(B)$ be a finite measure on $\mathcal{E}$. A Markov chain is called φ-irreducible, if every set $A \in \mathcal{E}$ for which $\varphi(A) > 0$ the quantity $\mathbb{P}_x\left[\tau_A^0 < \infty\right] > 0$ for all $x \in E$. In our case we may take $\psi(A) = \varphi(A) = P\left(1, x_0, A\right)$, $A \in \mathcal{E}$. The assumption that all measures of the form $A \mapsto P\left(t_0, x_0, A\right)$ are equivalent, makes the choice of $x_0 \in E$ irrelevant. Let $a := \left(a_k\right)_{k \in \mathbb{N}}$ be a sequence of non-negative real numbers which add up to one. Then we define the function $(x, A) \mapsto K_a(x, A)$, $(x, A) \in E \times \mathcal{E}$, by $K_a(x, A) = \sum_{k=0}^{\infty} a_k P^n(x, A)$. Note that $P(x, A) = P(1, x, A)$. Denote by $P(\mathbb{N})$ the collection of positive sequences which add up to 1. A subset $A \in \mathcal{E}$ is called "petite" if there exists a sequence $a \in P(\mathbb{N})$ and a non-trivial measure

ν_a such that $K_a(x,B) \geqslant \nu_a(B)$ for all $B \in \mathcal{E}$ and all $x \in A$. If we can find a of the form $a_k = \delta_n(k)$, $k \in \mathbb{N}$, and a corresponding non-trivial measure ν_n, for some $n \in \mathbb{N}$, then A is called "small"; this means that $P^n(x,B) = K_{\delta_n}(x,B) \geqslant \nu_n(B)$, $B \in \mathcal{E}$, and ν_n a non-trivial measure. It says that for Markov chains which are ψ-irreducible and aperiodic the collection of "small sets" coincides the collection of "petite sets". The Markov chain X is called topological, or a Markov T-chain, if the space E is a complete metrizable (locally compact) Hausdorff space, and the function $x \mapsto P(x,B) = P(1,x,B)$ is lower semi-continuous for every $B \in \mathcal{E}$. In fact the authors assume that for every $B \in \mathcal{E}$ the function $x \mapsto P(x,B)$ dominates a strictly positive lower semi-continuous function, whenever it itself is strictly positive. Observe that by the Markov property

$$\mathbb{P}_x\left[\tau_A^1 < \infty\right] = \mathbb{P}_x\left[\bigcup_{n=1}^{\infty}\{X(n) \in A\}\right] = \mathbb{E}_x\left[\mathbb{P}_{X(1)}\left[\bigcup_{n=1}^{\infty}\{X(n-1) \in A\}\right]\right],$$

and hence, by the strong Feller property, the function $x \mapsto \mathbb{P}_x\left[\tau_A^1 < \infty\right]$ is in fact continuous. For the notion of strong Feller property see Definitions 2.5 and 2.16. In the results, which we mention above and which will follow, the local compactness does not play a role.

As a corollary to (the proof of) Theorem 10.3 we have the following result.

Corollary 10.1. *Let the notation and assumptions be as in Theorem 10.3. Then the following equality holds for $f \in L^1(E,\pi)$:*

$$\lim_{n\to\infty}\int_{E\setminus A}\mathbb{E}_x\left[f\left(X(n)\right),\, \tau_A^0 \geqslant n\right]\,d\pi = 0. \tag{10.23}$$

A result, corresponding to Theorem 10.3 in the continuous time setting, reads as follows.

Theorem 10.4. *Let $\{(\Omega,\mathcal{F},\mathbb{P}_x),(X(t),t\geqslant 0),(\vartheta_t,t\geqslant 0),(E,\mathcal{E})\}$ be a strong Markov process with right-continuous paths. Fix $h > 0$. Let π be a σ-finite invariant measure for this Markov process. Then the following equality holds for all $f \in L^1(E,\pi)$, and for all Borel subsets A with the property that $\pi(A) < \infty$ and $\mathbb{P}_x\left[\tau_A^{0,h} < \infty\right] = 1$ for π-almost all $x \in E$:*

$$\int_A \mathbb{E}_x\left[\sum_{k=1}^{\tau_A^{1,h}/h} f\left(X\left(kh\right)\right)\right]d\pi(x) = \int f(x)d\pi(x). \tag{10.24}$$

In (10.24) the stopping times $\tau_A^{1,h}$, $h > 0$, *are defined by*

$$\tau_A^{1,h} = \inf\{\ell h : \ell \in \mathbb{N},\, \ell \geqslant 1,\, X(\ell h) \in A\} = h + \tau_A^{0,h} \circ \vartheta_h \tag{10.25}$$

where $\tau_A^{0,h} = \inf\{\ell h : \ell \in \mathbb{N},\, \ell \geqslant 0,\, X(\ell h) \in A\}$. *If there exists* $A \in \mathcal{E}$ *with the property that* $\int_A \mathbb{E}_x\left[\tau_A^{1,h}\right] d\pi(x) < \infty$ *for some* $h > 0$, *then the invariant measure* π *is finite, and* $h \times \pi(E) = \int_A \mathbb{E}_x\left[\tau_A^{1,h}\right] d\pi(x)$.

If $h = 1$ we write τ_A^1 instead of $\tau_A^{1,1}$: see formula (10.17) above. The proof of Theorem 10.4 is completely analogous to that of Theorem 10.3. Instead of the operator T, given by $Tf(x) = \mathbb{E}_x\left[f(X(1))\right] = \int f(y)P(x,dy)$, we now introduce the operators T_h, $h > 0$, $T_h f(x) = \mathbb{E}_x\left[f\left(X(h)\right)\right] = \int f(y)P(h,x,dy)$, where $P(t,x,B)$ is the probability transition function. We also need the operator $T_{A,h}$ defined by

$$T_{A,h}f(x) = \mathbb{E}_x\left[f\left(X(h)\right), \tau_A^{0,h} \geqslant h\right] = \int f(y)P_A(h,x,dy),$$

where $P_A(h,x,B) = \mathbb{P}_x\left[X(h) \in B,\, \tau_A^{0,h} \geqslant h\right]$. Again the proof yields the following corollary.

Corollary 10.2. *Let the notation and assumptions be as in Theorem 10.4. Then the following equality holds for* $f \in L^1(E,\pi)$:

$$\lim_{n\to\infty} \int_{E\setminus A} \mathbb{E}_x\left[f\left(X(nh)\right), \tau_A^{0,h} \geqslant nh\right] d\pi(x) = 0. \tag{10.26}$$

Proof. [Proof of Theorem 10.3.] Let $A \in \mathcal{E}$ be as in Theorem 10.3. We introduce two operators T and T_A, defined by respectively

$$Tf(x) = \mathbb{E}_x\left[f\left(X(1)\right)\right], \text{ and } T_A f(x) = \mathbb{E}_x\left[f\left(X(1)\right), \tau_A^0 \geqslant 1\right],\ \ f \in C_b(E). \tag{10.27}$$

Notice that $T_A f = Tf - \mathbf{1}_A Tf = \mathbf{1}_{E\setminus A} Tf$, so that $T_A f = 0$ on A. Then by induction with respect to n we see

$$\sum_{k=1}^n \mathbf{1}_A T T_A^{k-1} f + \mathbf{1}_{E\setminus A} T_A^n f = f + \sum_{k=1}^n (T - I)\, T_A^{k-1} f, \quad f \in C_b(E). \tag{10.28}$$

Hence, since π is an invariant measure the equality in (10.28) implies

$$\sum_{k=1}^n \int_A T T_A^{k-1} f(x)\, d\pi(x) + \int_{E\setminus A} T_A^n f(x)\, d\pi(x) = \int_E f(x)\, d\pi(x) \tag{10.29}$$

for $f \in L^1(E,\pi)$. Let $f \in L^\infty(E,\mathcal{E}) \bigcap L^1(E,\pi)$, $f \geqslant 0$. By the assumption that $\mathbb{P}_x\left[\tau_A^0 < \infty\right] = 1$, for π-almost all $x \in E$, we have $f = \lim_{n\to\infty}(f - T_A^n f)$, π-almost everywhere, and hence we obtain

$$\int_E f d\pi = \int_E \lim_{n\to\infty}(f - T_A^n f)\, d\pi = \int_E \liminf_{n\to\infty}(f - T_A^n f)\, d\pi$$

(Fatou's lemma)

$$\leqslant \liminf_{n\to\infty}\int_E (I - T_A^n) f\, d\pi = \liminf_{n\to\infty}\int_E (I - T_A)\sum_{k=1}^n T_A^{k-1} f\, d\pi$$

(employ the identity $T_A = T - \mathbf{1}_A T$)

$$= \liminf_{n\to\infty}\int_E (I - T + \mathbf{1}_A T)\sum_{k=1}^n T_A^{k-1} f\, d\pi$$

(the measure π is T-invariant)

$$= \liminf_{n\to\infty}\int_A T\sum_{k=1}^n T_A^{k-1} f\, d\pi = \sum_{k=1}^\infty \int_A T T_A^{k-1} f\, d\pi$$

$$= \int_A \mathbb{E}_x\left[\sum_{k=1}^{\tau_A^1} f(X(k))\right] d\pi(x). \tag{10.30}$$

The inequality in (10.30) shows that

$$\int_E f\, d\pi \leqslant \int_A \mathbb{E}_x\left[\sum_{k=1}^{\tau_A^1} f(X(k))\right] d\pi(x). \tag{10.31}$$

Of course, in (10.31) we assumed $\mathbb{P}_x\left[\tau_A^0 < \infty\right] = 1$, $x \in E$. On the other hand the equality in (10.28) yields:

$$\sum_{k=1}^n \int_A T T_A^{k-1} f\, d\pi$$

$$\leqslant \sum_{k=1}^n \int_A T T_A^{k-1} f\, d\pi + \int_{E\backslash A} T_A^n f\, d\pi = \int_E f\, d\pi + \int_E \sum_{k=1}^n (T - I) T_A^{k-1} f\, d\pi$$

$$= \int_E f\, d\pi. \tag{10.32}$$

From (10.32) we get, by letting $n \to \infty$ and using the Markov property several times

$$\int_A \mathbb{E}_x\left[\sum_{k=1}^{\tau_A^1} f(X(k))\right] d\pi(x) \leqslant \int_E f\, d\pi. \tag{10.33}$$

Combining (10.33) and (10.31) shows the equality:

$$\int_A \mathbb{E}_x \left[\sum_{k=1}^{\tau_A^1} f(X(k)) \right] d\pi(x) = \int_E f\, d\pi. \tag{10.34}$$

The equality in (10.34) completes the proof of Theorem 10.3. □

Remark 10.1. If in Theorem 10.3 we only assume π to be sub-invariant in the sense that $\int_E Tf(x)d\pi(x) \leqslant \int_E f(x)d\pi(x)$, $f \in L^1(E,\pi)$, $f \geqslant 0$, then for such functions we have

$$\sum_{k=1}^{n} \int_A TT_A^{k-1} f(x)\, d\pi + \int_{E\setminus A} T_A^n f(x)\, d\pi(x) \leqslant \int_E f(x)\, d\pi(x). \tag{10.35}$$

Proof. [Proof of Corollary 10.1.] The equality in (10.29) can be rewritten as follows:

$$\int_A \mathbb{E}_x \left[\sum_{k=1}^{\tau_A^1\wedge n} f(X(k)) \right] d\pi(x) + \int_{E\setminus A} \mathbb{E}_x \left[f(X(n)),\, \tau_A^0 \geqslant n \right] d\pi(x)$$
$$= \int_E f(x) d\pi(x). \tag{10.36}$$

The equality in (10.34) together with (10.36) yields the result in Corollary 10.1. To establish we need once more the fact that $\mathbb{P}_x\left[\tau_A^0 < \infty\right] = 1$ for π-almost all $x \in E$.

This completes the proof of Corollary 10.1. □

In the following corollary we give a result similar to the one in Theorem 10.3, but here we do not necessarily assume that $\mathbb{P}_x\left[\tau_A^0 < \infty\right] = 1$ for π-almost all $x \in E$.

Corollary 10.3. *Define the measures π_1 and π_∞ by the equalities:*

$$\int_E f(x)\, d\pi_1(x) = \inf_{\ell\in\mathbb{N}} \sup_{n\in\mathbb{N}} \int_A \mathbb{E}_x \left[\sum_{k=1}^{\tau_A^1\wedge n} T^\ell f(X(k)) \right] d\pi(x)$$
$$= \inf_{\ell\in\mathbb{N}} \sup_{n\in\mathbb{N}} \int_A \mathbb{E}_x \left[\sum_{k=1}^{\tau_A^1\wedge n} f(X(k+\ell)) \right] d\pi(x), \quad \textit{and} \tag{10.37}$$

$$\int_E f(x)\, d\pi_\infty(x) = \sup_{\ell\in\mathbb{N}} \inf_{n\in\mathbb{N}} \int_{E\setminus A} \mathbb{E}_x \left[T^\ell f(X(n)),\, \tau_A^0 \geqslant n \right] d\pi(x)$$
$$= \sup_{\ell\in\mathbb{N}} \inf_{n\in\mathbb{N}} \int_{E\setminus A} \mathbb{E}_x \left[f(X(n+\ell)),\, \tau_A^0 \geqslant n \right] d\pi(x), \tag{10.38}$$

where the function $f \geqslant 0$ *belongs to* $L^1(E,\pi)$*. Then the measures* π_1 *and* π_∞ *are* T*-invariant, and they split the measure* π*:*

$$\int_E f\,d\pi = \int_E f\,d\pi_1 + \int_E f\,d\pi_\infty, \quad f \in L^1(E,\pi). \tag{10.39}$$

If, $\mathbb{P}_x\left[\tau_A^0 < \infty\right] = 1$ *for* π*-almost all* $x \in E$*, then* $\pi_\infty = 0$ *and* $\pi_1 = \pi$*.*

Since $f \geqslant 0$, the infima and suprema in (10.37) and (10.38) are in fact limits. This observation follows from the equality in (10.40) and the invariance of the measure π.

Proof. From (10.36) we get:

$$\begin{aligned}
&\int_A \mathbb{E}_x\left[\sum_{k=1}^{\tau_A^1\wedge n} f\left(X(k+\ell)\right)\right] d\pi(x) + \int_{E\setminus A} \mathbb{E}_x\left[f\left(X(n+\ell)\right), \tau_A \geqslant n\right] d\pi(x) \\
&= \int_A \mathbb{E}_x\left[\sum_{k=1}^{\tau_A^1\wedge n} T^\ell f\left(X(k)\right)\right] d\pi(x) + \int_{E\setminus A} \mathbb{E}_x\left[T^\ell f\left(X(n)\right), \tau_A^0 \geqslant n\right] d\pi(x) \\
&= \int_E T^\ell f(x) d\pi(x) = \int_E f(x)\,d\pi(x).
\end{aligned} \tag{10.40}$$

The splitting in (10.39) follows from (10.40). If $\mathbb{P}_x\left[\tau_A^0 < \infty\right] = 1$ for π-almost all $x \in E$, then Corollary 10.1 yields $\pi_\infty = 0$, and hence $\pi_1 = \pi$.

This completes the proof of Corollary 10.3. □

10.2 Markov processes and invariant measures

In what follows we establish in the continuous time setting an analog to Theorem 10.3. The semigroup $\left\{e^{sL_A} : s \geqslant 0\right\}$ which we will use is defined by:

$$e^{sL_A} f(x) = \mathbb{E}_x\left[f\left(X(s)\right) : \tau_A > s\right], \quad f \in C_b(E). \tag{10.41}$$

Its generator L_A is pointwise defined by

$$L_A f(x) = \lim_{t\downarrow 0} \frac{e^{tL_A} f(x) - f(x)\mathbf{1}_{E\setminus A^r}(x)}{t} \tag{10.42}$$

for all functions $f \in C_b(E)$ for which these limits exist for all $x \in E$. Note that $e^{tL_A} f(x) = L_A f(x) = 0$ for $x \in A^r$. The semigroup e^{sL_A} lives on $E\setminus A^r$. Let $g \in C_b(E)$. Then the function Lg is defined to the extent to which the pointwise limit $Lg(x) = \lim_{h\downarrow 0} \dfrac{e^{hL} g(x) - g(x)}{h}$, $x \in E$, exists. In

Theorem 10.13, which is a consequence of Theorem 10.12 it will be shown under what conditions there exists, up to a multiplicative constant, a unique σ-finite invariant measure which is finite on compact subsets. The result should be compared with (10.24) in Theorem 10.4.

Theorem 10.5. *Suppose that the state space E of the irreducible time-homogeneous Markov process X is Polish. Let $A \in \mathcal{E}$ be such that $\mathbb{P}_x[\tau_A < \infty] = 1$ for π-almost all $x \in E$. Let π be a σ-finite invariant measure, and let $f \in L^1(E,\pi)$, $f \geqslant 0$. So it is assumed that the Borel measure π is such that for every compact subset K the following (in-)equalities hold: $0 \leqslant \int_E \mathbb{P}_x[X(h) \in K]d\pi(x) = \pi(K) < \infty$. It then follows that $\int_E e^{hL} f\, d\pi = \int_E f\, d\pi$, for all $f \in L^1(E,\mathcal{E},\pi)$. Then the following equalities hold:*

$$\begin{aligned}&\left(e^{hL} - e^{hL_A}\right)\int_0^t e^{sL_A} f\, ds + \int_0^h e^{\rho L_A} e^{tL_A} f\, d\rho \\ &= \int_0^h e^{\rho L_A} f\, d\rho + \left(e^{hL} - I\right)\int_0^t e^{sL_A} f\, ds, \end{aligned} \tag{10.43}$$

$$\left(e^{hL} - e^{hL_A}\right)\int_0^\infty e^{sL_A} f\, ds = \int_0^h e^{\rho L_A} f\, d\rho + \left(e^{hL} - I\right)\int_0^\infty e^{sL_A} f\, ds, \tag{10.44}$$

$$\begin{aligned}&\int_E \left(e^{hL} - e^{hL_A}\right)\int_0^t e^{sL_A} f\, ds\, d\pi + \int_E \int_0^h e^{\rho L_A} e^{tL_A} f\, d\rho\, d\pi \\ &= \int_E \int_0^h e^{\rho L_A} f\, d\rho\, d\pi, \end{aligned} \tag{10.45}$$

$$\lim_{h\downarrow 0} \int_E \frac{e^{hL} - e^{hL_A}}{h}\int_0^t e^{sL_A} f\, ds\, d\pi + \int_{A^r} f\, d\pi + \int_E e^{tL_A} f\, d\pi = \int_E f\, d\pi, \tag{10.46}$$

$$\int_E \left(e^{hL} - e^{hL_A}\right)\int_0^\infty e^{sL_A} f\, ds\, d\pi = \int_{E\setminus A^r} \int_0^h e^{\rho L_A} f\, d\rho\, d\pi, \tag{10.47}$$

$$\lim_{h\downarrow 0} \int_E \frac{e^{hL} - e^{hL_A}}{h}\int_0^\infty e^{sL_A} f\, ds\, d\pi + \int_{A^r} f\, d\pi = \int_E f\, d\pi, \quad \textit{and} \tag{10.48}$$

$$\lim_{t\to\infty} \int_E e^{tL_A} f\, d\pi = 0. \tag{10.49}$$

Suppose that the subset A possesses the additional properties that $\pi[A^r] < \infty$, and

$$\liminf_{h\downarrow 0} \frac{1}{h}\int_E \mathbb{E}_x\left[\mathbb{E}_{X(h)}[\tau_A],\, \tau_A \leqslant h\right] d\pi(x) < \infty. \tag{10.50}$$

Then the measure π is finite and

$$\pi[E] = \pi\left[A^r\right] + \lim_{h\downarrow 0}\frac{1}{h}\int_E \mathbb{E}_x\left[\mathbb{E}_{X(h)}\left[\tau_A\right],\, \tau_A \leqslant h\right] d\pi(x). \tag{10.51}$$

The discrete analog of formula (10.43) is the formula in (10.28). The finiteness result in (10.50) and hypothesis in Theorem 10.5 should be compared with the result in Corollary 10.5 under the assumption in (10.292). Here A^r stands for the collection of regular points of A:

$$A^r = \left\{x \in A : \mathbb{P}_x\left[\tau_A = 0\right] = 1\right\}.$$

From Blumenthal's zero-one law we know that $\mathbb{P}_x\left[\tau_A = 0\right] = 0$ or 1. Since the paths are right-continuous it follows that $A^r \subset \overline{A}$, where $\overline{A}$ is the (topological) closure of A.

Remark 10.2. Suppose that for any function $g \geqslant 0$ which is such that the function $\left(e^{hL} - I\right) g$ belongs to $L^1\left(E, \mathcal{E}, \pi\right)$, and is such that $\int \left(e^{hL} - I\right) g\, d\pi \leqslant 0$, then, by hypothesis, $\int \left(e^{hL} - I\right) g\, d\pi = 0$. The proof of Theorem 10.5 then shows that the equality in (10.49) holds: see equality (10.56) and the inequalities (10.64) and (10.65) below. However, such a hypothesis does not seem to be realistic. The present proof uses (10.256) in Proposition 10.7 below: see inequality (10.270).

Proof. [Proof of Theorem 10.5.] The equality in (10.43) is a consequence of

$$\left(I - e^{hL_A}\right)\int_0^t e^{sL_A} f\, ds = \int_0^h e^{\rho L_A} f\, d\rho\left(f - e^{tL_A}\right) f.$$

Notice that all terms in (10.43), except the last one, are non-negative provided that $f \geqslant 0$. The equality in (10.45) follows by integrating the left-hand and right-hand side of (10.43) with respect to the invariant measure π whereby the fact has been used that functions of the form $\int_0^t e^{sL_A} f\, ds$ belong to $L^1\left(E, \mathcal{E}, \pi\right)$ for $t \in (0, \infty)$. Let $B \in \mathcal{E}$, and let $g \in L^1\left(E, \mathcal{E}, \pi\right)$. Since

$$\lim_{h\downarrow 0}\int_B \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} g\, d\pi = \int_{B\bigcap(E\setminus A^r)} g\, d\pi \tag{10.52}$$

the equality in (10.46) follows from (10.45). Since, by assumption, A is recurrent, in (10.43) we can let t tend to ∞ to obtain (10.44). Moreover, from (10.45) we deduce

$$\begin{aligned}\int_E \frac{e^{hL} - e^{hL_A}}{h}\int_0^t e^{sL_A} f\, ds\, d\pi &\leqslant \int_E \frac{\int_0^h e^{\rho L_A}\, d\rho}{h}\left(\mathbf{1}_{E\setminus A^r} f\right) d\pi \\ &\leqslant \int_E \frac{\int_0^h e^{\rho L}\, d\rho}{h}\left(\mathbf{1}_{E\setminus A^r} f\right) d\pi = \int_{E\setminus A^r} f\, d\pi. \end{aligned}\tag{10.53}$$

Letting $t \uparrow \infty$ in (10.53) and invoking monotone convergence we see

$$\int_E \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} f \, ds \, d\pi \leqslant \int_{E \setminus A^r} f \, d\pi. \tag{10.54}$$

Next we show that $\int_E \left(e^{hL} - I\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi = 0$. For this purpose we write:

$$\int_E \left(e^{hL} - e^{hL_A}\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi = \lim_{t\to\infty} \int_E \left(e^{hL} - e^{hL_A}\right) \int_0^t e^{sL_A} f \, ds \, d\pi$$

(the measure π is invariant)

$$= \lim_{t\to\infty} \int_E \left(I - e^{hL_A}\right) \int_0^t e^{sL_A} f \, ds \, d\pi = \lim_{t\to\infty} \int_E \int_0^h e^{\rho L_A} \, d\rho \left(I - e^{tL_A}\right) f \, d\pi$$
$$= \int_E \int_0^h e^{\rho L_A} f \, d\rho \, d\pi - \lim_{t\to\infty} \int_E \int_0^h e^{\rho L_A} e^{tL_A} f \, d\rho \, d\pi. \tag{10.55}$$

From (10.55) we obtain:

$$\int_E \left(e^{hL} - I\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi$$
$$= \int_E \left(e^{hL} - e^{hL_A}\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi - \int_E \left(I - e^{hL_A}\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi$$
$$= \int_E \int_0^h e^{\rho L_A} f \, d\rho \, d\pi - \lim_{t\to\infty} \int_E \int_0^h e^{\rho L_A} e^{tL_A} f \, d\rho \, d\pi - \int_E \int_0^h e^{\rho L_A} f \, d\rho \, d\pi$$
$$= -\lim_{t\to\infty} \int_E \int_0^h e^{(\rho+t)L_A} \left(\mathbf{1}_{E\setminus A^r} f\right) \, d\rho \, d\pi. \tag{10.56}$$

From (10.56) we see that for $f \geqslant 0$, $f \in L^1(E, \mathcal{E}, \pi)$, the inequality

$$\int_E \left(e^{hL} - I\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi \leqslant 0 \tag{10.57}$$

holds. In order to prove the reverse inequality we proceed as follows. Let $g \in D(L_A) \bigcap L^1(E, \mathcal{E}, \pi)$ be arbitrary. Then we have

$$\int_E \int_0^h e^{\rho L_A} e^{tL_A} \left(\mathbf{1}_{E\setminus A^r} f\right) \, d\rho \, d\pi$$
$$= \int_E \int_0^h e^{(\rho+t)L_A} \left(\mathbf{1}_{E\setminus A^r} (f - L_A g)\right) \, d\rho \, d\pi + \int_E \int_0^h e^{(\rho+t)L_A} L_A g \, d\rho \, d\pi$$
$$\leqslant \int_E \int_0^h e^{(\rho+t)L_A} \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \, d\rho \, d\pi + \int_E e^{(h+t)L_A} g \, d\pi - \int_E e^{tL_A} g \, d\pi$$

$$\leqslant \int_E \int_0^h e^{(\rho+t)L} \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \, d\rho \, d\pi + \int_E e^{(h+t)L_A} g \, d\pi - \int_E e^{tL_A} g \, d\pi$$
$$= h \int_E \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \, d\pi + \int_E e^{(h+t)L_A} g \, d\pi - \int_E e^{tL_A} g \, d\pi. \tag{10.58}$$

From (10.56) and (10.58) we obtain

$$\int_E \left(e^{hL} - I\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi \tag{10.59}$$
$$\geqslant -h \int_E \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \, d\pi + \liminf_{t\to\infty} \left(\int_E e^{tL_A} g \, d\pi - \int_E e^{(h+t)L_A} g \, d\pi \right).$$

From the equality in (10.45) it follows that the limit $\lim_{t\to\infty} \int_E e^{tL_A} g \, d\pi$ exists and is finite for $g \in L^1(E, \mathcal{E}, \pi)$. Consequently, the inequality in (10.59) entails

$$\int_E \left(e^{hL} - I\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi \geqslant -h \int_E \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \, d\pi \tag{10.60}$$

for all $g \in D(L_A) \bigcap L^1(E, \mathcal{E}, \pi)$. Fix $\alpha > 0$ but small, and put $g_\alpha = -\int_0^\infty e^{-\alpha s} e^{sL_A} f \, ds$. Then the function g_α belongs to $L^1(E, \mathcal{E}, \pi)$; in fact $-\alpha \int_E g_\alpha \, d\pi \leqslant \int_{E\setminus A^r} f d\pi$. In addition, we have

$$\mathbf{1}_{E\setminus A^r} f - L_A g_\alpha$$
$$= \mathbf{1}_{E\setminus A^r} f - (\alpha I - L_A) \int_0^\infty e^{-\alpha s} e^{sL_A} f \, ds + \alpha \int_0^\infty e^{-\alpha s} e^{sL_A} f \, ds$$
$$= \mathbf{1}_{E\setminus A^r} f - \mathbf{1}_{E\setminus A^r} f + \alpha \int_0^\infty e^{-\alpha s} e^{sL_A} f \, ds = \alpha \int_0^\infty e^{-\alpha s} e^{sL_A} f \, ds. \tag{10.61}$$

By (10.256) in Proposition 10.7 below we see that

$$\lim_{\alpha\downarrow 0} \alpha \int_E \int_0^\infty e^{-\alpha s} e^{sL_A} f \, ds \, d\pi = 0. \tag{10.62}$$

Observe that the proof of Proposition 10.7 does not depend on Theorem 10.5: see inequality (10.270). From (10.57), (10.60), (10.61) and (10.62) we infer

$$\lim_{h\downarrow 0} \int_E \left(e^{hL} - I\right) \int_0^\infty e^{sL_A} f \, ds \, d\pi = 0. \tag{10.63}$$

Using this fact, and integrating the equality in (10.44) results in equality (10.47). Dividing the terms in (10.47) by $h > 0$, letting $h \downarrow 0$, and employing (10.52) leads to the equality in (10.48).

The proof of the equality in (10.49) follows from the arguments leading to (10.58). More precisely, let $h > 0$ be arbitrary. Then we have

$$\begin{aligned}\int_E e^{tL_A} f\, d\pi &= \int_E e^{tL_A}\left(f - \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\right) d\pi + \int_E e^{tL_A} \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\, d\pi \\ &\leqslant \int_E e^{tL}\left|\mathbf{1}_{E\setminus A^r} f - \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\right| d\pi + \int_E e^{tL_A} \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\, d\pi \\ &= \int_E \left|\mathbf{1}_{E\setminus A^r} f - \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\right| d\pi + \int_E e^{tL_A} \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\, d\pi.\end{aligned} \tag{10.64}$$

Then (10.45), the inequalities in (10.58) together with (10.61) and (10.62) applied to (10.64) implies

$$\lim_{t\to\infty} \int_E e^{tL_A} f\, d\pi \leqslant \int_E \left|\mathbf{1}_{E\setminus A^r} f - \frac{\int_0^h e^{\rho L_A}\, d\rho}{h} f\right| d\pi, \quad h > 0. \tag{10.65}$$

Since the right-hand side of (10.65) tends to 0 when $h \downarrow 0$ the equality in (10.49) follows.

The equality in (10.51) follows from (10.48) by putting $f = \mathbf{1}$. Notice the equality

$$\int_0^t e^{sL_A} f(x)\, ds = \mathbb{E}_x\left[\int_0^{t\wedge\tau_A} f(X(s))\, ds\right], \quad f \geqslant 0, \quad t > 0.$$

This completes the proof of Theorem 10.5. □

The following corollary is similar to Corollary 10.1 which in turn followed from the proof of Theorem 10.3. It is a direct consequence of (10.49).

Corollary 10.4. *Let the notation and assumptions be as in Theorem 10.5. Then the following equality holds for all* $f \in L^1(E, \mathcal{E}, \pi)$*:*

$$\lim_{t\to\infty} \int_{E\setminus A^r} \mathbb{E}_x\left[f(X(t)),\, \tau_A > t\right] d\pi(x) = 0. \tag{10.66}$$

10.2.1 *Some additional relevant results*

Theorem 10.6 is the most important result of the present subsection. We will assume that the squared gradient operator Γ_1 exists for functions in the domain of the generator of our Markov process $\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t), t \geqslant 0), (\vartheta_t, t \geqslant 0), (E, \mathcal{E})\}$. For the definition of squared gradient operator the reader is referred to the formulas (5.2) and (5.3) in

Chapter 5 or formula (7.7) in Chapter 7. Let v be a function in $D(L)$ and define the martingale $M_v(t)$ by $M_v(t) = v(X(t)) - v(X(0)) - \int_0^t Lv(X(s))$. A consequence of Proposition 5.3 is the following equality:

$$\begin{aligned} &\mathbb{E}_x \left[\left| v(X(t)) - v(X(0)) - \int_0^t Lv(X(s)) \right|^2 ds \right] \\ &= \mathbb{E}_x \left[\langle M_v, M_v \rangle (t) \right] = \int_0^t \mathbb{E}_x \left[\Gamma_1(v,v)(X(s)) \right] ds. \end{aligned} \tag{10.67}$$

Moreover, the separation property (a) in Proposition 9.1 has to be strengthened to inequality (10.68) in (a′):

(a′) For every $x \in E \backslash A^r$ and every $\varepsilon > 0$ there exists a function $v \in D(L)$ such that the following inequality holds for all $y \in A$:

$$\left(\Gamma_1(v,v)(x)\right)^{1/2} < \varepsilon \left(v(x) - v(y)\right). \tag{10.68}$$

The following remark serves to support the idea that in Theorem 10.5 it is not so obvious to take limits for $h \downarrow 0$.

Remark 10.3. It is tempting to take the limit for $h \downarrow 0$ in equality (10.53) of Theorem 10.5. This limit would be

$$(L - L_A) \int_0^t e^{sL_A} f\, ds + e^{tL_A} f = \mathbf{1}_{E \backslash A^r} f + L \int_0^t e^{sL_A} f\, ds. \tag{10.69}$$

However, it is not so clear how to define $(L - L_A) \int_0^t e^{sL_A} f\, ds(x)$ for $x \in A^r$. Under the condition (a′) we will show that

$$\lim_{h \downarrow 0} \frac{e^{hL} - e^{hL_A}}{h} \int_0^t e^{sL_A} f\, ds(x) = 0, \quad x \in E \backslash A^r. \tag{10.70}$$

Sometimes it is convenient to know circumstances under which an equality of the form $Lf - L_A f = \mathbf{1}_{A^r} Lf$, $f \in D(L)$, holds. Such an equality is a consequence of hypothesis (a′). In addition we have

$$L_A \int_0^t e^{sL_A} f\, ds = e^{tL_A} f - f, \quad \text{on } E \backslash A^r, \tag{10.71}$$

an equality which was also employed in the proof of Theorem 10.5. We will need the following lemma: it resembles Proposition 9.1. A somewhat more sophisticated version will show that the limit in (10.42) is in fact a strict limit on $E \backslash A^r$ provided that $E \backslash A^r$ is an open subset of E. The latter is the case if $A = A^r$ is a closed subset of E, in other words, if all points of the closed set A are regular for the Markov process.

Lemma 10.1. *Let $x \in E\backslash A^r$, and suppose that (a′) is satisfied. Then the following equalities hold:*

$$\lim_{t\downarrow 0} \frac{\mathbb{P}_x\left[\tau_A \leqslant t\right]}{t} = 0, \quad and \tag{10.72}$$

$$\left(L - L_A\right) f(x) = 0 \quad for\ f \in D(L). \tag{10.73}$$

In addition: $\left(L - L_A\right) f = \mathbf{1}_{A^r} Lf$, and consequently $L_A f = \mathbf{1}_{E\backslash A^r} Lf$ for $f \in D(L)$.

If A is closed, then the limit in (10.42) and (10.72) can be taken uniformly on compact subsets of the open subset $E\backslash A$, so that for $f \in D(L)$ the limit in (10.42) is in fact a limit in terms of the strict topology on $E\backslash A$.

Remark 10.4. Suppose that the subset A is closed. Then the convergence in (10.72) is uniformly for x in compact subsets of $E\backslash A$. Since we have the following inclusion of events $\{X(s) \in A\} \subset \{\tau_A \leqslant s\}$, $s > 0$, it follows that the paths of the Markov process $\left\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t), t \geqslant 0), (\vartheta_t, t \geqslant 0), (E, \mathcal{E})\right\}$ are necessarily $\mathbb{P}_x$-almost surely continuous. For this result see Proposition 4.6 in §4.4.

Proof. [Proof of Lemma 10.1.] Fix $x \in E\backslash A^r$ and let $\varepsilon > 0$. Then by assumption (a′) there exists a function $v \in D(L)$ such that on the event $\{\tau_A \leqslant t\}$ the following inequality holds $\mathbb{P}_x$-almost surely:

$$\begin{aligned}&\left(v(x) - \sup_{y\in A^r} v(y) + t \inf_{y\in E\backslash A^r} \min\left(Lv(y), 0\right)\right) \mathbf{1}_{\{\tau_A \leqslant t\}} \\ &\leqslant \left(v\left(X(0)\right) - v\left(X\left(\tau_A\right)\right) + \int_0^{\tau_A} Lv\left(X(s)\right)\, ds\right) \mathbf{1}_{\{\tau_A \leqslant t\}}. \end{aligned}\tag{10.74}$$

In (10.74) we write $M_v(t) = v\left(X(t)\right) - v\left(X(0)\right) - \int_0^t Lv\left(X(s)\right) ds$, and we take expectations to obtain by the martingale property of the process $t \mapsto M_v(t)$:

$$\begin{aligned}&\left(v(x) - \sup_{y\in A^r} v(y) + t \inf_{y\in E\backslash A^r} \min\left(Lv(y), 0\right)\right) \mathbb{P}_x\left[\tau_A \leqslant t\right] \\ &\leqslant \mathbb{E}_x\left[-M_v\left(\tau_A\right), \tau_A \leqslant t\right] = \mathbb{E}_x\left[-M_v\left(t\right), \tau_A \leqslant t\right]. \end{aligned}\tag{10.75}$$

Applying the Cauchy-Schwarz' inequality to the right-hand side of (10.75) and employing the equality in (10.67) yields

$$\begin{aligned}&\left(v(x) - \sup_{y\in A^r} v(y) + t \inf_{y\in E\backslash A^r} \min\left(Lv(y), 0\right)\right) \mathbb{P}_x\left[\tau_A \leqslant t\right] \\ &\leqslant \left(\mathbb{E}_x\left[\left|M_v(t)\right|^2\right]\right)^{1/2} \left(\mathbb{P}_x\left[\tau_A \leqslant t\right]\right)^{1/2} \\ &= \left(\mathbb{E}_x\left[\int_0^t \Gamma_1(v,v)\left(X(s)\right) ds\right]\right)^{1/2} \left(\mathbb{P}_x\left[\tau_A \leqslant t\right]\right)^{1/2}. \end{aligned}\tag{10.76}$$

By continuity of the function $y \mapsto \Gamma_1(v,v)(y)$, $y \in E$, and using assumption (a$'$) from (10.76) we obtain for $0 < t \leqslant t_\varepsilon$:

$$\begin{aligned} &\left(v(x) - \sup_{y\in A^r} v(y) + t \inf_{y\in E\backslash A^r} \min\left(Lv(y),0\right)\right) \left(\mathbb{P}_x\left[\tau_A \leqslant t\right]\right)^{1/2} \\ &\leqslant \varepsilon t^{1/2}\left(v(x) - \sup_{y\in A^r} v(y) + t \inf_{y\in E\backslash A^r} \min\left(Lv(y),0\right)\right), \qquad (10.77) \end{aligned}$$

and hence $\mathbb{P}_x\left[\tau_A \leqslant t\right] \leqslant \varepsilon^2 t$ for $0 < t \leqslant t_\varepsilon$. This proves equality (10.72) in Lemma 10.1. We use the equality in (10.72) to prove that $(L - L_A) f(x) = 0$. Therefore we write

$$\begin{aligned} &\left|e^{tL} f(x) - e^{tL_A} f(x)\right| = \left|\mathbb{E}_x\left[f(X(t))\right] - \mathbb{E}_x\left[f(X(t)),\, \tau_A > t\right]\right| \\ &= \left|\mathbb{E}_x\left[f(X(t)),\, \tau_A \leqslant t\right]\right| \leqslant \|f\|_\infty \mathbb{P}_x\left[\tau_A \leqslant t\right]. \qquad (10.78) \end{aligned}$$

The equality in (10.73) follows from (10.72) and (10.78). Finally let $f \in D(L)$. Equality (10.73) shows that $(L - L_A) f(y) = 0$ for $y \in E\backslash A^r$. If $y \in A^r$, then $L_A f(y) = 0$. Consequently $(L - L_A) f = \mathbf{1}_{A^r} Lf$.

In order to finish the proof of Lemma 10.1 we have to prove that the indicated limits are uniform on compact subsets of $E\backslash A$ when A is a closed subset of E. So from now on A is a closed subset of E. First we do this for the limit in (10.72). Let K be a compact subset of $E\backslash A$, and let $\varepsilon > 0$ be arbitrary. Then there exist functions $v_j \in D(L)$, $1 \leqslant j \leqslant N$, such that

$$K \subset \bigcup_{j=1}^N \left\{x \in E\backslash A :\, \left(\Gamma_1(v_j,v_j)(x)\right)^{1/2} < \varepsilon\left(v_j(x) - \sup_{y\in E\backslash A} v_j(y)\right)\right\}. \qquad (10.79)$$

By uniform continuity of the functions $(t,y) \mapsto \dfrac{\int_0^t \mathbb{E}_y\left[\Gamma_1(v_j,v_j)(X(s))\right] ds}{t}$, $1 \leqslant j \leqslant N$, on the compact subset $[0,t_0] \times K$, for $t_0 > 0$, the inclusion in (10.79) entails that there exists a strictly real number $t_\varepsilon > 0$ such that $K = \bigcup_{j=1}^N K_j$ where $x \in K$ belongs to K_j if and only if the inequality

$$\begin{aligned} &\left(\mathbb{E}_x\left[\int_0^t \Gamma_1(v_j,v_j)(X(s))\, ds\right]\right)^{1/2} \\ &< \varepsilon t^{1/2}\left(v_j(x) - \sup_{y\in A} v_j(y) + t \inf_{y\in E\backslash A} \min\left(Lv(y),0\right)\right) \qquad (10.80) \end{aligned}$$

holds for $0 < t \leqslant t_\varepsilon$. Here, of course, $\dfrac{\int_0^t \mathbb{E}_y\left[\Gamma_1(v_j,v_j)(X(s))\right] ds}{t}$ is interpreted as $\Gamma_1(v_j,v_j)(y)$ when $t = 0$. Let $x \in K_j$. As in the proof of (10.72) we obtain $\mathbb{P}_x\left[\tau_A \leqslant t\right] \leqslant \varepsilon^2 t$ for $0 < t \leqslant t_\varepsilon$: see the inequalities in

(10.76) and (10.77). It follows that for $0 < t \leqslant t_\varepsilon$ and $x \in K$ we have $\mathbb{P}_x\left[\tau_A \leqslant t\right] \leqslant \varepsilon^2 t$. As a consequence we see that the limit in (10.72) is uniform for $x \in K$, where K is an arbitrary compact subset of $E\backslash A$, i.e.

$$\lim_{t\downarrow 0}\sup_{x\in K}\frac{\mathbb{P}_x\left[\tau_A \leqslant t\right]}{t} = 0, \quad \text{for all compact subsets } K \text{ of } E\backslash A. \tag{10.81}$$

The equality in (10.81) together with the inequality in (10.78) shows that the convergence in (10.42) is uniform for x in compact subsets of $E\backslash A$.

This concludes the proof of Lemma 10.1. □

The following lemma shows that equality (10.71) holds.

Lemma 10.2. *The equality in (10.71) holds for $f \in C_b(E)$.*

Proof. The proof follows a standard procedure. We write

$$\begin{aligned}
&e^{hL_A}\int_0^t e^{sL_A} f\, ds(x) - \int_0^t e^{sL_A} f(x)\, ds \\
&= \int_0^t e^{(s+h)L_A} f(x)\, ds - \int_0^t e^{sL_A} f(x)\, ds \\
&= \int_h^{t+h} e^{sL_A} f(x) ds - \int_0^t e^{sL_A} f(x)\, ds \\
&= \int_t^{t+h} e^{sL_A} f(x) ds - \int_0^h e^{sL_A} f(x)\, ds.
\end{aligned} \tag{10.82}$$

Upon dividing (10.82) by h and sending h to zero we obtain the equality in (10.71).

This completes the proof of Lemma 10.2. □

The following lemma shows that equality (10.70) holds.

Lemma 10.3. *The limit in (10.70) converges uniformly on compact subsets of $E\backslash A^r$.*

Proof. Without loss of generality we assume that $f \geqslant 0$, $f \in C_b(E)$. In order to prove (10.70) we write

$$\begin{aligned}
0 &\leqslant \frac{e^{hL} - e^{hL_A}}{h}\int_0^t e^{sL_A} f\, ds(x) = \frac{1}{h}\mathbb{E}_x\left[\left[\int_0^t e^{sL_A} f\left(X(h)\right)\, ds,\, 0 \leqslant \tau_A \leqslant h\right]\right. \\
&\leqslant t\left\|f\right\|_\infty \frac{\mathbb{P}_x\left[\tau_A \leqslant h\right]}{h}.
\end{aligned} \tag{10.83}$$

The equality in (10.70) follows from Lemma 10.1 and (10.83).

This shows the claim in Lemma 10.3. □

The following theorem transfers properties of the invariant measure to properties on $E\backslash A^r$ provided that certain conditions are satisfied. As in Theorem 10.5 the result should be compared with (10.24) in Theorem 10.4.

Theorem 10.6. *Let the hypotheses in Theorem 10.5 be strengthened with hypothesis (a′). It is understood that the measure π is L-invariant with properties as described in Theorem 10.5. Then the following equality holds for $f \in L^1(E, \mathcal{E}, \pi)$:*

$$\lim_{h\downarrow 0,\, h>0} \frac{1}{h} \int_{A^r} e^{hL} \int_0^\infty e^{sL_A} f\, ds\, d\pi + \int_{A^r} f\, d\pi = \int_E f\, d\pi. \tag{10.84}$$

Notice that $\mathbf{1}_{A^r} g = 0$ π-almost everywhere whenever g belongs the L^1-domain of L_A. This fact is used in (10.86) below.

Proof. Let $f \geqslant 0$ belong to $L^1(E, \mathcal{E}, \pi)$. From equality (10.47) in Theorem 10.5 it follows that, in order to obtain (10.84), it suffices to prove that

$$\lim_{h\downarrow 0} \int_{E\backslash A^r} \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} f\, ds\, d\pi = 0. \tag{10.85}$$

To achieve the equality in (10.85) we choose $g \in D(L) \bigcap D(D_A)$ arbitrarily, and notice

$$\begin{aligned}
&\frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} f\, ds \\
&= \frac{e^{hL} - e^{hL_A}}{h} \left(\int_0^\infty e^{sL_A} (f - L_A g)\, ds + \int_0^\infty e^{sL_A} L_A g\, ds \right) \\
&= \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} (f - L_A g)\, ds - \frac{e^{hL} - e^{hL_A}}{h} \left(\mathbf{1}_{E\backslash A^r} g \right) \\
&= \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} (f - L_A g)\, ds - \frac{e^{hL} - I}{h} g + \frac{e^{hL_A} - I}{h} g.
\end{aligned} \tag{10.86}$$

Next we integrate (10.86) with respect to π and invoke (10.47) to obtain

$$\begin{aligned}
&\int_{E\backslash A^r} \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} f\, ds\, d\pi \\
&\quad + \int_{E\backslash A^r} \frac{e^{hL} - I}{h} g\, d\pi - \int_{E\backslash A^r} \frac{e^{hL_A} - I}{h} g\, d\pi \\
&= \int_{E\backslash A^r} \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} (f - L_A g)\, ds\, d\pi
\end{aligned}$$

$$\leqslant \int_E \frac{\int_0^h e^{\rho L_A}\,d\rho}{h} \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \,d\pi \leqslant \int_E \frac{\int_0^h e^{\rho L}\,d\rho}{h} \left|\mathbf{1}_{E\setminus A^r} f - L_A g\right| \,d\pi$$
$$= \int_{E\setminus A^r} |f - L_A g| \,d\pi. \tag{10.87}$$

In (10.87) we let $h \downarrow 0$ and deduce, for $g \in D(L) \bigcap D(L_A)$ arbitrary,

$$\limsup_{h\downarrow 0} \int_{E\setminus A^r} \frac{e^{hL} - e^{hL_A}}{h} \int_0^\infty e^{sL_A} f \,ds\,d\pi + \int_{E\setminus A^r} Lg\,d\pi - \int_{E\setminus A^r} L_A g\,d\pi$$
$$\leqslant \int_{E\setminus A^r} |f - L_A g| \,d\pi. \tag{10.88}$$

By taking g of the form $g = g_\alpha = -\int_0^\infty e^{-\alpha s} e^{sL_A} f\,ds$, and applying (10.256) in Proposition 10.7 below we get that the right-hand side of (10.88) is as small as we please: see (10.61) in the proof of Theorem 10.5. Using the arguments in Lemma 10.1, which depends directly on the hypothesis in (a$'$), we see that $Lg_\alpha = L_A g_\alpha$ on $E\setminus A^r$. The inequality in (10.88) then entails the equality in (10.85), which completes the proof of Theorem 10.6. □

10.2.2 *An attempt to construct an invariant measure*

In this subsection we will try to give a construction of an, up to a multiplicative constant, unique σ-finite measure provided the hypotheses of Theorem 10.5 are fulfilled. We will use Dynkin's formula, and we will employ resolvent techniques: see (10.21). We will begin with establishing a number of relevant formulas, which we collect in Proposition 10.5 below. In what follows we employ the following notation:

$$\begin{aligned} R(\lambda)f(x) &= (\lambda I - L)^{-1} f(x) = \int_0^\infty e^{-\lambda s} e^{sL} f(x)\,ds \\ &= \int_0^\infty e^{-\lambda s} \mathbb{E}_x \left[f(X(s)) \right] ds, \end{aligned} \tag{10.89}$$

$$\begin{aligned} R_A(\lambda)f(x) &= (\lambda I - L_A)^{-1} f(x) \\ &= \mathbb{E}_x \left[\int_0^{\tau_A} e^{-\lambda s} f(X(s))\,ds \right] \\ &= \int_0^\infty e^{-\lambda s} \mathbb{E}_x \left[f(X(s)) : \tau_A > s \right] ds \\ &= \int_0^\infty e^{-\lambda s} e^{sL_A} f(x) ds, \end{aligned} \tag{10.90}$$

$$\begin{aligned} P_A(\lambda) &= (L - \lambda I) R_A(\lambda) + I = \mathbf{1}_{A^r} \left((L - \lambda I) R_A(\lambda) + I \right) \\ &= (\lambda I - L) H_A(\lambda) R(\lambda) = \mathbf{1}_{A^r} (\lambda I - L) H_A(\lambda) R(\lambda), \end{aligned} \tag{10.91}$$

where

$$H_A(\lambda)f(x) = \begin{cases} \mathbb{E}_x\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right], & x \in E\backslash A^r, \\ f(x), & x \in A^r. \end{cases} \tag{10.92}$$

The equalities in (10.91) follow from equality (10.130) in Proposition 10.4.

Lemma 10.4. *Let $f \in D(L)$ be such that $H_A(\lambda)f \in D(L)$. Then the equalities in (10.91) yield*

$$P_A(\lambda)\left(L - \lambda I\right) f = \left(L - \lambda I\right) H_A(\lambda) f \quad \textit{and} \tag{10.93}$$

$$\lambda R(\lambda)\left\{\left(e^{-\lambda' h} e^{hL} - I\right) R_A\left(\lambda'\right) + \int_0^h e^{-\lambda' s} e^{sL} ds\right\}\left(L - \lambda' I\right) f \tag{10.94}$$

$$= \lambda R(\lambda)\left(e^{-\lambda' h} e^{hL} - I\right) H_A\left(\lambda'\right) f \tag{10.95}$$

$$= \lambda(\lambda R(\lambda) - I)\int_0^h e^{-\lambda' s} e^{sL} ds\, H_A(\lambda') f - \lambda'\lambda R(\lambda)\int_0^h e^{-\lambda' s} e^{sL} ds\, H_A(\lambda') f,$$

for $h > 0$, $\lambda' > 0$, and $\lambda > 0$.

Proof. The equality in (10.93) is an immediate consequence of (10.91). The equality in (10.94) is a consequence of the following identities:

$$R_A\left(\lambda'\right)\left(L - \lambda' I\right) f + f = H_A\left(\lambda'\right) f \quad \text{and} \tag{10.96}$$

$$\int_0^h e^{-\lambda' s} e^{sL}\left(L - \lambda' I\right) f ds = e^{-\lambda' h} e^{hL} f - f. \tag{10.97}$$

The equalities in (10.96) and (10.97) hold for $f \in D(L)$, $\lambda' > 0$, and $h > 0$. The equality in (10.95) is closely related to (10.93). This can be seen as follows:

$$\lambda R(\lambda)\left\{\left(e^{-\lambda' h} e^{hL} - I\right) R_A\left(\lambda'\right) + \int_0^h e^{-\lambda' s} e^{sL} ds\right\}\left(L - \lambda' I\right) f$$

$$= \lambda R(\lambda)\left\{\left(L - \lambda' I\right)\int_0^h e^{-\lambda' s} e^{sL} ds R_A\left(\lambda'\right) + \int_0^h e^{-\lambda' s} e^{sL} ds\right\}\left(L - \lambda' I\right) f$$

$$= \lambda\left(L - \lambda' I\right) R(\lambda)\int_0^h e^{-\lambda' s} e^{sL} ds\left\{R_A\left(\lambda'\right)\left(L - \lambda' I\right) + I\right\} f$$

$$= \lambda\left(L - \lambda' I\right) R(\lambda)\int_0^h e^{-\lambda' s} e^{sL} ds\, H_A\left(\lambda'\right) f. \tag{10.98}$$

Since $LR(\lambda) = \lambda R(\lambda) - I$, equality (10.94) is a consequence of (10.98). In order to obtain (10.98) we also used the identity

$$\left(L - \lambda' I\right)\int_0^h e^{-\lambda' s} e^{sL} f\, ds = e^{-\lambda' h} e^{hL} f - f, \quad f \in C_b(E).$$

This completes the proof of Lemma 10.4. □

Next, fix $h > 0$, $\lambda > 0$, $\mu \in \mathcal{M}(A)$, and $f \in C_b(E)$. Here $\mathcal{M}(A)$ is the space of those (complex) measures $\mu \in \mathcal{E}$ which are concentrated on A; i.e. $|\mu|(E\backslash A) = 0$ (see notation introduced prior to Theorem 2.2). We will also need the following stopping times, operators, and functionals:

$$\tau_A^{0,h} = \inf\{kh : k \geqslant 0, k \in \mathbb{N}, X(kh) \in A\}, \tag{10.99}$$

$$\tau_A^{1,h} = \inf\{kh : k \geqslant 1, k \in \mathbb{N}, X(kh) \in A\} = h + \tau_A^{0,h} \circ \vartheta_h, \tag{10.100}$$

$$\begin{aligned} L^h(\lambda)f(x) &= \frac{1}{h}\left(e^{-\lambda h}e^{hL} - I\right)f(x) = \frac{1}{h}\left(e^{-\lambda h}\mathbb{E}_x\left[f(X(h))\right] - f(x)\right) \\ &= \frac{1}{h}\left(\mathbb{E}_x\left[e^{-\lambda h}f(X(h)) - f(X(0))\right]\right), \end{aligned} \tag{10.101}$$

$$H_A^h(\lambda)f(x) = \mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}}f\left(X\left(\tau_A^{0,h}\right)\right)\right], \tag{10.102}$$

$$\begin{aligned} R_A^h(\lambda)f(x) &= h\mathbb{E}_x\left[\sum_{k=0}^{\tau_A^{0,h}/h-1} e^{-\lambda kh}f(X(kh)), \tau_A^{0,h} \geqslant h\right] \\ &= h\sum_{k=0}^{\infty}\mathbb{E}_x\left[e^{-\lambda kh}f(X(kh)), \tau_A^{0,h} \geqslant (k+1)h\right], \end{aligned} \tag{10.103}$$

$$P_A^h(\lambda)(f) = L^h(\lambda)H_A^h(\lambda)L^h(\lambda)^{-1}f = L^h(\lambda)R_A^h(\lambda)f + f. \tag{10.104}$$

$$\Lambda_{A,\mu}^h(\lambda)(f) = \int L^h(\lambda)H_A^h(\lambda)L^h(\lambda)^{-1}f(x)\,d\mu(x). \tag{10.105}$$

The second equality in (10.104) follows from equality (10.106) in Proposition 10.2 below.

Instead of $\Lambda_{A,\delta_{x_0}}^h(\lambda)$ we write $\Lambda_{A,x_0}^h(\lambda)$, when $\mu = \delta_{x_0}$ is the Dirac measure at x_0. Instead of $L^h(0)$ we write L^h. For the definition of the stopping times $\tau_A^{0,h}$ and $\tau_A^{1,h}$ see (10.25) in Theorem 10.4. Put

$$\tau_A = \inf\{s > 0 : X(s) \in A\} = \inf_{h>0}\tau_A^{1,h} = \lim_{h\downarrow 0}\tau_A^{1,h}.$$

Proposition 10.2. *The following identity holds:*

$$H_A^h(\lambda) = I + R_A^h(\lambda)L^h(\lambda), \quad h > 0,\ \lambda > 0. \tag{10.106}$$

Moreover, the equality in (10.106) is equivalent to equality (10.120) below. In addition, $H_A^h(\lambda)R_A^h(\lambda) = 0$, and hence $H_A^h(\lambda)^2 = H_A^h(\lambda)$. If $\mathbb{P}_x\left[\tau_A^{0,h} < \infty\right] = 1$ for all $x \in E$, then equality (10.106) holds for $\lambda = 0$. If A is recurrent in the sense that $\mathbb{P}_x[\tau_A < \infty] = 1$ for all $x \in E$, then $H_A(0) = I + R_A(0)L$, and $H_A(0)^2 = H_A(0)$, where $H_A(0)f(x) = \mathbb{E}_x[f(X(\tau_A)), \tau_A < \infty]$.

Proof. Let us check the equality in (10.106). To this end we consider:

$$
\begin{aligned}
&R_A^h(\lambda)L^h(\lambda)f(x)\\
&= h\sum_{k=0}^{\infty}\mathbb{E}_x\left[e^{-\lambda kh}L^h(\lambda)f\left(X(kh)\right),\, \tau_A^{0,h}\geqslant (k+1)h\right]\\
&= \sum_{k=0}^{\infty}\mathbb{E}_x\left[e^{-\lambda kh}\left\{e^{-\lambda h}\mathbb{E}_{X(kh)}\left[f\left(X(h)\right)\right]-f\left(X(kh)\right)\right\},\, \tau_A^{0,h}\geqslant (k+1)h\right]\\
&= \sum_{k=0}^{\infty}e^{-\lambda(k+1)h}\mathbb{E}_x\left[\mathbb{E}_{X(kh)}\left[f\left(X(h)\right),\, \tau_A^{0,h}\geqslant (k+1)h\right]\right]\\
&\quad - \sum_{k=0}^{\infty}e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h}\geqslant (k+1)h\right]
\end{aligned}
$$

(observe that the event $\left\{\tau_A^{0,h}\geqslant (k+1)h\right\}$ is $\mathcal{F}_{kh}$-measurable) and employ the strong Markov property

$$
\begin{aligned}
&= \sum_{k=0}^{\infty}e^{-\lambda(k+1)h}\mathbb{E}_x\left[f\left(X\left((k+1)h\right)\right),\, \tau_A^{0,h}\geqslant (k+1)h\right]\\
&\quad - \sum_{k=0}^{\infty}e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h}\geqslant (k+1)h\right]\\
&= \sum_{k=0}^{\infty}e^{-\lambda kh}\mathbb{E}_x\left[f\left(X\left(kh\right)\right),\, \tau_A^{0,h}\geqslant kh\right]-f(x)\\
&\quad - \sum_{k=0}^{\infty}e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h}\geqslant (k+1)h\right]\\
&= \sum_{k=0}^{\infty}e^{-\lambda kh}\mathbb{E}_x\left[f\left(X\left(kh\right)\right),\, \tau_A^{0,h}= kh\right]-f(x)\\
&= \mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}}f\left(X\left(\tau_A^{0,h}\right)\right)\right]-f(x) = H_A^h(\lambda)f(x)-f(x).
\end{aligned}
\tag{10.107}
$$

The equality in (10.107) is the same as (10.106).

If $x\in A$, then $\mathbb{P}_x\left[\tau_A^{0,h}=0\right]=1$ and so $R_A^h(\lambda)f(x)=0$. Since $X\left(\tau_A^{0,h}\right)\in A$ it follows that, for $\lambda>0$,

$$
H_A^h(\lambda)R_A^h(\lambda)f(x)=\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}}R_A^hf\left(X\left(\tau_A^{0,h}\right)\right)\right]=0. \tag{10.108}
$$

The assertions for $h=\lambda=0$ follow by taking limits with respect to $\lambda\downarrow 0$ and $h\downarrow 0$ in the corresponding equality (10.106).

This completes the proof of Proposition 10.2. □

Proposition 10.3. *The following equalities hold for $f \in C_b(E)$, $\lambda \geqslant 0$, and $h > 0$:*

$$
\begin{aligned}
&L^h(\lambda)R_A^h(\lambda)f(x) + f(x)\\
&= \mathbb{E}_x\left[\sum_{k=0}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \tau_A^{0,h} = 0\right]\\
&= \mathbb{E}_x\left[\sum_{k=0}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \tau_A^{0,h} = 0\right]\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]\\
&= \mathbb{E}_x\left[\sum_{k=0}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right)\right]\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]\\
&= e^{-\lambda h}e^{hL}\mathbb{E}_{(\cdot)}\left[\sum_{k=0}^{h^{-1}\tau_A^{0,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \tau_A^{0,h} \geqslant h\right](x)\cdot\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]\\
&\qquad + f(x)\cdot\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]\\
&= \left(\frac{1}{h}e^{-\lambda h}e^{hL}R_A^h(\lambda)f(x) + f(x)\right)\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]\\
&= \left(L^h(\lambda)R_A^h(\lambda)f(x) + f(x)\right)\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]\\
&= \lim_{n\to\infty}\sum_{k=0}^{n} e^{-\lambda(k+1)h}e^{hL}\mathbb{E}_{(\cdot)}\Big[f(X(kh)),\, \tau_A^{0,h} \geqslant (k+1)h\Big](x)\cdot\mathbb{P}_x\Big[\tau_A^{0,h} = 0\Big]\\
&\quad + f(x)\cdot\mathbb{P}_x\left[\tau_A^{0,h} = 0\right].
\end{aligned} \tag{10.109}
$$

In addition, the following assertions are true:

(a) *The following formula holds:*

$$\left(L^h(\lambda)R_A^h(\lambda) + I\right)^2 = \mathbf{1}_A\left(L^h(\lambda)R_A^h(\lambda) + I\right). \tag{10.110}$$

This formula is also valid with $\lambda = 0$ provided that A is h-recurrent in the sense that $\mathbb{P}_x\left[\tau_A^{0,h} < \infty\right] = 1$ for all $x \in E$. If A is recurrent, i.e. if $\mathbb{P}_x\left[\tau_A < \infty\right] = 1$ for all $x \in E$, then the operator $P_A(0) = LR_A(0+) + I$ is a projection operator.

(b) *For $f \geqslant 0$ the function $P_A^h(\lambda)f$ is non-negative, and the function $\lambda \mapsto P_A^h(\lambda)f(x)$ increases when λ decreases. In addition, by the fifth equality in (10.109) it follows that with $h = h_n = 2^{-n}h'$ the sequence $n \mapsto P_A^{2^{-n}h'}(\lambda)$ increases where $h' > 0$ and $\lambda \geqslant 0$ are fixed.*

The equalities in (10.109) are the same as those in (10.119). They will be employed to prove that the invariant measure we will introduce is σ-finite.

Proof. We use the definitions of the operators $L^h(\lambda)$ and $R_A^h(\lambda)$ to obtain

$$
\begin{aligned}
&L^h(\lambda)R_A^h(\lambda)f(x) + f(x)\\
&= \sum_{k=0}^{\infty} e^{-\lambda kh}\left(e^{-\lambda h}e^{hL} - I\right)\mathbb{E}_{(\cdot)}\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right](x) + f(x)\\
&= \sum_{k=0}^{\infty} e^{-\lambda(k+1)h}e^{hL}\mathbb{E}_{(\cdot)}\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right](x)\\
&\quad - \sum_{k=0}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right] + f(x)\\
&= \sum_{k=0}^{\infty} e^{-\lambda(k+1)h}\mathbb{E}_x\left[\mathbb{E}_{X(h)}\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right]\right]\\
&\quad - \sum_{k=0}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right] + f(x)
\end{aligned}
$$

(Markov property)

$$
\begin{aligned}
&= \sum_{k=0}^{\infty} e^{-\lambda(k+1)h}\mathbb{E}_x\left[f\left(X\left(k+1\right)h\right),\, \tau_A^{0,h}\circ\vartheta_h \geqslant (k+1)h\right]\\
&\quad - \sum_{k=0}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right] + f(x)\\
&= \sum_{k=1}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X\left(kh\right)\right),\, h + \tau_A^{0,h}\circ\vartheta_h \geqslant (k+1)h\right]\\
&\quad - \sum_{k=0}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right] + f(x)\\
&= \sum_{k=1}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X\left(kh\right)\right),\, \tau_A^{1,h} \geqslant (k+1)h\right]\\
&\quad - \sum_{k=1}^{\infty} e^{-\lambda kh}\mathbb{E}_x\left[f\left(X(kh)\right),\, \tau_A^{0,h} \geqslant (k+1)h\right] + f(x)\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]
\end{aligned}
$$

(a sum of the form $\sum_{k=k_1}^{k_2} \alpha_k$ is interpreted as 0 if $k_2 < k_1$)

$$
= \mathbb{E}_x\left[\sum_{k=\left(h^{-1}\tau_A^{0,h}\right)\vee 1}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right)\right] + f(x)\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]
$$

$$= \mathbb{E}_x \left[\sum_{k=\left(h^{-1}\tau_A^{0,h}\right)\vee 1}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \, \tau_A^{0,h} = 0 \right] + f(x)\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]$$
$$+ \mathbb{E}_x \left[\sum_{k=\left(h^{-1}\tau_A^{0,h}\right)\vee 1}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \, \tau_A^{0,h} \geqslant h \right]$$

(on the event $\left\{\tau_A^{0,h} \geqslant h\right\}$ the equality $\tau_A^{1,h} = \tau_A^{0,h}$ holds $\mathbb{P}_x$-almost surely)

$$= \mathbb{E}_x \left[\sum_{k=1}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \, \tau_A^{0,h} = 0 \right] + f(x)\mathbb{P}_x\left[\tau_A^{0,h} = 0\right]$$

$$= \mathbb{E}_x \left[\sum_{k=0}^{h^{-1}\tau_A^{1,h}-1} e^{-\lambda kh} f\left(X(kh)\right), \, \tau_A^{0,h} = 0 \right]. \tag{10.111}$$

The equality in (10.111) shows the first equality in (10.109). The second and third equality follow from the equality $\mathbb{P}_x\left[\tau_A^{0,h} = 0\right] = 1$ which is true if and only if $x \in A$. The fourth equality in (10.109) is a consequence of the Markov property in conjunction with the third equality. The fifth equality just follows from the definition of the operator $R_A^h(\lambda)$. The sixth equality follows from Lebesgue's dominated convergence theorem, or from the monotone convergence theorem if $f \geqslant 0$.

(a). In order to prove assertion (a) in Proposition 10.3 we consider

$$\begin{aligned}\left(L^h(\lambda)R_A^h(\lambda) + I\right)^2 &= \left(L^h(\lambda)R_A^h(\lambda) + I\right) L^h(\lambda)R_A^h(\lambda) + L^h(\lambda)R_A^h(\lambda) + I \\ &= L^h(\lambda)\left(R_A^h(\lambda)L^h(\lambda) + I\right) R_A^h(\lambda) + L^h(\lambda)R_A^h(\lambda) + I\end{aligned}$$

(apply the equality in (10.106) and the final assertion in Proposition 10.2)

$$\begin{aligned}&= L^h(\lambda)H_A^h(\lambda)R_A^h(\lambda) + L^h(\lambda)R_A^h(\lambda) + I \\ &= L^h(\lambda)R_A^h(\lambda) + I.\end{aligned} \tag{10.112}$$

By taking limits as $\lambda \downarrow 0$, and $h \downarrow 0$ in

$$\left(L^h(\lambda)R_A^h(\lambda) + I\right)^2 = L^h(\lambda)R_A^h(\lambda) + I$$

the final conclusion in assertion (a) of Proposition 10.3 follows. Since $\mathbf{1}_A(x) = \mathbb{P}_x\left[\tau_A^{0,h} = 0\right]$ the final equality in (10.110) follows from (10.109).

(b). Observe that by the equalities in (10.109) and the definition of the operator $R_A^h(\lambda)$ we have

$$L^h(\lambda)R_A^h(\lambda) + I = \mathbf{1}_A\left(L^h(\lambda)R_A^h(\lambda) + I\right) = \mathbf{1}_A\left(\frac{e^{-\lambda h}e^{hL} - I}{h}R_A^h(\lambda) + I\right)$$
$$= \frac{1}{h}\mathbf{1}_A\left(e^{-\lambda h}e^{hL}R_A^h(\lambda) + I\right). \tag{10.113}$$

From (10.113) it follows that for $f \geqslant 0$ the function $P_A^h(\lambda)f$ is non-negative, and that the function $\lambda \mapsto P_A^h(\lambda)f(x)$ increases when λ decreases. In addition, by the fifth equality in (10.109) it follows that for $h = 2^{-n}h'$ the sequence $n \mapsto P_A^{2^{-n}h'}(\lambda)$ where $h' > 0$ and $\lambda \geqslant 0$ are fixed. The equality in (10.112) and the latter observations complete the proof of Proposition 10.3. □

Fix $\mu \in \mathcal{P}(E)$. An attempt to define an invariant measure π goes as follows. It is determined by the functional

$$\begin{aligned}\Lambda_{A,\mu} : f \mapsto &\lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\int \left(\lambda I - L^h(\lambda)\right) H_A^h(\lambda)\left(\lambda I - L^h(\lambda)\right)^{-1} f(x)\, d\mu(x) \\ &= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\int L^h(\lambda)\left(H_A^h(\lambda) - I\right) L^h(\lambda)^{-1} f(x)\, d\mu(x) + \int f(x)\, d\mu(x) \\ &= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\int L^h(\lambda)R_A^h(\lambda)f(x)\, d\mu(x) + \int f(x)\, d\mu(x) \\ &= \lim_{h\downarrow 0}\int L^h R_A^h(0)f(x)\, d\mu(x) + \int f(x)\, d\mu(x).\end{aligned} \tag{10.114}$$

In (10.114) we employed equality (10.106). Let us try to check the L-invariance of the functional in (10.114). To this end we fix $f \in D(L)$. Then

$$Lf = \mathcal{T}_\beta\text{-}\lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} L^h(\lambda)f = Lf, \quad \text{and}$$

$$\begin{aligned}\Lambda_{A,\mu}Lf &= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\Lambda_{A,\mu}^h(\lambda)\left(L^h(\lambda)f\right) = \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\int L^h(\lambda)H_A^h(\lambda)f\, d\mu \\ &= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\frac{1}{h}\int_A \left(\mathbb{E}_x\left[e^{-\lambda h}H_A^h(\lambda)f(X(h))\right] - H_A^h(\lambda)f(x)\right) d\mu(x) \\ &= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\frac{1}{h}\int_A \left(\mathbb{E}_x\left[e^{-\lambda\tau_A^{1,h}} f\left(X\left(\tau_A^{1,h}\right)\right)\right]\right. \\ &\qquad \left. -\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right) d\mu(x)\end{aligned}$$

(for $x \in E\backslash A$ the equality $\tau_A^{1,h} = \tau_A^{0,h}$ holds $\mathbb{P}_x$-almost surely)

$$= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\frac{1}{h}\int_E \left(\mathbb{E}_x\left[e^{-\lambda\tau_A^{1,h}} f\left(X\left(\tau_A^{1,h}\right)\right)\right]\right.$$

$$
\begin{aligned}
&\quad -\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right) d\mu(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\frac{1}{h}\int_E \left(\mathbb{E}_x\left[e^{-\lambda\tau_A^{1,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\circ\vartheta_h\right]\right.\\
&\quad \left. -\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right) d\mu(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\frac{e^{-\lambda h}}{h}\int_E \left(\mathbb{E}_x\left[e^{-\lambda\tau_A^{h}\circ\vartheta_h} f\left(X\left(\tau_A^{0,h}\right)\right)\circ\vartheta_h\right]\right.\\
&\quad \left. -\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right) d\mu(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0}\frac{e^{-\lambda h}}{h}\int_E \left(\mathbb{E}_x\left[\mathbb{E}_{X(h)}\left[e^{-\lambda\tau_A^{h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right]\right.\\
&\quad \left. -\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right) d\mu(x)\\
&= \lim_{h\downarrow 0}\frac{1}{h}\int_E \left(\mathbb{E}_x\left[\mathbb{E}_{X(h)}\left[f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right]\right.\\
&\quad \left. -\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}} f\left(X\left(\tau_A^{0,h}\right)\right)\right]\right) d\mu(x)\\
&= \lim_{h\downarrow 0}\frac{1}{h}\int_E \left(e^{hL}-I\right) H_A^h f(x)\, d\mu(x). \qquad (10.115)
\end{aligned}
$$

Hence, if μ were an invariant Borel measure, then the expression in (10.115) would vanish. So the expression for $\Lambda_{A,\mu}$ does not automatically lead to an invariant measure. So that there is a problem with the invariance, although (10.114) yields a measure. In order to take care of that problem we will assume that for every $f \in C_b(E)$ there exists a sequence of strictly positive real numbers $(\lambda_n)_{n\in\mathbb{N}}$, which decreases to zero, and is such that $Pf := \mathcal{T}_\beta\text{-}\lim_{n\to\infty} \lambda_n R(\lambda_n) f$ exists for all $f \in C_b(E)$. Notice $R(P) = N(L)$ and that, by the resolvent identity $P^2 = P$, i.e. P is a projection on the zero-space of the operator L. Fix $x_0 \in A^r$. Then, in general, the formula

$$
\begin{aligned}
f &\mapsto \lim_{n\to\infty} \lambda_n\left(\lambda_n I - L\right) H_A\left(\lambda_n\right) R\left(\lambda_n\right)^2 f\left(x_0\right)\\
&= \lim_{n\to\infty}\Big(\lambda_n\left(\lambda_n I - L\right)\left(H_A\left(\lambda_n\right) - I\right) R\left(\lambda_n\right)^2 f\left(x_0\right) + \lambda_n R\left(\lambda_n\right) f\left(x_0\right)\Big)\\
&= \lim_{n\to\infty}\left(L-\lambda_n\right) R_A\left(\lambda_n\right)\left(\lambda_n R\left(\lambda_n\right) f\right)\left(x_0\right) + \lim_{n\to\infty}\lambda_n R\left(\lambda_n\right) f\left(x_0\right)\\
&= LR_A(0+)Pf\left(x_0\right) + Pf\left(x_0\right) \qquad (10.116)
\end{aligned}
$$

does not provide an invariant measure either. Suppose e.g. that $N(L)$ consists of the constant functions. Then by taking $f = \mathbf{1}$ in (10.118) below we have

$$
LR_A(0+)\mathbf{1}(x) + \mathbf{1}(x) = \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} L^h(\lambda) H_A^h(\lambda) L^h(\lambda)^{-1}\mathbf{1}(x)
$$

$$
\begin{aligned}
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} L^h(\lambda)H_A^h(\lambda)L^h(\lambda)^{-1}\mathbf{1}(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} \frac{1}{1-e^{-\lambda h}}\left(I-e^{-\lambda h}e^{hL}\right)H_A^h(\lambda)\mathbf{1}(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} \frac{1}{1-e^{-\lambda h}}\left(\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}}-e^{-\lambda\tau_A^{1,h}}\right]\right)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} \frac{\lambda}{1-e^{-\lambda h}}\frac{\mathbb{E}_x\left[e^{-\lambda\tau_A^{0,h}}-e^{-\lambda\tau_A^{1,h}},\, \tau_A^{1,h}>\tau_A^{0,h}\right]}{\lambda}\\
&= \lim_{h\downarrow 0}\frac{1}{h}\left(\mathbb{E}_x\left[\tau_A^{1,h}-\tau_A^{0,h},\, \tau_A^{1,h}>\tau_A^{0,h}\right]\right)
\end{aligned}
$$

($\mathbb{E}_x\left[\tau_A^{0,h}\right]=0$ for $x\in A$, and $\mathbb{P}_x\left[\tau_A^{1,h}=\tau_A^{0,h}\right]=1$ for $x\in E\backslash A$)

$$
\begin{aligned}
&= \lim_{h\downarrow 0}\frac{1}{h}\mathbf{1}_A(x)\left(\mathbb{E}_x\left[\tau_A^{1,h}\right]\right)\\
&= \lim_{h\downarrow 0}\frac{1}{h}\mathbf{1}_A(x)\left(\mathbb{E}_x\left[h+\tau_A^{h}\circ\vartheta_h\right]\right)
\end{aligned}
$$

(Markov property)

$$
\begin{aligned}
&= \lim_{h\downarrow 0}\frac{1}{h}\mathbf{1}_A(x)e^{hL}E_{(\cdot)}\left[\tau_A^{0,h}\right](x)+1\\
&= \mathbf{1}_A(x)LE_{(\cdot)}\left[\tau_A\right](x)+1 \qquad (10.117)
\end{aligned}
$$

where $\tau_A=\inf\left\{\tau_A^{0,h}: h>0\right\}$. If possible choose $x_0\in A$ in such a way that $LE_{(\cdot)}\left[\tau_A\right](x_0)=\infty$. Then $LR_A(0+)\mathbf{1}(x_0)+\mathbf{1}(x_0)=\infty$. It follows that for $f\in C_b(E)$, $f\geqslant 0$, the expression $LR_A(0+)Pf(x_0)+Pf(x_0)$ is either ∞, in case $Pf\neq 0$, or 0, in case $Pf=0$. Observe that, under the hypothesis "the space $N(L)$ consists of the constant functions", Pf is a constant $\geqslant 0$.

The reader is cautioned that the symbol $LR_A(0+)f(x)+f(x)$ is a shorthand notation for the following limit:

$$
\begin{aligned}
LR_A(0+)f(x)+f(x) &= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} L^h(\lambda)R_A^h(\lambda)f(x)+f(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} L^h(\lambda)\left(H_A^h(\lambda)-I\right)L^h(\lambda)^{-1}f(x)+f(x)\\
&= \lim_{h\downarrow 0}\lim_{\lambda\downarrow 0} L^h(\lambda)H_A^h(\lambda)L^h(\lambda)^{-1}f(x). \qquad (10.118)
\end{aligned}
$$

The symbols $L^h(\lambda)$, $R_A^h(\lambda)$, and $H_A^h(\lambda)$ are explained in (10.105). Since $PLf=0$ it follows that the "measure" determined by (10.116) is L-invariant, but not necessarily σ-finite; the expression in (10.118) is either 0 or ∞. In case the measure π is σ-finite, then there exist functions $f \in C_b(E)$, $f \geqslant 0$, such that $0 < \int f \, d\pi < \infty$.

Suppose that for every sequence $(f_n : n \in \mathbb{N})$ which decreases pointwise to zero the sequence $(\sup_{0<\lambda<1} \lambda R(\lambda) f_n : n \in \mathbb{N})$ decreases to the zero-function uniformly on compact subsets. Then the family $\{\lambda R(\lambda) : 0 < \lambda < 1\}$ is $\mathcal{T}_\beta$-equi-continuous, and $\mathcal{T}_\beta\text{-}\lim_{\lambda \downarrow 0} \lambda R(\lambda) f = Pf$ exists for all $f \in C_b(E)$, provided that the vector sum $R(L) + N(L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$. Some of the formulas we need are the following ones:

$$\begin{aligned}
&L^h(\lambda) R_A^h(\lambda) f(x) + f(x) \\
&= \mathbb{E}_x \left[\sum_{k=0}^{h^{-1}\tau_A^{1,h} - 1} e^{-\lambda k h} f\left(X(kh)\right) \right] \cdot \mathbb{P}_x\left[\tau_A^{0,h} = 0\right] \\
&= e^{-\lambda h} e^{hL} \mathbb{E}_{(\cdot)} \left[\sum_{k=0}^{h^{-1}\tau_A^{0,h} - 1} e^{-\lambda k h} f\left(X(kh)\right), \, \tau_A^{0,h} \geqslant h \right](x) \cdot \mathbb{P}_x\left[\tau_A^{0,h} = 0\right] \\
&\qquad + f(x) \mathbb{P}_x\left[\tau_A^{0,h} = 0\right],
\end{aligned} \tag{10.119}$$

$$\left(H_A^h(\lambda) - I\right) L^h(\lambda)^{-1} = R_A^h(\lambda), \tag{10.120}$$

$$- R_A^h(\lambda) L^h(\lambda) = I - e^{-\lambda h} e^{hL} H_A^h(\lambda), \quad H_A^h(\lambda) = \left(H_A^h(\lambda)\right)^2, \quad \text{and} \tag{10.121}$$

$$\lim_{h\downarrow 0} L^h R_A^h(0) L^h f\left(x_0\right) + L^h f\left(x_0\right) = \lim_{h \downarrow 0} L^h H_A^h(0) f\left(x_0\right) = 0, \quad x_0 \in E \backslash A^r. \tag{10.122}$$

We also write $P_A^h(\lambda) = L^h(\lambda) R_A^h(\lambda) + I$. Then

$$P_A^h(\lambda) L^h(\lambda) = L^h(\lambda) \left(R_A(\lambda) L^h(\lambda) + I\right) = L^h(\lambda) H_A^h(\lambda). \tag{10.123}$$

Observe that the equalities in (10.119) are proved in Proposition 10.3: see (10.109). Notice that (10.106) is equivalent to (10.120), and that the equalities in (10.107) prove this equality. The second equality in (10.123) is a consequence of (10.120). Put

$$P_A(0) f = L R_A(0+) f + f = \lim_{h\downarrow 0} \lim_{\lambda \downarrow 0} L^h(\lambda) R_A^h(\lambda) f + f.$$

Then from (10.123) we infer informally that $P_A(0) L f = L H_A(0) f$, $f \in D(L)$. More precisely, for $f \in D(L)$, and $\lambda > 0$ we have

$$\lambda R(\lambda) P_A(0) L f = \lambda L R(\lambda) H_A(0) f = \lambda \left(\lambda R(\lambda) - I\right) H_A(0) f. \tag{10.124}$$

The expression in (10.124) is uniformly bounded in $\lambda > 0$, and converges uniformly to zero when $\lambda \downarrow 0$. Some other ideas will be proposed next. Let

$\mu \geqslant 0$ be any positive measure on $\mathcal{E}$. Then we define the measure π via the functional:

$$f \mapsto \lim_{\lambda \downarrow 0} \lambda \int R(\lambda) P_A(0) f \, d\mu, \quad f \geqslant 0, \; f \in C_b(E). \tag{10.125}$$

Then for $f \in D(L)$ and μ a bounded Borel measure we have

$$\begin{aligned} \lim_{\lambda \downarrow 0} \lambda \int R(\lambda) P_A(0) L f \, d\mu &= \lim_{\lambda \downarrow 0} \lambda \int R(\lambda) L H_A(0) f \, d\mu \\ &= \lim_{\lambda \downarrow 0} \lambda \int \left(\lambda R(\lambda) - I\right) H_A(0) f \, d\mu = 0. \end{aligned} \tag{10.126}$$

If μ is a probability measure which is concentrated on $A^r \subset A$, then the expression in (10.125) can be employed to define a non-trivial invariant measure π. So that

$$\int f \, d\pi = \lim_{\lambda \downarrow 0} \lambda \int R(\lambda) P_A(0) f \, d\mu.$$

The invariance follows from (10.126). The existence follows from the assumption that the subspace $R(L) + N(L)$ is $\mathcal{T}_\beta$-dense in $C_b(E)$. The non-triviality follows from the fact that $\int \mathbf{1} d\pi \geqslant \int \mathbf{1} d\mu = 1$: compare with (10.117). The σ-finiteness follows from the assumption that the subset A is recurrent, i.e. $\mathbb{P}_x\left[\tau_A < \infty\right] = 1$ for all $x \in E$ together with (10.118). Suppose $x \in A^r$. Then the limits in (10.118) are in fact suprema, provided the numbers h are taken of the form $2^{-n}h'$, $h' > 0$ fixed, and $n \to \infty$. Moreover, the expression in (10.118) vanishes for $x \in A$. In addition, we need the fact that

$$\tau_A = \inf_{h'>0} \lim_{n\to\infty} \tau_A^{1,2^{-n}h'} = \inf_{h'>0} \inf_{n\in\mathbb{N}} \tau_A^{1,2^{-n}h'} = \inf\left\{s > 0 : \; X(s) \in A\right\}. \tag{10.127}$$

If A is an open subset, then in (10.127) we may fix h'; e.g. $h' = 1$ will do. As throughout this book we assume that the paths are $\mathbb{P}_x$-almost surely right-continuous. In order to finish the arguments we need Choquet's capacity theorem, which states that for $x \in A^r$ the stopping time τ_A can be approximated from above by hitting times of compact subsets K of A, and form below by hitting times of open subsets:

$$\inf_{K\subset A,\, K \text{ compact}} \tau_K = \tau_A = \sup_{U\supset A,\, U \text{ open}} \tau_U, \quad \mathbb{P}_\mu\text{-almost surely}. \tag{10.128}$$

For more details see §4.5. In particular, see the proof of Theorem 4.6; the equality in (4.248) is quite relevant.

In the remaining part of this subsection the operators L and L_A have to be interpreted in the pointwise sense. For $x \in E$ we have

$$Lf(x) = \lim_{t\downarrow 0} \frac{e^{tL} f(x) - f(x)}{t} = \lim_{t\downarrow 0} \frac{1}{t} \left(\mathbb{E}_x\left[f(X(t)) - f(X(0))\right]\right),$$

and

$$L_A f(x) = \lim_{t\downarrow 0} \frac{e^{tL_A} f(x) - f(x)\mathbf{1}_{E\setminus A^r}(x)}{t}$$
$$= \lim_{t\downarrow 0} \frac{1}{t} \left(\mathbb{E}_x \left[f(X(t)),\, \tau_A > t\right] - \mathbb{E}_x \left[f(X(0)),\, \tau_A > 0\right]\right). \quad (10.129)$$

As a consequence $L_A f(x) = 0$ for $x \in A^r$.

Proposition 10.4. *For $\lambda > 0$ and $x \in E$ the equality*

$$(\lambda I - L)\, H_A(\lambda) f(x) = \mathbf{1}_{A^r}(x)\, (\lambda I - L)\, H_A(\lambda) f(x) \quad (10.130)$$

holds for $f \in C_b(E)$ with the property that $H_A(\lambda)f$ belongs to the pointwise domain of L. Moreover, the function $R_A(\lambda)f$ belongs to the (pointwise) domain of L if and only if the same is true for the function $H_A(\lambda)R(\lambda)f$.

Proof. On $E\setminus A^r$ the equality in (10.130) follows from Lemma 10.1 equality (10.72): see Proposition 10.11 below as well. More precisely, for $x \in E\setminus A^r$, $\lambda > 0$ and $h > 0$ we have

$$\left(I - e^{-\lambda h} e^{hL}\right) H_A(\lambda) f(x)$$
$$= \mathbb{E}_x \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right)\right] - e^{-\lambda h} \mathbb{E}_x \left[\mathbb{E}_{X(h)} \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right)\right]\right]$$

(Markov property)

$$= \mathbb{E}_x \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right)\right] - \mathbb{E}_x \left[e^{-\lambda h - \lambda \tau_A \circ \vartheta_h} f\left(X\left(h + \tau_A \circ \vartheta_h\right)\right)\right]$$

(on the event $\{\tau_A > h\}$ the equality $h + \tau_A \circ \vartheta_h = \tau_A$ holds $\mathbb{P}_x$-almost surely)

$$= \mathbb{E}_x \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right) - e^{-\lambda h - \lambda \tau_A \circ \vartheta_h} f\left(X\left(h + \tau_A \circ \vartheta_h\right)\right),\, \tau_A \leqslant h\right]. \quad (10.131)$$

The equality in (10.130) now follows from Lemma 10.1 equality (10.72).

From the strong Markov property the Dynkin's formula follows:

$$R(\lambda) f(x) - R_A(\lambda) f(x) = H_A(\lambda) R(\lambda) f(x), \quad f \in C_b(E), \quad x \in E, \quad (10.132)$$

or equivalently,

$$f = (\lambda I - L)\left(R_A(\lambda) f + H_A(\lambda) R(\lambda) f\right). \quad (10.133)$$

Let $f \in C_b(E)$. Hence, from (10.133) it follows that the function $R_A(\lambda)f$ belongs to the (pointwise) domain of L if and only if the same is true for the function $H_A(\lambda)R(\lambda)f$.

This completes the proof of Proposition 10.4. □

If $A \in \mathcal{E}$, then τ_A denotes its first hitting time: $\tau_A = \inf\{s > 0 : X(s) \in A\}$.

Proposition 10.5. *Suppose that the Borel subset A is such that it possesses the almost separation property as defined in Definition 9.2 with $D = D(L)$. The following identity holds for all $\lambda > 0$ and $f \in C_b(E)$:*

$$R(\lambda)f - R_A(\lambda)f = R(\lambda)\left(L - L_A\right)R_A(\lambda) + R(\lambda)\left(\mathbf{1}_{A^r}f\right). \tag{10.134}$$

Let $f \in C_b(E)$ and $\lambda > 0$ be such that the function

$$x \mapsto R_A(\lambda)f(x) = \left(\lambda I - L_A\right)^{-1}f(x) = \mathbb{E}_x\left[\int_0^{\tau_A} e^{-\lambda s} f\left(X\left(s\right)\, ds\right)\right] \tag{10.135}$$

belongs to the pointwise domain of the operator L. Then the following equality holds:

$$\begin{aligned}&\left(L - L_A\right)\mathbb{E}_{(\cdot)}\left[\int_0^{\tau_A} e^{-\lambda s} f\left(X(s)\right)\, ds\right](x)\\ &= \left(\lambda I - L\right)\mathbb{E}_{(\cdot)}\left[e^{-\lambda\tau_A}\left(\lambda I - L\right)^{-1} f\left(X\left(\tau_A\right)\right)\right](x) - \mathbf{1}_{A^r}(x)f(x).\end{aligned} \tag{10.136}$$

Proof. The equality in (10.136) is just a rewriting of (10.134). The equality in (10.134) can be obtained by noticing that

$$\left(\lambda I - L_A\right)R_A(\lambda)f = \mathbf{1}_{E\backslash A^r}f. \tag{10.137}$$

The equality in (10.137) follows from the following identities:

$$\begin{aligned}&R_A(\lambda)f(x) - e^{-\lambda\delta}e^{\delta L_A}R_A(\lambda)f(x)\\ &= \int_0^{\infty} e^{-\lambda s}e^{sL_A}f(x)\,ds - \int_\delta^{\infty} e^{-\lambda s}e^{sL_A}f(x)\,ds = \int_0^{\delta} e^{-\lambda s}e^{sL_A}f(x)\,ds.\end{aligned} \tag{10.138}$$

After dividing by $\delta > 0$ and letting δ tend to 0 we see that (10.137) follows. The equality in (10.134) then follows from $L - L_A = \left(\lambda I - L_A\right) - \left(\lambda I - L\right)$ together with Dynkin's formula (10.132).

This completes the proof of Proposition 10.5. □

10.2.3 *Auxiliary results*

Theorem 10.12 below yields the existence of an invariant σ-finite Borel measure provided that there exists a compact recurrent subset A. Originally it was assumed that the set A^r, i.e. the set of regular points of A coincides with A. Instead of the operator Q_A, as defined in (10.210), the operator $H_A : C_b([0,\infty) \times E) \to C_b(E)$ defined by

$$H_A g(x) = \mathbb{E}_x \left[g\left(\tau_A, X\left(\tau_A\right)\right)\right] = \mathbb{E}_x \left[g\left(\tau_A, X\left(\tau_A\right)\right), \tau_A < \infty \right], \quad (10.139)$$

was employed. Only in case $A^r = A$ one can be sure that the function $H_A g$ is continuous whenever g is a bounded continuous function on $[0,\infty) \times E$. For some of the consequences of the assumption $A = A^r$ see Lemma 10.12 below: notice that the equality in (10.139) is the same as (10.280) in Lemma 10.12. The operator Q_A assigns bounded continuous functions to bounded continuous functions automatically. Recall that a Markov process with transition function $P(t,x,B)$ is strong Feller whenever every function $x \mapsto P(t,x,B)$, $(t,B) \in (0,\infty) \times E$, is continuous: see Definitions 2.5 and 2.16 as well.

The following result reduces the existence of an invariant measure for the Markov process given by (9.14) to that of a Markov chain. In fact our approach is inspired by results due to Azema, Kaplan-Duflo and Revuz [Azéma *et al.* (1966, 1965/1966, 1967)]. Basically, the process $t \mapsto X(t)$ is replaced by the chain $(n,\omega,\lambda) \mapsto X(T_n(\lambda),\omega)$, $(n,\omega,\lambda) \in \mathbb{N} \times \Omega \times \Lambda$, where the process $(n,\lambda) \mapsto T_n(\lambda)$, $(n,\lambda) \in \mathbb{N} \times \Lambda$, are the jump times of an independent Poisson process of intensity $\lambda_0 > 0$

$$\left\{ \left(\Lambda, \mathcal{G}, \pi_t\right)_{t \geqslant 0}, \left(N(t), t \geqslant 0\right), \left(\vartheta_t^P : t \geqslant 0\right), [0,\infty) \right\}. \quad (10.140)$$

This means that $T_n = \inf\{t > 0 : N(t) \geqslant n\}$, $n \in \mathbb{N}$. The process $n \mapsto T_n$ can be realized as a random walk: $T_n = \sum_{k=1}^n Z_k$. Here the sequence $(Z_k : k \in \mathbb{N}, k \geqslant 1)$ is a sequence of independent variables each exponentially distributed with parameter α_0. This technique is also described in Chapter 20 of [Meyn and Tweedie (1993b)] second version. Employing probabilistic techniques, e.g. Poisson variables, in order to approximate semigroups and represent resolvent operators also occurs in [Chung (1962)].

Lemma 10.5. *The process given by*

$$\left\{ \left(\Omega \times \Lambda, \mathcal{F} \otimes \mathcal{G}, \mathbb{P}_x \otimes \pi_0\right), \left(X\left(T_n(\lambda),\omega\right), n \in \mathbb{N}\right), \left(\vartheta_n^P(\lambda), n \in \mathbb{N}\right), (E,\mathcal{E}) \right\} \quad (10.141)$$

is a Markov chain. Its transition kernel is given by

$$\mathbb{P}_x \otimes \pi_0 \left[X\left(T_1\right) \in B \right] = \alpha_0 \int_0^\infty e^{-\alpha_0 t} P(t,x,B)\, dt \quad (10.142)$$

provided that T_1 is exponentially distributed with parameter α_0.

The $\mathbb{P}_x \otimes \pi_0$-distribution of the state variable $X(T_n)$ can be expressed in terms of the $\mathbb{P}_x$-distribution of the process $X(t)$:

$$\begin{aligned}\mathbb{P}_x \otimes \pi_0 \left[X(T_n) \in B\right] &= \alpha_0 \int_0^\infty \frac{(\alpha_0 t)^{n-1}}{(n-1)!} e^{-\alpha_0 t} P(t,x,B)\, dt \\ &= \alpha_0 \int_0^\infty \frac{(\alpha_0 t)^{n-1}}{(n-1)!} e^{-\alpha_0 t} \mathbb{P}_x \left[X(t) \in B\right] dt = \left(\alpha_0 R(\alpha_0)\right)^n \mathbf{1}_B(x). \quad (10.143)\end{aligned}$$

In (10.141) we have $X(t) \circ \vartheta_m^P(\omega,\lambda) = X(t + T_m(\lambda), \omega)$, $n, m \in \mathbb{N}$, $(\omega,\lambda) \in \Omega \times \Lambda$. The time translation operators $\vartheta_m^P(\lambda)$ satisfy

$$X(t) \circ \vartheta_m^P(\omega,\lambda) = X(t + T_m(\lambda), \omega), \quad (\omega,\lambda) \in \Omega \times \Lambda.$$

Relative to π_0 the variables $T_n - T_m$ and $T_{n-m} - T_0 = T_{n-m}$, $n > m$, have the same distributions, and the measure π_t is the measure π_0 translated over time t, i.e.

$$\begin{aligned}&\int_\Lambda F(T_1 - T_0, \dots, T_n - T_{n-1})\, d\pi_t \\ &= \int_\Lambda F(T_1 - T_0 + t, \dots, T_n - T_{n-1} + t)\, d\pi_0,\end{aligned}$$

where $F : [0,\infty)^n \to \mathbb{R}$ is any bounded Borel measurable function. It follows that the probability measures $\left(\pi_{T_m(\lambda)} : m \in \mathbb{N},\, \lambda \in \Lambda\right)$ satisfy $(n > m)$:

$$\begin{aligned}\int_\Lambda f(T_n - T_m)\, d\pi_{T_m}(\lambda) &= \int_\Lambda f\left((T_n - T_m)(\lambda') + T_m(\lambda)\right) d\pi_0(\lambda') \\ &= \int_\Lambda f\left(T_{n-m}(\lambda') + T_m(\lambda)\right) d\pi_0(\lambda'). \quad (10.144)\end{aligned}$$

Proof. [Proof of Lemma 10.5.] For brevity we write $\widetilde{\mathbb{P}}_x = \mathbb{P}_x \otimes \pi_0$. Put $Y_k = X(T_k)$, $k \in \mathbb{N}$, and let $f_j : E \to \mathbb{R}$, $1 \leqslant j \leqslant n+1$, be bounded Borel measurable functions. In order to show that the process in (10.141) is a Markov chain we have to prove the equality:

$$\widetilde{\mathbb{E}}_x \left[\prod_{j=1}^{n+1} f_j(Y_j)\right] = \widetilde{\mathbb{E}}_x \left[\prod_{j=1}^{n} f_j(Y_j)\, \widetilde{\mathbb{E}}_{Y_n}\left[f_{n+1}(Y_1)\right]\right]. \qquad (10.145)$$

Employing Fubini's theorem shows that the right-hand side of (10.145) can be rewritten as

$$\begin{aligned}&\widetilde{\mathbb{E}}_x \left[\prod_{j=1}^{n} f_j(Y_j)\, \widetilde{\mathbb{E}}_{Y_n}\left[f_{n+1}(Y_1)\right]\right] \\ &= \int_\Lambda d\pi_0(\lambda) \int_\Lambda d\pi_0(\lambda') \mathbb{E}_x \left[\prod_{j=1}^{n} f_j(X(T_j(\lambda))) \mathbb{E}_{X(T_n(\lambda))}\left[f_{n+1}(X(T_1(\lambda')))\right]\right]\end{aligned}$$

(the process in (9.14) is a Markov process)

$$= \int_\Lambda d\pi_0(\lambda) \int_\Lambda d\pi_0\left(\lambda'\right) \mathbb{E}_x \left[\prod_{j=1}^{n} f_j\left(X\left(T_j(\lambda)\right)\right) f_{n+1}\left(X\left(T_n(\lambda) + T_1\left(\lambda'\right)\right)\right) \right]$$

(the variables $T_{n+1} - T_n$ and T_1 have the same π_0-distribution)

$$= \int_\Lambda d\pi_0(\lambda) \int_\Lambda d\pi_0\left(\lambda'\right)$$
$$\mathbb{E}_x \left[\prod_{j=1}^{n} f_j\left(X\left(T_j(\lambda)\right)\right) f_{n+1}\left(X\left(T_n(\lambda) + T_{n+1}\left(\lambda'\right) - T_n\left(\lambda'\right)\right)\right) \right]$$

(the variables $T_{n+1} - T_n$ and T_j, $1 \leqslant j \leqslant n$, are π_0-independent)

$$= \int_\Lambda d\pi_0(\lambda) \mathbb{E}_x \left[\prod_{j=1}^{n} f_j\left(X\left(T_j(\lambda)\right)\right) f_{n+1}\left(X\left(T_n(\lambda) + T_{n+1}(\lambda) - T_n(\lambda)\right)\right) \right]$$
$$= \int_\Lambda d\pi_0(\lambda) \mathbb{E}_x \left[\prod_{j=1}^{n} f_j\left(X\left(T_j(\lambda)\right)\right) f_{n+1}\left(X\left(T_{n+1}(\lambda)\right)\right) \right]$$
$$= \widetilde{\mathbb{E}}_x \left[\prod_{j=1}^{n+1} f_j\left(Y_j\right) \right]. \tag{10.146}$$

The equality in (10.146) proves the Markov chain property of the process in (10.141).

Next we will show equality (10.142). Therefore we write

$$\mathbb{P}_x \otimes \pi_0 \left[X\left(T_1\right) \in B\right] = \int_\Lambda P\left(T_1(\lambda), x, B\right) d\pi_0(\lambda) = \alpha_0 \int_0^\infty e^{-\alpha_0 t} P\left(t, x, B\right) dt. \tag{10.147}$$

In the final step in (10.147) we used the exponential distribution of the variable T_1 with parameter $\alpha_0 > 0$.

The equalities in (10.146) and (10.147) complete the proof of Lemma 10.5. □

Lemma 10.6. *Put*

$$N(t, \lambda) = \sum_{n=0}^{\infty} n \mathbf{1}_{[T_n(\lambda), T_{n+1}(\lambda))}(t) = \#\left\{k \geqslant 1 : T_k(\lambda) \leqslant t\right\}$$
$$= \max\left\{k \geqslant 0 : T_k(\lambda) \leqslant t\right\}. \tag{10.148}$$

Suppose that the variables $T_{k+1} - T_k$, $k \in \mathbb{N}$, are π_0-independent and identically exponentially distributed random variables with parameter α_0 attaining their values in $[0, \infty)$. Then with respect to π_0 the process $N(t)$, $t \geqslant 0$, is a Poisson process of intensity α_0 and with jumping times T_n.

Proof. Fix $k \in \mathbb{N}$ and $t > 0$. Then we have:

$$\begin{aligned}\pi_0 [N(t) = k] &= \pi_0 [T_k \leqslant t < T_{k+1}] = \pi_0 [0 \leqslant t - T_k < T_{k+1} - T_k] \\ &= \int_\Lambda d\pi_0 \int_{t-T_k}^\infty \alpha_0 e^{-\alpha_0 s} ds\, \mathbf{1}_{\{T_k < t\}} \\ &= \int_\Lambda d\pi_0 e^{-\alpha_0(t-T_k)} \mathbf{1}_{\{T_k \leqslant t\}} = \frac{(\alpha_0 t)^k}{k!} e^{-\alpha_0 t}. \end{aligned} \tag{10.149}$$

By writing $T_k = \sum_{j=1}^k (T_j - T_{j-1})$, and using the independence of the increments $T_j - T_{j-1}$, $1 \leqslant j \leqslant k$, the ultimate equality in (10.149) can be proved by induction with respect to k and using the exponential distribution of $T_j - T_{j-1}$. This completes the proof of Lemma 10.6 □

Lemma 10.7. *Let the process $(T_k : k \in \mathbb{N})$ be the process of jump times of a Poisson process*

$$\{(\Lambda, \mathcal{G}, \pi_n)_{n\in\mathbb{N}} : (N(t), t \geqslant 0), (\vartheta_t : t \geqslant 0), \mathbb{N}\}.$$

Let the initial measure π_0 be exponentially distributed with parameter $\alpha_0 > 0$. Let B be a Borel subset of $[0, \infty)$ of Lebesgue measure ∞. Then

$$\pi_0 \left[\bigcap_{n\in\mathbb{N}} \bigcup_{m\geqslant n} \{T_m \in B\} \right] = 1. \tag{10.150}$$

Proof. Put $B_n = B \bigcap (n, n+1]$, and $E_n = \bigcup_{k=1}^\infty \{T_k \in B_n\}$. Let $\mathcal{H}_t$ be the σ-field generated by $(N(s) : s \leqslant t)$. Since the event

$$E_n = \{\text{there is a jump in } B_n\}$$

contains the event

$$\{n + T_1 \circ \vartheta_n \in B_n\} = \{\text{the first jump after } n \text{ occurs in } B_n\},$$

we have

$$\pi_0 \left[E_n \mid \mathcal{H}_n \right] \geqslant \pi_0 \left[n + T_1 \circ \vartheta_n \in B_n \mid \mathcal{H}_n \right]$$

(Markov property of the process $N(t)$)

$$= \pi_{N(n)} [n + T_1 \in B_n] = \pi_0 [B_n - n], \tag{10.151}$$

where in the ultimate equality in (10.151) we used that fact that the distribution of the first jumping time of a Poisson process does not depend on the initial position. From (10.151) it follows that

$$\pi_0 \left[E_n \mid \mathcal{H}_n \right] \geqslant \pi_0 [T_1 \in B_n - n] = \alpha_0 \int_{B_n - n} e^{-\alpha_0 t} dt$$

$$\geqslant \alpha_0 e^{-\alpha_0} \int_{B_n - n} \mathbf{1} dt = \alpha_0 e^{-\alpha_0} m(B_n) \tag{10.152}$$

where $m(B_n)$ is the Lebesgue measure of B_n. Since the variables T_k are stopping times relative to the process $t \mapsto N(t)$, the events E_n are $\mathcal{H}_{n+1}$-measurable, and hence an application of the generalized Borel-Cantelli theorem yields $\pi_0 \left[\bigcap_n \bigcup_{m \geqslant n} E_n\right] = 1$. Since $\bigcap_n \bigcup_{m \geqslant n} E_n \subset \bigcap_n \bigcup_{m \geqslant n} \{T_n \in B\}$ the equality in (10.150) follows. For a precise formulation of the generalized Borel-Cantelli lemma see e.g. [Shiryayev (1984)] Corollary VII 5.2 or the equality in (9.49) in Theorem 9.3.

This completes the proof of Lemma 10.7. □

The following theorem appears as Theorem 1 in [Kaspi and Mandelbaum (1994)]. For the notion of strong Feller property see Definitions 2.5 and 2.16.

Theorem 10.7. *Let the strong Markov process be as in (9.14) of Theorem 9.2. Suppose that this time-homogeneous Markov process on the Polish space E has transition probability function $P(t, x, \cdot)$, $t \geqslant 0$, $x \in E$, which is conservative in the sense that $P(t, x, E) = 1$ for all $t \geqslant 0$ and $x \in E$. In addition, assume that the process $X(t)$ is strong Feller. Then the following assertions are equivalent:*

(a) There exists a non-zero σ-finite Borel measure μ such that for all $B \in \mathcal{E}$, $\mu(B) > 0$ implies $\mathbb{P}_x\left[\int_0^\infty \mathbf{1}_B(X(t))\, dt = \infty\right] = 1$ for all $x \in E$.

(b) There exists a non-zero σ-finite Borel measure ν such that for all $B \in \mathcal{E}$, $\nu(B) > 0$ implies $\mathbb{P}_x[\tau_B < \infty] = 1$ for all $x \in E$.

Here $\tau_B = \inf\{t > 0 : X(t) \in B\}$ is the first hitting time of B. Moreover, $\{\tau_B < \infty\} = \bigcup_{t>0} \{X(t) \in B\}$. The measure μ in assertion (a) could be called a Harris recurrence measure, and the measure ν in assertion (b) could be called a recurrence measure. In the proof of Theorem 10.7 we need Lemma 10.8 below.

Remark 10.5. In (10.171) below we will see that the measure

$$\mu(B) = \int_E \mathbb{E}_x \left[\int_0^\infty e^{-t} \mathbf{1}_B(X(t))\, dt\right] d\nu(x) \tag{10.153}$$

conforms to assertion (a), provided ν conforms to (b). If μ is given by $\mu(B) = P(t_0, x_0, B)$, $B \in \mathcal{E}$, for some fixed $(t_0, x_0) \in (0, \infty) \times E$. Then ν is given by $\nu(B) = e^{t_0} \int_{t_0}^\infty e^{-s} P(s, x_0, B)\, ds$.

Remark 10.6. If all measures $B \mapsto P(s,x,B)$, $B \in \mathcal{E}$, $(s,x) \in (0,\infty) \times E$, are equivalent, and if any (all) of these measures serves as a recurrence measure, then for ν we may also choose one of these transition probabilities. Fix $(t_0,x_0) \in (0,\infty) \times E$. In fact, if all such measures are equivalent, and $\nu(B) = P(t_0,x_0,B)$, $B \in \mathcal{E}$, then the measure μ in (10.153) is given by

$$\begin{aligned}\mu(B) &= \int_0^\infty e^{-t} \int_E \mathbb{E}_y\left[\mathbf{1}_B(X(t))\right] P(t_0,x_0,dy)\, dt \\ &= \int_0^\infty e^{-t} \mathbb{E}_{x_0}\left[\mathbb{E}_{X(t_0)}\left[\mathbf{1}_B(X(t))\right]\right] dt\end{aligned}$$

(Markov property)

$$\begin{aligned}&= \int_0^\infty e^{-t} \mathbb{E}_{x_0}\left[\mathbf{1}_B(X(t+t_0))\right] dt \\ &= e^{t_0} \int_{t_0}^\infty e^{-t} \mathbb{E}_{x_0}\left[\mathbf{1}_B(X(t))\right] dt. \qquad (10.154)\end{aligned}$$

From (10.154) it easily follows that μ is also equivalent to the measure $B \mapsto P(t_0,x_0,B)$, $B \in \mathcal{E}$. Therefore, let $B \in \mathcal{E}$ be such that $\mu(B) = 0$. Then there exists $(t,x) \in (0,\infty) \times E$ such that $P(t,x,B) = 0$. By equivalence we see $P(t_0,x_0,B) = 0$.

Let $\alpha \geqslant 0$. We also have a need for α-excessive functions.

Definition 10.3. A non-negative function $f : E \to [0,\infty)$ is called α-excessive if $t \mapsto \mathbb{E}_x\left[e^{-\alpha t} f(X(t))\right]$ increases to $f(x)$ for all $x \in E$ whenever $t \downarrow 0$. If $\alpha = 0$, then f is called excessive.

Let $f : E \to [0,\infty)$ be an α-excessive function, and let $0 \leqslant t_1 < t_2 < \infty$. The (in-)equalities

$$\begin{aligned}&\mathbb{E}_x\left[e^{-\alpha t_2} f(X(t_2)) \,\middle|\, \mathcal{F}_{t_1}\right] - e^{-\alpha t_1} f(X(t_1)) \\ &= e^{-\alpha t_2} \mathbb{E}_{X(t_1)}\left[f(X(t_2 - t_1))\right] - e^{-\alpha t_1} f(X(t_1)) \\ &= e^{-\alpha t_1}\left(e^{-\alpha(t_2-t_1)} \mathbb{E}_{X(t_1)}\left[f(X(t_2-t_1))\right] - f(X(t_1))\right) \leqslant 0 \qquad (10.155)\end{aligned}$$

show that the process $t \mapsto e^{-\alpha t} f(X(t))$ is a $\mathbb{P}_x$-super-martingale relative to the filtration $(\mathcal{F}_t)_{t \geqslant 0}$. From the (super-)martingale convergence theorem it then follows that $\lim_{t\to\infty} e^{-\alpha t} f(X(t))$ exists $\mathbb{P}_x$-almost surely for all $x \in E$. Let $\tau : \Omega \to [0,\infty]$ be any stopping time such that $\mathbb{P}_x[\tau < \infty] = 1$. Then it also follows that

$$\mathbb{E}_x\left[\lim_{t\to\infty} e^{-\alpha t} f(X(t))\right] \leqslant \lim_{t\to\infty} \mathbb{E}_x\left[e^{-\alpha t} f(X(t))\right]$$

$$\begin{aligned}
&= \lim_{t\to\infty} \mathbb{E}_x \left[e^{-\alpha t} f\left(X(t)\right), \tau \leqslant t \right] \\
&= \lim_{t\to\infty} \mathbb{E}_x \left[\mathbb{E}_x \left[e^{-\alpha t} f\left(X(t)\right) \bigm| \mathcal{F}_{\tau\wedge t} \right], \tau \leqslant t \right]
\end{aligned}$$

(Doob's optional sampling theorem for super-martinagales)

$$\begin{aligned}
&\leqslant \lim_{t\to\infty} \mathbb{E}_x \left[e^{-\alpha(\tau\wedge t)} f\left(X(\tau\wedge t)\right), \tau \leqslant t \right] \\
&= \mathbb{E}_x \left[e^{-\alpha\tau} f\left(X(\tau)\right), \tau < \infty \right]. \qquad (10.156)
\end{aligned}$$

If τ_A denotes the first hitting time of $A \in \mathcal{E}$, and $\alpha > 0$, then the function $x \mapsto \mathbb{E}_x\left[e^{-\alpha\tau_A}\right]$ is α-excessive, and the function $x \mapsto \mathbb{P}_x\left[\tau_A < \infty\right]$ is excessive. These assertions follow from the Markov property, and the fact that $t + \tau_A \circ \vartheta_t$ decreases to τ_A when $t \downarrow 0$. Recall that $\tau_A = \inf\{s > 0 : X(s) \in A\}$ which is the first hitting time of A.

Lemma 10.8. *Let ν be a recurrence measure for the Markov process in (9.14) of Theorem 9.2, and let $L \geqslant 0$ be an increasing right-continuous additive process on Ω such that $L(0+) = L(0) = 0$. Then either $L(\infty) := \lim_{t\to\infty} L(t) = \infty$ $\mathbb{P}_x$-almost surely for all $x \in E$, or $L(\infty) = 0$ $\mathbb{P}_\nu$-almost surely. These assertions are mutually exclusive.*

The defining property of an adapted additive process $t \mapsto L(t)$ is the equality $L(s) + L(t) \circ \vartheta_s = L(s+t)$, which should hold $\mathbb{P}_x$-almost surely for all $x \in E$ and for all $s, t \geqslant 0$. For more details on the notion of additive processes see Definition 9.6. For our purpose relevant additive processes are given by $L(t) = \int_0^t \mathbf{1}_B\left(X(s)\right) ds$ with $B \in \mathcal{E}$. Let $t \mapsto L(t)$ be an increasing positive additive process, and fix $\varepsilon > 0$. Suppose that $L(0+) = 0$, and define the stopping time τ_ε by

$$\tau_\varepsilon = \inf\{t > 0 : L(t) > \varepsilon\}. \qquad (10.157)$$

Then the function $x \mapsto \mathbb{P}_x\left[\tau_\varepsilon < \infty\right]$ is excessive. This can be seen as follows. First observe that

$$\begin{aligned}
t + \tau_\varepsilon \circ \vartheta_t &= \inf\{t + s : s > 0, \ L(s) \circ \vartheta_t > \varepsilon\} \\
&= \inf\{t + s : s > 0, \ L(t) + L(s) \circ \vartheta_t > \varepsilon + L(t)\} \\
&= \inf\{t + s : s > 0, \ L(t+s) > \varepsilon + L(t)\} \\
&= \inf\{s > t, \ L(s) > \varepsilon + L(t)\}, \qquad (10.158)
\end{aligned}$$

which decreases to

$$\inf\{s > 0 : L(s) > \varepsilon + L(0+)\} = \inf\{s > 0 : L(s) > \varepsilon\} = \tau_\varepsilon. \qquad (10.159)$$

From (10.158) together with (10.159) and the Markov property it follows that

$$\mathbb{E}_x\left[\mathbb{P}_{X(t)}\left[\tau_\varepsilon<\infty\right]\right]=\mathbb{P}_x\left[t+\tau_\varepsilon\circ\vartheta_t<\infty\right]\uparrow\mathbb{P}_x\left[\tau_\varepsilon<\infty\right] \tag{10.160}$$

when $t\downarrow 0$. From (10.160) we see that the function $x\mapsto\mathbb{P}_x\left[\tau_\varepsilon<\infty\right]$ is excessive.

Proof. [Proof of Lemma 10.8.] Fix $\varepsilon>0$ and define τ_ε as in (10.157). By the right-continuity of the process $s\mapsto L(s)$ we obtain

$$\begin{aligned}\lim_{t\downarrow s}\mathbb{E}_x\left[\mathbb{P}_{X(t)}\left[\tau_\varepsilon<\infty\right]\right]&=\lim_{t\downarrow s}\mathbb{E}_x\left[t+\tau_\varepsilon\circ\vartheta_t<\infty\right]\\&=\mathbb{E}_x\left[s+\tau_\varepsilon\circ\vartheta_s<\infty\right]=\mathbb{E}_x\left[\mathbb{P}_{X(s)}\left[\tau_\varepsilon<\infty\right]\right].\end{aligned} \tag{10.161}$$

From (10.161) it follows that we may, and shall, assume that the supermartingale $t\mapsto\mathbb{P}_{X(t)}\left[\tau_\epsilon<\infty\right]$ is right-continuous. We have the $\mathbb{P}_x$-almost sure equality of events:

$$\{t<\tau_\varepsilon<\infty\}=\{\tau_\varepsilon>t,\,\tau_\varepsilon\circ\vartheta_t<\infty\}. \tag{10.162}$$

Conditioning (10.162) on $\mathcal{F}_t$ and employing the Markov property yields:

$$\begin{aligned}\mathbb{P}_x\left[t<\tau_\varepsilon<\infty\,\middle|\,\mathcal{F}_t\right]&=\mathbf{1}_{\{\tau_\varepsilon>t\}}\mathbb{P}_x\left[\tau_\varepsilon\circ\vartheta_t<\infty\,\middle|\,\mathcal{F}_t\right]\\&=\mathbf{1}_{\{\tau_\varepsilon>t\}}\mathbb{P}_{X(t)}\left[\tau_\varepsilon<\infty\right].\end{aligned} \tag{10.163}$$

Next we let $t\uparrow\infty$ in (10.163) to obtain:

$$0=\mathbf{1}_{\{\tau_\varepsilon=\infty\}}\lim_{t\to\infty}\mathbb{P}_{X(t)}\left[\tau_\varepsilon<\infty\right],\ \mathbb{P}_x\text{-almost surely.} \tag{10.164}$$

Consider the sets

$$F_\varepsilon=\left\{x\in E:\,\mathbb{P}_x\left[\tau_\varepsilon<\infty\right]=0\right\},\quad\text{and}$$
$$G_{\varepsilon,\delta}=\left\{x\in E:\,\mathbb{P}_x\left[\tau_\varepsilon<\infty\right]>\delta\right\},\quad\delta>0.$$

First assume $\nu\left(E\backslash F_\varepsilon\right)>0$. Then $\nu\left(G_{\varepsilon,\delta}\right)>0$ for some $\delta>0$. Since ν is a recurrence measure, it follows that $\limsup_{t\to\infty}\mathbf{1}_{G_{\varepsilon,\delta}}\left(X(t)\right)=1$ $\mathbb{P}_x$-almost surely, and hence $\lim_{t\to\infty}\mathbb{P}_{X(t)}\left[\tau_\varepsilon<\infty\right]\geqslant\delta$ $\mathbb{P}_x$-almost surely. Thus (10.164) implies $\mathbb{P}_x\left[\tau=\infty\right]=0$, which is equivalent to $\mathbb{P}_x\left[\tau<\infty\right]=1$. Consequently,

$$\nu\left(E\backslash F_\varepsilon\right)>0\Longrightarrow\tau_\varepsilon<\infty\ \mathbb{P}_x\text{-almost surely for all }x\in E. \tag{10.165}$$

Next assume that $\nu\left(F_\varepsilon\right)>0$. Let τ_{F_ε} be the (first) hitting time of F_ε: $\tau_{F_\varepsilon}=\inf\left\{s>0:\,X(s)\in F_\varepsilon\right\}$. Then, since ν is a recurrence measure, we have $\mathbb{P}_x\left[\tau_{F_\varepsilon}<\infty\right]=1$. From (10.156) with $\tau=\tau_{F_\varepsilon}$, $\alpha=0$, and $f(x)=\mathbb{P}_x\left[\tau_\varepsilon<\infty\right]$ we see that $\lim_{t\to\infty}\mathbb{P}_{X(t)}\left[\tau_\varepsilon<\infty\right]=0$ $\mathbb{P}_x$-almost surely for all $x\in E$, and so $\tau_\varepsilon=\infty$ $\mathbb{P}_x$-almost surely for all $x\in E$. We repeat the latter conclusion:

$$\nu\left(F_\varepsilon\right)>0\Longrightarrow\tau_\varepsilon=\infty\ \mathbb{P}_x\text{-almost surely for all }x\in E. \tag{10.166}$$

There are two mutually exclusive possibilities:

(i) either there exists $\varepsilon > 0$ such that $\nu(E \backslash F_\varepsilon) > 0$, and (10.165) holds for some $\varepsilon > 0$,

(ii) or for every $\varepsilon > 0$ $\nu(F_\varepsilon) > 0$, and (10.166) holds for all $\varepsilon > 0$.

If (10.165) holds for some $\varepsilon > 0$, then for such $\varepsilon > 0$ the equality $\mathbb{P}_x[\tau_\varepsilon < \infty] = 1$ holds for all $x \in E$. Then we proceed as follows. By induction we introduce the following sequence of stopping times:

$$\begin{aligned}&\eta_0 = 0, \quad \eta_1 = \tau_\varepsilon, \quad \text{and for } n \geqslant 1\\ &\eta_n = \eta_{n-1} + \tau_\varepsilon \circ \vartheta_{\eta_{n-1}} = \inf\{s > \eta_{n-1} : L(s) > \varepsilon + \eta_{n-1}\}. \end{aligned} \tag{10.167}$$

From (10.167) it follows that $\{\eta_n < \infty\} \subset \{L(\infty) > n\varepsilon\}$, and hence for all $n \geqslant 1$ and $x \in E$ we have

$$\mathbb{P}_x[\eta_n < \infty] \leqslant \mathbb{P}_x[L(\infty) > n\varepsilon]. \tag{10.168}$$

In addition, by the strong Markov property we have

$$\begin{aligned}\mathbb{P}_x[\eta_n < \infty] &= \mathbb{P}_x\left[\eta_{n-1} < \infty, \tau_\varepsilon \circ \vartheta_{\eta_{n-1}} < \infty\right]\\ &= \mathbb{E}_x\left[\eta_{n-1} < \infty, \mathbb{P}_{X(\eta_{n-1})}[\tau_\varepsilon < \infty]\right]. \end{aligned} \tag{10.169}$$

Since $\mathbb{P}_y[\tau_\varepsilon < \infty] = 1$ for all $y \in E$ by induction with respect to $n \in \mathbb{N}$ (10.169) yields $\mathbb{P}_x[\eta_n < \infty] = 1$ for all $x \in E$ and all $n \in \mathbb{N}$. From (10.168) we then infer $\mathbb{P}_x[L(\infty) = \infty] = 1$ for all $x \in E$. This is the first alternative in Lemma 10.8. If, on the other hand, (10.166) holds for all $\varepsilon > 0$, then we have

$$\mathbb{P}_\nu[L(\infty) = 0] = \lim_{\varepsilon \downarrow 0} \mathbb{P}_\nu[L(\infty) < \varepsilon] = \lim_{\varepsilon \downarrow 0} \mathbb{P}_\nu[\tau_\varepsilon = \infty] = 1. \tag{10.170}$$

The equality (10.170) yields the second alternative of Lemma 10.8.

Altogether this completes the proof of Lemma 10.8. □

Now we are ready to prove Theorem 10.7.

Proof. [Proof of Theorem 10.7.] The implication (a) $\Longrightarrow$ (b) follows with $\nu = \mu$. Let ν be such that assertion (b) holds with the measure ν. Then we will prove that (a) holds with

$$\mu(B) = \mathbb{E}_\nu\left[\int_0^\infty e^{-t}\mathbf{1}_B(X(t))\,dt\right] = \int_E \mathbb{E}_x\left[\int_0^\infty e^{-t}\mathbf{1}_B(X(t))\,dt\right] d\nu(x). \tag{10.171}$$

Let $B \in \mathcal{E}$ be such that $\mu(B) > 0$, where μ is as in (10.171). Put $L(t) = \int_0^t \mathbf{1}_B(X(s))\,ds$. Then $L(\infty) = 0$ cannot be true $\mathbb{P}_\nu$-almost everywhere. From Lemma 10.8 it follows that $L(\infty) = \infty$ $\mathbb{P}_x$-almost surely for all $x \in E$.

This completes the proof of Theorem 10.7. □

The following theorem is the Markov chain analogue of Theorem 10.7. Its proof can be adapted from the proof of Theorem 10.7, and the required Lemma 10.8 with $L(k) = \sum_{j=1}^{k} \mathbf{1}_{\{X(j)\in B\}}$ where $B \in \mathcal{E}$. The time τ_ε is replaced by $\tau_1 = \inf\{k \geqslant 1 : L(k) \geqslant 1\}$. The equalities in (10.171) are replaced with e.g.

$$\begin{aligned} \mu(B) &= (1-r)\sum_{k=1}^{\infty} r^k \mathbb{E}_\nu\left[\sum_{k=1}^{\infty} \mathbf{1}_B\left(X(k)\right)\right] \\ &= (1-r)\sum_{k=1}^{\infty} r^{k-1}\int_E \mathbb{E}_x\left[\mathbf{1}_B\left(X(k)\right)\right] d\nu(x), \end{aligned} \tag{10.172}$$

for some $0 < r < 1$. A version of the following theorem was first proved by Meyn and Tweedie in [Meyn and Tweedie (1993a)] Theorem 1.1. In fact Theorem 10.8 is a consequence of Proposition 9.1.1 in [Meyn and Tweedie (1993b)], which reads as follows.

Proposition 10.6. *Suppose some subset $B \in \mathcal{E}$ has the following property. For every $x \in B$ the equality $\mathbb{P}_x\left[\tau_B^1 < \infty\right] = 1$ holds. Then*

$$\mathbb{P}_x\left[\sum_{k=1}^{\infty} \mathbf{1}_B\left(X(k)\right) = \infty\right] = \mathbb{P}_x\left[\tau_B^1 < \infty\right], \quad \text{for all } x \in E. \tag{10.173}$$

Proof. Put $\tau_B^0 = \inf\{n \geqslant 0 : X(n) \in B\}$,

$$\tau_B^1 = 1 + \tau_B^0 \circ \vartheta_1 = \inf\left\{n > \tau_B^0 : X(n) \in B\right\} \quad \text{and}$$
$$\tau_B^k = \inf\left\{n > \tau_B^{k-1} : X(n) \in B\right\}, \quad k \geqslant 2.$$

Then $\tau_B^k = \tau_B^{k-1} + \tau_B^1 \circ \vartheta_{\tau_B^{k-1}}$, and hence by the strong Markov property and our assumption on τ_B^1 we have

$$\begin{aligned} \mathbb{P}_x\left[\sum_{\ell=1}^{\infty} \mathbf{1}_B\left(X(\ell)\right) \geqslant k+1\right] &= \mathbb{P}_x\left[\tau_B^{k+1} < \infty\right] \\ &= \mathbb{P}_x\left[\tau_B^1 \circ \vartheta_{\tau_B^k} < \infty,\, \tau_B^k < \infty\right] \\ &= \mathbb{E}_x\left[\mathbb{P}_{X\left(\tau_B^k\right)}\left[\tau_B^1 < \infty\right], \tau_B^k < \infty\right]. \end{aligned} \tag{10.174}$$

Assuming that $\mathbb{P}_y\left[\tau_B^1 < \infty\right] = 1$, $y \in B$, then (10.174) implies

$$\mathbb{P}_x\left[\sum_{\ell=1}^{\infty} \mathbf{1}_B\left(X(\ell)\right) \geqslant k+1\right] = \mathbb{P}_x\left[\tau_B^{k+1} < \infty\right] = \mathbb{P}_x\left[\tau_B^k < \infty\right]. \tag{10.175}$$

By induction with respect to k we see that (10.175) implies

$$\mathbb{P}_x\left[\sum_{\ell=1}^{\infty} \mathbf{1}_B\left(X(\ell)\right) \geqslant k\right] = \mathbb{E}_x\left[\tau_B^1 < \infty\right], \quad \text{for all } x \in E. \tag{10.176}$$

Let k tend to ∞ to obtain (10.173) from (10.176), which completes the proof of Proposition 10.6. □

Theorem 10.8. *Let*

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in E}, (X(k), k \in \mathbb{N}), (\vartheta_k, k \in \mathbb{N}), (E, \mathcal{E})\} \tag{10.177}$$

be a Markov chain with probability transition function $P(x, B)$ which is conservative in the sense that $P(x, E) = 1$ for all $x \in E$. Then the following assertions are equivalent:

(a) There exists a non-zero σ-finite Borel measure μ on $\mathcal{E}$ such that for all $B \in \mathcal{E}$, $\mu(B) > 0$ implies $\mathbb{P}_x\left[\sum_{k=1}^{\infty} \mathbf{1}_B(X(k)) = \infty\right] = 1$ for all $x \in E$.
(b) There exists a non-zero σ-finite Borel measure ν such that for all $B \in \mathcal{E}$, $\nu(B) > 0$ implies $\mathbb{P}_x\left[\tau_B^1 < \infty\right] = 1$ for all $x \in E$.

Here $\tau_B^1 = \inf\{k \geqslant 1 : X(k) \in B\}$.

Proof. Again the implication (a) $\Longrightarrow$ (b) is evident with $\nu = \mu$. Fix $0 < r < 1$. Repeating the arguments in the proof of Theorem 10.7 the reverse implication can be proved with μ given by e.g.

$$\mu(B) = (1 - r) \sum_{k=1}^{\infty} r^{k-1} \int \mathbb{P}_x [X(k) \in B] \, d\nu(x), \quad B \in \mathcal{E}, \tag{10.178}$$

provided that ν is a measure which accommodates assertion (b). However, using Proposition 10.6 we see that implication (a) follows from (b) with $\mu = \nu$ where ν conforms assertion (b).

This completes the proof of Theorem 10.8. □

Remark 10.7. If all probability measures $B \mapsto \mathbb{P}_x [X(1) \in B] = P(x, B)$, $x \in E$, are equivalent, and if (b) is satisfied with $B \mapsto P(x_0, B)$, then (a) holds with the same measure. To see this, consider $\nu(B) = \mathbb{P}_{x_0}[X(1) \in B] = P(x_0, B)$. Then by the Markov property μ in (10.178) is given by

$$\begin{aligned}
\mu(B) &= (1 - r) \sum_{k=1}^{\infty} r^{k-1} \int \mathbb{P}_x [X(k) \in B] \, d\mu(x) \\
&= (1 - r) \sum_{k=1}^{\infty} r^{k-1} \int \mathbb{E}_{x_0} \left[\mathbb{P}_{X(1)} [X(k) \in B]\right] \\
&= (1 - r) \sum_{k=1}^{\infty} r^{k-1} \int \mathbb{P}_{x_0} [X(k+1) \in B], \qquad (10.179)
\end{aligned}$$

and hence, if $\mu(B) = 0$, then $\mathbb{E}_{x_0}\left[\mathbb{P}_{X(1)}\left[X(1) \in B\right]\right] = \mathbb{P}_{x_0}\left[X(2) \in B\right] = 0$. Thus, we see that $\mathbb{P}_{X(1)}\left[X(1) \in B\right] = 0$, $\mathbb{P}_{x_0}$-almost surely. Therefore there exists at least one $x \in E$ such that $P(x, B) = \mathbb{P}_x\left[X(1) \in B\right] = 0$. Since, by assumption, all measures $B \mapsto P(y \in B)$, $y \in E$, are equivalent in follows that $P(x_0, B) = 0$.

The following theorem reduces (Harris) recurrence problems for time-continuous Markov processes with the Feller property and sample space Ω to Markov chains on a larger sample space $\Omega \times \Lambda$ where the continuous time is replaced with the time jump process

$$\left\{(\Lambda, \mathcal{G}, \pi_0), (T_n, n \in \mathbb{N}), \left(\vartheta_n^P, n \in \mathbb{N}\right)\right\}$$

of a Poisson process. Here the variable T_n has π_0-distribution function

$$t \mapsto \pi_0\left[T_n \leqslant t\right] = \pi_0\left[N(t) \geqslant n\right] = \sum_{k=n}^{\infty} \frac{(\alpha_0 t)^k}{k!} e^{-\alpha_0 t}.$$

In the following theorem we see that for certain conservative strong Feller processes the notions of recurrent and Harris recurrent coincide.

Theorem 10.9. *Let the process*

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t), t \geqslant 0), (\vartheta_t, t \geqslant 0), (E, \mathcal{E})\} \tag{10.180}$$

be a Markov process with the strong Feller property, and with a conservative probability transition function $P(t, x, B)$, $(t, x, B) \in [0, \infty) \times E \times \mathcal{E}$. Suppose that all Borel measures $B \mapsto P(t, x, B)$, $(t, x) \in (0, \infty) \times E$, are equivalent i.e. have the same negligible sets. Let the Markov chain

$$\left\{(\Omega \times \Lambda, \mathcal{F} \otimes \mathcal{G}, \mathbb{P}_x \otimes \pi_0), (X(T_n(\lambda), \omega), n \in \mathbb{N}), \left(\vartheta_n^P(\lambda), n \in \mathbb{N}\right), (E, \mathcal{E})\right\} \tag{10.181}$$

be as in (10.141) of Lemma 10.5. Then the following assertions are equivalent:

(a) The Markov process in (10.180) is Harris recurrent in the sense that for any Borel subset B for which $P(t_0, x_0, B) > 0$ for some $(t_0, x_0) \in (0, \infty) \times E$ the equality

$$\mathbb{P}_x\left[\int_0^{\infty} \mathbf{1}_{\{X(t) \in B\}} dt = \infty\right] = 1 \tag{10.182}$$

holds for all $x \in E$.

(b) *The Markov process in (10.180) is recurrent in the sense that for any Borel subset B for which $P(t_0, x_0, B) > 0$ for some $(t_0, x_0) \in (0, \infty) \times E$ the equality*

$$\mathbb{P}_x \left[\tau_B < \infty \right] = 1 \tag{10.183}$$

holds for all $x \in E$.

(c) *The Markov chain in (10.181) is Harris recurrent in the sense that for any Borel subset B for which $P(t_0, x_0, B) > 0$ for some $(t_0, x_0) \in (0, \infty) \times E$ the equality*

$$\mathbb{P}_x \otimes \pi_0 \left[\sum_{k=0}^{\infty} \mathbf{1}_{\{X(T_k) \in B\}} = \infty \right] = 1 \tag{10.184}$$

holds for all $x \in E$.

(d) *The Markov chain in (10.181) is recurrent in the sense that for any Borel subset B for which $P(t_0, x_0, B) > 0$ for some $(t_0, x_0) \in (0, \infty) \times E$ the equality*

$$\mathbb{P}_x \otimes \pi_0 \left[\tau_B^1 < \infty \right] = 1 \tag{10.185}$$

holds for all $x \in E$.

Here $\tau_B = \inf \{ t > 0 : X(t) \in B \}$, and $\tau_B^1 = \inf \{ n \geqslant 1 : X(T_n) \in B \}$.

For the notion of strong Feller property the reader is referred to Definitions 2.5 and 2.16.

Proof. First we observe that the measures $B \mapsto \mathbb{P}_x \otimes \pi_0 [X(T_n) \in B]$, $n \in \mathbb{N}$, $n \geqslant 1$, and $x \in E$, are equivalent to the measures $B \mapsto P(t, x, B) = \mathbb{P}_x [X(t) \in B]$, $(t, x) \in (0, \infty) \times E$. The reason for this equivalence is the following equality:

$$\mathbb{P}_x \otimes \pi_0 \left[X(T_n) \in B \right] = \alpha_0 \int_0^\infty \frac{(\alpha_0 t)^{n-1}}{(n-1)!} e^{-\alpha_0 t} P(t, x, B) \, dt \tag{10.186}$$

which can be found in (10.143). Now we are ready to prove Theorem 10.9.

(a) $\Longleftrightarrow$ (b). This equivalence is a consequence of Theorem 10.7 with $\mu(B) = \nu(B) = P(t_0, x_0, B)$, $B \in \mathcal{B}$.

(c) $\Longleftrightarrow$ (d). This equivalence is a consequence of Theorem 10.8 with $\nu(B) = \mathbb{P}_x \otimes \pi_0 [X(T_1 \in B)]$, and $\mu = \nu$, or

$$\mu(B) = (1 - r) \sum_{k=1}^{\infty} r^{k-1} \mathbb{P}_x \otimes \pi_0 \left[X(T_{k+1}) \in B \right]$$

$$= (1-r)\sum_{k=1} \left(\alpha_0 R\left(\alpha_0\right)\right)^{k+1} \mathbf{1}_B(x). \tag{10.187}$$

For this result the reader is referred to the equalities (10.143) and (10.179), and to Theorem 10.8. Since the measures ν and μ in (10.187) are equivalent to the measure $B \mapsto P\left(t_0, x_0, B\right)$ assertions (c) and (d) are equivalent with the measure $B \mapsto P\left(t_0, x_0, B\right)$.

(d) $\Longrightarrow$ (b). From the definitions of the stopping times τ_B and τ_B^1 it follows the following $\mathbb{P}_{x_0} \otimes \pi_0$-sure inclusion of events:

$$\left\{\tau_B < \infty\right\} \supset \left\{T_{\tau_B^1} < \infty\right\} = \left\{\tau_B^1 < \infty\right\}, \tag{10.188}$$

and hence

$$\mathbb{P}_{x_0}\left[\tau_B < \infty\right] = \mathbb{P}_{x_0} \otimes \pi_0\left[\tau_B < \infty\right] \geqslant \mathbb{P}_{x_0} \otimes \pi_0\left[\tau_B^1 < \infty\right] = 1. \tag{10.189}$$

Assertion (b) is a consequence of (10.189).

(a) $\Longrightarrow$ (c). Let $A \in \mathcal{E}$ be such that $\left(\alpha_0 R\left(\alpha_0\right)\right)^n \mathbf{1}_A\left(x_0\right) > 0$, for some $n \in \mathbb{N}$, $n \geqslant 1$, which by assumption is equivalent to $\mathbb{P}_{x_0}\left[\left(X\left(t_0\right)\right) \in A\right] = P\left(t_0, x_0, A\right) > 0$. Let $\omega \in \Omega$ and put $B_\omega = \{t \geqslant 0 : X(t,\omega) \in A\}$. By assumption (a) we know that $\mathbb{P}_x\left[\int_0^\infty \mathbf{1}_A\left(X(t)\right)\, dt = \infty\right] = 1$ for all $x \in E$. Hence it follows that the Lebesgue measure of B_ω is ∞ for $\mathbb{P}_x$-almost all $\omega \in \Omega$ and for all $x \in A$. An application of equality (10.150) in Lemma 10.7 in the penultimate equality in (10.190) below yields:

$$\begin{aligned}
&\mathbb{P}_x \otimes \pi_0\left[\bigcap_n \bigcup_{m \geqslant n} \left\{X\left(T_m\right) \in A\right\}\right] \\
&= \int_\Omega d\mathbb{P}_x(\omega) \int_\Lambda d\pi_0(\lambda) \limsup_{n\to\infty} \mathbf{1}_{\{X(T_n(\lambda)) \in A\}}(\omega) \\
&= \int_\Omega d\mathbb{P}_x(\omega) \int_\Lambda d\pi_0(\lambda) \limsup_{n\to\infty} \mathbf{1}_{\{T_n \in B_\omega\}}(\lambda) \\
&= \int_\Omega d\mathbb{P}_x(\omega) \int_\Lambda d\pi_0(\lambda) \mathbf{1} = 1.
\end{aligned} \tag{10.190}$$

From (10.190) assertion (c) follows.

This completes the proof of Theorem 10.9 □

Lemma 10.9. *Let the notation and hypotheses be as in Theorem 10.7. Suppose that all measures $B \mapsto P(t,x,B)$, $(t,x) \in (0,\infty) \times E$ are equivalent, and that B is recurrent whenever $P(t,x,B) > 0$ for some pair $(t,x) \in (0,\infty) \times E$. Then all Borel subsets B for which $P(t,x,B) > 0$ for some pair $(t,x) \in (0,\infty) \times E$ are recurrent for the chain described in (10.141) of Lemma 10.5.*

Proof. From equality (10.142) it follows that all transition probability measures $B \mapsto \mathbb{P}_x \otimes \pi_0 [X(T_1) \in B]$, $x \in E$, are equivalent. $\square$

Lemma 10.10. *Let $(e^{tL} : t \geqslant 0)$ be the semigroup associated to the Markov process in (9.14). Put for $\alpha > 0$ and $f \in C_b(E)$*

$$R(\alpha)f(x) = \int_0^\infty e^{-\alpha t} e^{tL} f(x) dt = \int_0^\infty e^{-\alpha t} \mathbb{E}_x [f(X(t))] \, dt, \qquad (10.191)$$

and fix $\alpha_0 > 0$. Let μ be σ-finite Radon measure on the Borel filed of E. Then the following assertions are equivalent:

(1) The measure μ is L-invariant, i.e. $\int Lf \, d\mu = 0$ for all $f \in D(L)$ which belong to $L^1(E,\mu)$;
(2) There exists $\alpha_0 > 0$ such that $\alpha_0 \int R(\alpha_0) f \, d\mu = \int f \, d\mu$ for all $f \geqslant 0$ which are Borel measurable;
(3) For all $\alpha > 0$ and for all Borel measurable functions $f \geqslant 0$ the equality $\alpha R(\alpha) f \, d\mu = \int f \, d\mu$;
(4) The measure μ is invariant for the semigroup $\{e^{tL} : t \geqslant 0\}$, i.e. $\int e^{tL} f d\mu = \int f \, d\mu$ for all $f \geqslant 0$ which are Borel measurable and for all $t \geqslant 0$.

Proof. (1) $\Longrightarrow$ (2). Let the positive σ-finite Radon measure μ be such that $\int Lf \, d\mu = 0$ for $f \in D(L) \bigcap L^1(E,\mu)$, and fix $\alpha_0 > 0$. Then we have for $f \in L^1(E,\mu)$

$$\int \alpha_0 R(\alpha_0) f \, d\mu - \int f \, d\mu = \int LR(\alpha_0) f \, d\mu = 0. \qquad (10.192)$$

From (10.192) we infer $\int \alpha_0 R(\alpha_0) f \, d\mu = \int f \, d\mu$, $f \in L^1(E,\mu)$. This proves the implication (1) $\Longrightarrow$ (2).

(2) $\Longrightarrow$ (3). Let $f \geqslant 0$ belong to $L^1(E,\mu)$, and α_0 as in (2). By the resolvent equation, we have

$$\alpha_0 R(\alpha) + \alpha_0 (\alpha - \alpha_0) R(\alpha_0) R(\alpha) = \alpha_0 R(\alpha_0),$$

an so for $\alpha > \alpha_0$

$$\alpha_0 \int R(\alpha) f \, d\mu + \alpha_0 (\alpha - \alpha_0) \int R(\alpha_0) R(\alpha) f \, d\mu = \alpha_0 \int R(\alpha_0) f \, d\mu,$$

and hence by assertion (2)

$$\alpha_0 \int R(\alpha) f \, d\mu + (\alpha - \alpha_0) \int R(\alpha) f \, d\mu = \int f \, d\mu. \qquad (10.193)$$

From the equality in (10.193) we see that $\alpha \int R(\alpha) f \, d\mu = \int f \, d\mu$, $f \in L^1(E, \mu)$, and $\alpha > \alpha_0$. This shows the implication (2) $\Longrightarrow$ (3) for $\alpha > 0$ and large.

(3) $\Longrightarrow$ (4). Under the restriction that we know (3) for all large α we will show that (4) holds. Let $f \geqslant 0$ be a member of $L^1(E, \mu)$. For all $\alpha > 0$ large (3) entails

$$\begin{aligned} \alpha \int_0^\infty e^{-\alpha\rho} \int e^{\rho L} e^{tL} f \, d\mu \, d\rho &= \alpha \int R(\alpha) e^{tL} f \, d\mu \\ &= \int e^{tL} f \, d\mu = \alpha \int_0^\infty e^{-\alpha\rho} \int e^{tL} f \, d\mu \, d\rho \end{aligned} \tag{10.194}$$

for all $t \geqslant 0$. By uniqueness of Laplace transforms the equality in (10.195) implies $\int e^{\rho L} e^{tL} f \, d\mu = \int e^{tL} f \, d\mu$ for almost all $\rho > 0$. Here the "almost all" depends on $t \geqslant 0$. Next fix $t > 0$. Then by what is proved above we get

$$\int_0^\infty e^{-\alpha\rho} \int e^{\rho L} e^{tL} f \, d\mu \, d\rho = \int_0^\infty e^{-\alpha\rho} \int e^{tL} f d \, d\mu \, d\rho. \tag{10.195}$$

From (10.195) we infer

$$\int_0^\infty e^{-\alpha(\rho+t)} \int e^{(\rho+t)L} f \, d\mu \, d\rho = \int_0^\infty e^{-\alpha(\rho+t)} \int e^{tL} f d \, d\mu \, d\rho, \tag{10.196}$$

or, what amounts to the same, from (10.195) we infer

$$\int_t^\infty e^{-\alpha\rho} \int e^{\rho L} f \, d\mu \, d\rho = \int_t^\infty e^{-\alpha\rho} \int e^{tL} f d \, d\mu \, d\rho. \tag{10.197}$$

As a consequence of (10.197) together with the fact that for almost all $\rho > 0$ the equality $\int e^{\rho L} f d\mu = \int f \, d\mu$ holds we obtain

$$\int_0^\infty e^{-\alpha\rho} \int f \, d\mu \, d\rho = \int_0^\infty e^{-\alpha\rho} \int e^{\rho L} f \, d\mu \, d\rho = \int_0^\infty e^{-\alpha\rho} \int e^{(t \wedge \rho)L} f \, d\mu \, d\rho. \tag{10.198}$$

Again by uniqueness of Laplace transforms the equality in (10.198) implies that for almost all $\rho > 0$ we get $\int e^{(t\wedge\rho)L} f \, d\mu = \int f \, d\mu$. Upon choosing $\rho > t$ we get $\int e^{tL} f \, d\mu = \int f \, d\mu$. Since $t > 0$ is arbitrary assertion (4) is a consequence of the latter equality.

The implications (4) $\Longrightarrow$ (3) $\Longrightarrow$ (2) are easy. The implication (2) $\Longrightarrow$ (1) can be obtained by noting that $D(L)$ is the range of the operator $R(\alpha_0)$, and writing $LR(\alpha_0) f = \alpha_0 R(\alpha_0) f - f$, $f \in C_b(E)$.

This completes the proof of Lemma 10.10. □

The proof of the following theorem will be based on the Markov chain constructed in Theorem 10.9 and on the corresponding result for recurrent Markov chains as exhibited by e.g. Meyn and Tweedie in [Meyn and Tweedie (1993b)]. In fact for Markov chains the result goes back to Harris [Harris (1956)]. Our proof will follow the arguments for the proof of Theorem I.3 in [Azéma *et al.* (1967)]. In case the invariant measures are finite He and Ying [He and Ying (2009)] have a relatively short argument to prove uniqueness.

Theorem 10.10. *Let the Markov process in (10.180) be recurrent in the sense of Theorem 10.9. Moreover, suppose that the hypotheses of Theorem 10.9 are satisfied. Then the process in (10.180) admits an invariant σ-finite measure which is unique op to a multiplicative constant. This measure μ has the property that $\mu(B) > 0$ if and only if $P(t_0, x_0, B) > 0$ for some (all) $(t_0, x_0) \in (0, \infty) \times E$.*

The following theorem can be found in [Harris (1956)]. It is a consequence of Theorem 10.3.

Theorem 10.11. *Suppose that for the time-homogenous Markov chain*

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(i), i \in \mathbb{N}), (\vartheta_i, i \in \mathbb{N}), (E, \mathcal{E})\} \tag{10.199}$$

there exists a σ-finite measure m such that $m(B) > 0$ implies

$$\mathbb{P}_x \left[\sum_{k=0}^{\infty} \mathbf{1}_B (X(k)) = \infty \right] = 1 \quad \text{for all } x \in E.$$

In other words the Markov chain in (10.199) is m-recurrent. Then there exists a σ-finite invariant measure μ which is unique up to a multiplicative constant, and which is such that μ is absolutely continuous with respect to μ, i.e. $\mu(A) = 0$ implies $\mu(A) = 0$, $A \in \mathcal{E}$.

For much more explanation about Markov chains see e.g. [Meyn and Tweedie (1993b)]. We will take Theorem 10.11 for granted as we did with Theorem 10.3.

Proof. [Proof of Theorem 10.10.] Fix $(t_0, x_0) \in (0, \infty) \times E$, and put $m(B) = P(t_0, x_0, B)$, $B \in \mathcal{E}$. From Theorem 10.9 it follows that the Markov chain in (10.181) is m-recurrent if and only if the Markov process in (10.180) is m-recurrent. So by Theorem 10.11 the Markov chain in (10.181), i.e.

$$\left\{(\Omega \times \Lambda, \mathcal{F} \otimes \mathcal{G}, \mathbb{P}_x \otimes \pi_0), (X(T_n(\lambda), \omega), n \in \mathbb{N}), \left(\vartheta_n^P(\lambda), n \in \mathbb{N}\right), (E, \mathcal{E})\right\} \tag{10.200}$$

admits a σ-finite invariant measure μ which is equivalent to the measure m. By Lemma 10.10 the measure μ is also invariant for the Markov process in (10.180) of Theorem 10.9, i.e. for

$$\left\{\left(\Omega,\mathcal{F},\mathbb{P}_x\right),\left(X\left(t\right),t\geqslant 0\right),\left(\vartheta_t,t\geqslant 0\right),\left(E,\mathcal{E}\right)\right\}. \tag{10.201}$$

Since the σ-finite invariant measures for the processes in (10.200) are unique up to multiplicative constants, the same is true for the σ-finite invariant measures for the Markov process in (10.201). Moreover, by Theorem 10.11 these invariant measures are equivalent with the measure m. Altogether this proves Theorem 10.10. □

10.2.4 *Actual construction of an invariant measure*

Theorem 9.4, which is one of the most important results in Chapter 9, gives sufficient conditions in order that the Markov process in (10.180) possesses a compact recurrent subset. This assumption of the existence of such a compact subset is made in the following theorem.

Theorem 10.12. *Suppose that there exists a compact recurrent subset A, and suppose that the Markov process in (10.180) is irreducible and strong Feller. In addition, suppose that all measures $B \mapsto P\left(t,x,B\right)$, $x \in E$, $t > 0$, are equivalent. Then there exists a non-trivial σ-finite invariant measure π, and the vector sum $R(L) + N(L)$ is dense in $C_b(E)$ for the strict topology. In fact the measure π has the property that $f \in C_b(E)$, $f \geqslant 0$, $f \neq 0$, implies $\int f\,d\pi > 0$. Moreover, $\pi(B) = 0$ if and only if $P\left(t,x,B\right) = 0$ for all pairs (some pair) $(t,x) \in (0,\infty) \times E$. Moreover, the measure π is unique up to a multiplicative constant.*

For the notion of strong Feller property see Definitions 2.5 and 2.16. A combination of Theorem 10.12 and Theorem 9.4 in Chapter 9 yields the following result.

Theorem 10.13. *Let L be the generator of a strong Markov process which almost separate points and closed subsets, in the sense that for every $x \in U$ with U open there exists a function $v \in D(L)$ such that $v(x) > \sup_{y\in E\setminus U} v(y)$. Suppose that every non-empty open subset is recurrent, and that all measures of the form $B \mapsto P(t,x,B)$, $B \in \mathcal{E}$, $(t,x) \in (0,\infty) \times E$, are equivalent probability measures. Then there exists a non-trivial σ-finite invariant measure π, provided that all functions of the form $(t,x) \mapsto P(t,x,B)$, $(t,x) \in (0,\infty) \times E$, $B \in \mathcal{E}$, are continuous. Moreover, the invariant measure π is unique up to a multiplicative constant.*

In addition, suppose that there exists a recurrent subset A such that $\sup_{x\in A}\mathbb{E}_x\left[h+\tau_A\circ\vartheta_h\right]<\infty$ for some $h>0$. Then the invariant measure is finite, and may be chosen to be a probability measure.

Proof. Under the hypotheses of Theorem 10.13 there exists a compact recurrent subset by Theorem 9.4. Theorem 10.12 yields the desired conclusion in the first statement of Theorem 10.13. The second one is a consequence of Corollary 10.5 below because our extra assumption is the same as the finiteness assumption in (10.292). □

Remark 10.8. From the proof of Theorem 10.12 it follows that for every compact subset K there exists an open subset $K_\varepsilon \supset K$, and hence a function $f_K \in C_b(E)$ such that

$$\mathbf{1}_K \leqslant f_K \leqslant \mathbf{1}_{K_\varepsilon}, \quad \text{and} \quad \int f_K\, d\pi < \infty \tag{10.202}$$

where π is the invariant measure. It also follows that

$$E = \bigcup_{K,\, K\text{compact}} \{f_K > 0\}.$$

Since the space E is second countable, the family $\{f_K : K\,\text{compact}\}$ in (10.202) may be chosen countable, while still satisfying $E = \bigcup_{n\in\mathbb{N}}\{f_{K_n}>0\}$. This can be seen as follows. The second countability implies that there exists a sequence of open subsets $(U_n)_{n\in\mathbb{N}}$ such that for every compact subset K of E there a countable subset $(U_{K,k})_{k\in\mathbb{N}} \subset (U_n)_{n\in\mathbb{N}}$ such that $\{f_K>0\} = \bigcup_{k\in\mathbb{N}} U_{K,k}$. For every $n\in\mathbb{N}$ we choose a compact subset K_n such that $U_n \subset \{f_{K_n}>0\}$. We only take into account those open subsets U_n for which such f_{K_n} exists. Then the sequence $(f_{K_n})_{n\in\mathbb{N}}$ will be such that $E = \bigcup_{n\in\mathbb{N}}\{f_{\alpha_n}>0\}$.

Here, the space $C_b(E)$ is supplied with the strict topology. A sequence $(f_n)_{n\in\mathbb{N}}$ converges with respect to the strict topology if it is uniformly bounded and if it converges to a function $f\in C_b(E)$ uniformly on compact subsets of the space E. The symbol $R(L)$ stands for the range of L, and $N(L)$ stands for the null space of L.

Proof. [Proof of Theorem 10.12.] We sketch a proof. Fix $h>0$, $\lambda>0$, $\mu\in\mathcal{M}(A)$, and $f\in C_b(E)$. Here $\mathcal{M}(A)$ is the space of those (complex) measures $\mu\in\mathcal{E}$ which are concentrated on A; i.e. $|\mu|(E\backslash A)=0$. We will also need the following stopping times:

$$\tau_A^h = \inf\{s>h:\, X(s)\in A\} = h+\tau_A\circ\vartheta_h \quad \text{where } \tau_A \text{ is the hitting time}$$
$$\tau_A = \inf\{t>0:\, X(t)\in A\}. \tag{10.203}$$

Therefore we will rewrite the equality:

$$\begin{aligned}&\int_0^h e^{-\lambda' s}e^{sL}\,ds\left\{\left(L-\lambda' I\right)R_A\left(\lambda'\right)+I\right\}f\left(x\right)\\&=\left\{\left(e^{-\lambda' h}e^{hL}-I\right)R_A\left(\lambda'\right)+\int_0^h e^{-\lambda' s}e^{sL}\,ds\right\}f\left(x\right).\end{aligned}\tag{10.204}$$

The expression in (10.204) is equal to

$$\begin{aligned}&\left(e^{-h\lambda'}e^{hL}-I\right)R_A\left(\lambda'\right)f(x)+\int_0^h e^{-\lambda' s}e^{sL}f(x)\,ds\\&=e^{-h\lambda'}\mathbb{E}_x\left[R_A\left(\lambda'\right)f\left(X(h)\right)\right]-R_A\left(\lambda'\right)f(x)+\int_0^h e^{-\lambda' s}e^{sL}f(x)\,ds\\&=e^{-h\lambda'}\mathbb{E}_x\left[\mathbb{E}_{X(h)}\left[\int_0^{\tau_A}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho\right]\right]-\mathbb{E}_x\left[\int_0^{\tau_A}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho\right]\\&\quad+\mathbb{E}_x\left[\int_0^h e^{-\lambda' s}f\left(X(s)\right)\,ds\right]\end{aligned}$$

(Markov property)

$$\begin{aligned}&=\mathbb{E}_x\left[\int_0^{\tau_A\circ\vartheta_h}e^{-\lambda'(h+\rho)}f\left(X(h+\rho)\right)\,d\rho\right]-\mathbb{E}_x\left[\int_0^{\tau_A}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho\right]\\&\quad+\mathbb{E}_x\left[\int_0^h e^{-\lambda' s}f\left(X(s)\right)\,ds\right]\\&=\mathbb{E}_x\left[\int_h^{h+\tau_A\circ\vartheta_h}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho\right]-\mathbb{E}_x\left[\int_0^{\tau_A}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho\right]\\&\quad+\mathbb{E}_x\left[\int_0^h e^{-\lambda' s}f\left(X(s)\right)\,ds\right]\end{aligned}$$

(τ_A is a terminal stopping time: on $\{\tau_A > h\}$ the equality $h+\tau_A\circ\vartheta_h=\tau_A$ holds $\mathbb{P}_x$-almost surely)

$$\begin{aligned}&=\mathbb{E}_x\left[\int_h^{h+\tau_A\circ\vartheta_h}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho,\,\tau_A\leqslant h\right]\\&\quad+\mathbb{E}_x\left[\int_h^{\tau_A}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho,\,\tau_A>h\right]-\mathbb{E}_x\left[\int_0^{\tau_A}e^{-\lambda'\rho}f\left(X(\rho)\right)\,d\rho\right]\\&\quad+\mathbb{E}_x\left[\int_0^h e^{-\lambda' s}f\left(X(s)\right)\,ds,\tau_A\leqslant h\right]+\mathbb{E}_x\left[\int_0^h e^{-\lambda' s}f\left(X(s)\right)\,ds,\tau_A>h\right]\end{aligned}$$

$$= \mathbb{E}_x \left[\int_{\tau_A}^{h+\tau_A\circ\vartheta_h} e^{-\lambda'\rho} f\left(X(\rho)\right)\, d\rho,\ \tau_A \leqslant h \right]$$

$$= \mathbb{E}_x \left[e^{-\lambda'\tau_A} \int_0^{h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}\circ\vartheta_{\tau_A}} e^{-\lambda'\rho} f\left(X\left(\rho+\tau_A\right)\right)\, d\rho,\ \tau_A \leqslant h \right]$$

(strong Markov property)

$$= \mathbb{E}_x \left[e^{-\lambda'\tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right],\ \tau_A \leqslant h \right] \tag{10.205}$$

$$= \mathbb{E}_x \left[e^{-\lambda' h} \mathbb{E}_{X(\tau_A)} \left[\int_0^{\tau_A\circ\vartheta_{h-\tau_A}} e^{-\lambda'\rho} f\left(X\left(\rho+h-\tau_A\right)\right)\, d\rho \right],\ \tau_A \leqslant h \right]$$
$$+ \mathbb{E}_x \left[e^{-\lambda'\tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{h-\tau_A} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right],\ \tau_A \leqslant h \right]$$

(Markov property once more)

$$= \mathbb{E}_x \left[e^{-\lambda' h} \mathbb{E}_{X(\tau_A)} \left[\mathbb{E}_{X(h-\tau_A)} \left[\int_0^{\tau_A} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right]\right],\ \tau_A \leqslant h \right]$$
$$+ \mathbb{E}_x \left[e^{-\lambda'\tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{h-\tau_A} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right],\ \tau_A \leqslant h \right]. \tag{10.206}$$

It is perhaps useful to explain the way the expectations in (10.206) have to be understood. The second term should be read as follows:

$$\mathbb{E}_x \left[e^{-\lambda'\tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{h-\tau_A} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right],\ \tau_A \leqslant h \right]$$
$$= \mathbb{E}_x \left[\omega \mapsto e^{-\lambda'\tau_A(\omega)} \mathbb{E}_{X(\tau_A)(\omega)} \left[\int_0^{h-\tau_A(\omega)} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right] \mathbf{1}_{\{\tau_A\leqslant h\}}(\omega) \right]$$

where $X\left(\tau_A\right)(\omega) = X\left(\tau_A(\omega)\right)(\omega) = X\left(\tau_A(\omega),\omega\right)$. The first term in (10.206) has to be interpreted in the following manner:

$$\mathbb{E}_x \left[e^{-\lambda' h} \mathbb{E}_{X(\tau_A)} \left[\mathbb{E}_{X(h-\tau_A)} \left[\int_0^{\tau_A} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right]\right],\ \tau_A \leqslant h \right]$$
$$= \mathbb{E}_x \left[\omega \mapsto e^{-\lambda' h} \mathbb{E}_{X(\tau_A(\omega),\omega)} \left[\omega' \mapsto \mathbb{E}_{X(h-\tau_A(\omega),\omega')} \left[\int_0^{\tau_A} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho \right]\right] \right.$$
$$\left. \mathbf{1}_{\{\tau_A\leqslant h\}}(\omega) \right].$$

The equality of the expressions in (10.205) and (10.206) will be used to prove the existence and uniqueness (up to scalar multiples) of an invariant measure. A crucial role will be played by Proposition 10.8.

The equality in (10.204) will also be used to prove that the invariant measure π is strictly positive in the sense that $\int f\,d\pi > 0$ whenever $f \in C_b(E)$ is such that $f \geqslant 0$ and $f \neq 0$. This claim follows from the first equality in (10.255) in Proposition 10.7 together with the first inequality in (10.241) in Lemma 10.11 below. Here we also need the irreducibility of the Markov process X. So let $f \geqslant 0$, $f \neq 0$, $f \in C_b(E) \bigcap L^1(E, \mathcal{E}, \pi)$. Then, from the first equality in (10.255) we see:

$$\begin{aligned}
& h \int_E f(x)\,d\pi(x) \\
&= \int_E \mathbb{E}_x \left[\mathbb{E}_{X(\tau_A)} \left[\int_0^{h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}} f(X(\rho))\,d\rho \right], \tau_A \leqslant h \right] d\pi(x) \\
&\geqslant \int_E \mathbb{E}_x \left[\mathbb{E}_{X(\tau_A)} \left[\int_0^{\frac{1}{2}h} f(X(\rho))\,d\rho \right], \tau_A \leqslant \frac{1}{2}h \right] d\pi(x) \\
&\geqslant \inf_{y\in A} \mathbb{E}_y \left[\int_0^{\frac{1}{2}h} f(X(\rho))\,d\rho \right] \int_E \mathbb{E}_x \left[\tau_A \leqslant \frac{1}{2}h \right] d\pi(x) \\
&= \mathbb{E}_{y_0} \left[\int_0^{\frac{1}{2}h} f(X(\rho))\,d\rho \right] \int_E \mathbb{E}_x \left[\tau_A \leqslant \frac{1}{2}h \right] d\pi(x)
\end{aligned} \tag{10.207}$$

for some $y_0 \in A$. By irreducibility we have

$$\mathbb{E}_{y_0} \left[\int_0^{\frac{1}{2}h} f(X(\rho))\,d\rho \right] > 0. \tag{10.208}$$

The combination of the first inequality in (10.241) in Lemma 10.11 and (10.207) shows that $\int_E f(x)\,d\pi(x) > 0$, where $f \geqslant 0$, $f \neq 0$, $f \in C_b(E) \bigcap L^1(E, \mathcal{E}, \pi)$: see (10.207). As a consequence we have that the corresponding measure π is strictly positive in the sense that $\pi(O) > 0$ for every non-empty open subset O of E. In addition, we have $\pi(B) = 0$, $B \in \mathcal{E}$, if and only if $P(t,x,B) = 0$ for some $(t,x) \in (0,\infty) \times E$. If $P(t_0, x_0, B) = 0$ for some $(t_0, x_0) \in (0,\infty) \times E$, then $P(t,x,B) = 0$ for all $(t,x) \in (0,\infty) \times E$, and hence $\pi(B) = \int P(t,x,B)\,d\pi(x) = \pi(B) = 0$. Conversely, suppose $B \in \mathcal{E}$ is such that $\pi(B) = 0$. Then $\int P(t_0,x,B)\,d\pi(x) = 0$ (by invariance). Since, by the strong Feller property the function $x \mapsto P(t_0,x,B)$ is continuous it follows by the strict positiveness of the measure π that $P(t_0,x,B) = 0$ for some $x \in E$. Since all the measures $B \mapsto P(t_0,x,B)$,

$x \in E$, are equivalent it follows that $P(t_0, x_0, B) = 0$. For the notion of strong Feller property see Definitions 2.5 and 2.16 as well.

First let us embark on the existence of the invariant measure π. We will use a Hahn-Banach argument to obtain such a measure. Recall that

$$\tau_A^h = \inf\{s > h : X(s) \in A\} = h + \tau_A \circ \vartheta_h$$

where $\tau_A = \inf\{s > 0 : X(s) \in A\}$. Since the compact subset A is recurrent we see that

$$\mathbb{P}_x\left[\tau_A^h < \infty\right] = \mathbb{P}_x\left[\tau_A \circ \vartheta_h < \infty\right] = \mathbb{E}_x\left[\mathbb{P}_{X(h)}\left[\tau_A < \infty\right]\right] = \mathbb{E}_x\left[\mathbf{1}\right] = 1, \tag{10.209}$$

and hence the stopping time τ_A^h is finite $\mathbb{P}_x$-almost surely for all $x \in E$. Define the operator $Q_A : C(A) \to C(A)$ by

$$Q_A f(x) = \mathbb{E}_x\left[f\left(X\left(\tau_A^h\right)\right)\right] = \mathbb{E}_x\left[\mathbb{E}_{X(h)}\left[f\left(X\left(\tau_A\right)\right)\right]\right], \quad f \in C(A). \tag{10.210}$$

By the strong Feller property of the Markov process $X(t)$ it follows that the operator Q_A in (10.210) is a positivity preserving linear mapping from $C(A)$ to $C(A)$. For the notion of strong Feller property see Definitions 2.5 and 2.16. Moreover, $Q_A\mathbf{1} = \mathbf{1}$. Fix $x_0 \in E$. By the Hahn-Banach extension theorem there exists a positive linear functional $\Lambda_{x_0} : C(A) \to \mathbb{R}$ such that for $f \in C(A)$, $f \geqslant 0$,

$$\liminf_{r\uparrow 1}(1-r)\sum_{k=0}^{\infty} r^k Q_A^k f\left(x_0\right) \leqslant \Lambda_{x_0}(f) \leqslant \limsup_{r\uparrow 1}(1-r)\sum_{k=0}^{\infty} r^k Q_A^k f\left(x_0\right). \tag{10.211}$$

To obtain Λ, apply the analytic version of the Theorem of Hahn-Banach to the functional:

$$f \mapsto \inf_{g\in C(A), g\geqslant 0} \limsup_{r\uparrow 1}(1-r)\sum_{k=0}^{\infty} r^k\left(Q_A^k\left(f+g\right)\left(x_0\right) - Q_A^k g\left(x_0\right)\right). \tag{10.212}$$

From (10.211) it follows that $\Lambda_{x_0}\left(\mathbf{1}_A\right) = 1$. From Hahn-Banach's theorem it also follows that the second inequality in (10.211) holds for all $f \in C(A)$. Consequently, we have

$$\begin{aligned}\Lambda_{x_0}\left(f - Q_A f\right) &\leqslant \limsup_{r\uparrow 1}(1-r)\sum_{k=0}^{\infty} r^k Q_A^k\left(I - Q_A\right) f\left(x_0\right) \\ &= \limsup_{r\uparrow 1}\left((1-r) f\left(x_0\right) - (1-r)^2\sum_{k=0}^{\infty} r^k Q_A^k f\left(x_0\right)\right) = 0. \end{aligned} \tag{10.213}$$

From (10.213) we infer $\Lambda_{x_0}(f - Q_A f) \leqslant 0$. The latter inequality is also true for $-f$ instead of f, and hence the functional Λ_{x_0} is Q_A-invariant. Since the subset A is compact, by the Riesz representation theorem the functional Λ_{x_0} can be represented by a measure π_{x_0}: $\Lambda_{x_0}(f) = \int_A f(x)d\,d\pi_{x_0}(x)$, $f \in C(A)$. In order to see the uniqueness we use Orey's theorem 10.2. First we introduce the sequence of stopping times: $\tau_A^{k+1,h} = \tau_A^{k,h} + \tau_A^{1,h} \circ \vartheta_{\tau_A^{k,h}}$, where $\tau_A^{1,h} = \tau_A^h$, the stopping time defined in (10.203) and not the time defined in (10.100). We will employ the reference measure $B \mapsto \mathbb{P}_x\left[X\left(\tau_A^{1,h}\right) \in B\right]$. We need the fact that all measures of the form

$$\begin{aligned} \mathbb{P}_x\left[X\left(\tau_A^{k+1,h}\right) \in B\right] &= \mathbb{E}_x\left[\mathbb{P}_{X\left(\tau_A^{k,h}\right)}\left[X\left(\tau_A^{1,h}\right) \in B\right]\right] \\ &= \mathbb{E}_x\left[\mathbb{E}_{X\left(\tau_A^{k,h}\right)}\left[\mathbb{P}_{X(h)}\left[X\left(\tau_A\right) \in B\right]\right]\right], \quad k \in \mathbb{N}, \end{aligned} \tag{10.214}$$

are equivalent. Suppose that B is such that the very first member in (10.214) vanishes. Then there exists $y = X\left(\tau_A^{k,h}\right) \in A$ such that

$$\mathbb{E}_y\left[\mathbb{P}_{X(h)}\left[X\left(\tau_A\right) \in B\right]\right] = 0. \tag{10.215}$$

Since all measures of the form $B \mapsto P(h, y, B)$, $y \in E$, are equivalent, (10.215) implies that the quantity in (10.215) vanishes for all $y \in E$. It follows that

$$\mathbb{P}_{X\left(\tau_A^{\ell,h}\right)}\left[X\left(\tau_A^{1,h}\right) \in B\right] = 0 \tag{10.216}$$

for all $\ell \in \mathbb{N}$. As a consequence we see that the process $k \mapsto X\left(\tau_A^{k,h}\right)$ is Harris recurrent relative to the measure $B \mapsto \mathbb{P}_y\left[X\left(\tau_A^{k,h}\right) \in B\right]$, $B \in \mathcal{E}$. Then Orey's theorem yields that for all pairs of probability measures (μ_1, μ_2) on the Borel field of A the following limit vanishes (see (10.15) in Theorem 10.2):

$$\lim_{n\to\infty} \iint \operatorname{Var}\left(Q_A^n(x,\cdot) - Q_A^n(y,\cdot)\right) d\mu_1(x)\, d\mu_2(y) = 0. \tag{10.217}$$

Consequently, we see that Q_A-invariant probability measures on the Borel field of A are unique. We call such an invariant measure π_A. The existence was established using the Hahn-Banach theorem. It then follows that for all $f \in C(A)$ and uniformly for $x \in A$

$$\lim_{r\uparrow 1}(1-r)\sum_{k=0}^{\infty} r^k Q_A^k f(x) - \int_A f\, d\pi_A \mathbf{1}_A = 0. \tag{10.218}$$

Assertions (b), (c), (d), and (e) in Proposition 10.8 below then show the existence and uniqueness (up to scalar multiplications of e^{tL}-invariant measures) on the Borel field of E.

Next we prove that the invariant measure π on E, the existence of which is established by Proposition 10.8, is in fact a σ-finite, and strictly positive invariant Radon measure which is equivalent to the measures $B \mapsto P(t,x,B)$. This will be the subject of the remaining part of the proof.

The σ-finiteness of the measure π follows from Lemma 10.11. More precisely, put

$$A_{m,n} = \left\{ x \in E : \mathbb{P}_x \left[\tau_A \leqslant m \right] > \frac{1}{n} \right\}, \ m, n \in \mathbb{N}. \tag{10.219}$$

Then $E = \bigcup_{n,m\in\mathbb{N}} A_{m,n}$. Since by Lemma 10.11 $\int_E \mathbb{P}_x \left[\tau_A \leqslant m \right] d\pi(x) < \infty$, it follows that $\pi(A_{m,n}) < \infty$ for all $m, n \in \mathbb{N} \setminus \{0\}$.

From (10.273) in assertion (f) of Proposition 10.8 and (10.205) it follows that for $f \in C_b(E)$, $f \geqslant 0$,

$$h \int_E f(x) d\pi(x) \leqslant \sup_{y\in A} \mathbb{E}_y \left[\int_0^{h+\tau_A\circ\vartheta_h} f(X(\rho))\, d\rho \right] \int_E \mathbb{P}_x \left[\tau_A \leqslant h \right] d\pi(x). \tag{10.220}$$

From (10.205) and (10.220) we will infer that the measure π is σ-finite, and that it is a Radon measure. In the proof of this result we will adapt the proof of Theorem 9.5 in Chapter 9. In particular the inequality in (9.53) is relevant. The precise arguments run as follows. Let K be a compact subset of E such that $A \subset K$. Then there exists $\varepsilon_0 > 0$ with the property that

$$\sup_{y\in A} \mathbb{P}_y \left[X(t) \notin K_\varepsilon \ \text{ for all } t \in [h, h+\tau_A\circ\vartheta_h) \right] > 0, \tag{10.221}$$

for all $0 < \varepsilon < \varepsilon_0$. Below we will show that under the hypotheses of Theorem 10.12 the inequality in (10.221) is satisfied indeed: see (10.238). Here $K_\varepsilon := \{x \in E : d(x,K) \leqslant \varepsilon\}$ stands for an ε-neighborhood of K: d denotes a compatible metric on the Polish space E. We are going to show that

$$\sup_{y\in E} \mathbb{E}_y \left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(\rho))\, d\rho \right] < \infty \tag{10.222}$$

for some $\varepsilon > 0$. Let τ_ε be the first hitting time of K_ε. From (10.221) it follows that for every $\varepsilon \in (0,\varepsilon_0)$ there exists $y_\varepsilon \in A$ such that

$$\mathbb{P}_{y_\varepsilon} \left[\tau_\varepsilon\circ\vartheta_h \geqslant \tau_A\circ\vartheta_h \right] = \mathbb{P}_{y_\varepsilon} \left[X(t) \notin K_\varepsilon \ \text{ for all } t \in [h, h+\tau_A\circ\vartheta_h) \right] > 0. \tag{10.223}$$

The function $y \mapsto \mathbb{P}_y \left[\tau_\varepsilon\circ\vartheta_h \geqslant \tau_A\circ\vartheta_h \right] = \mathbb{E}_y \left[\mathbb{P}_{X(h)} \left[\tau_\varepsilon \geqslant \tau_A \right] \right]$ is continuous, and so there exists a neighborhood V_ε of y_ε such that

$$\alpha_\varepsilon := \inf_{x\in V_\varepsilon} \mathbb{P}_x \left[\tau_\varepsilon\circ\vartheta_h \geqslant \tau_A\circ\vartheta_h \right] > 0. \tag{10.224}$$

and such that

$$\inf_{x\in K_\varepsilon} P\left(t_0,x,V_\varepsilon\right) > 0 \tag{10.225}$$

for some fixed but arbitrary $t_0 > h$. In (10.225) we used the irreducibility of the Markov process and the continuity of the function $x \mapsto P\left(t_0,x,V_\varepsilon\right)$ for $\varepsilon > 0$. If necessary we choose a smaller neighborhood V_ε of y_ε and a smaller ε, which we are entitled to do, because (10.223) holds for every $\varepsilon \in (0,\varepsilon_0)$. Choose $y \in K_\varepsilon$. Then by the Markov property we have

$$\begin{aligned}
&\mathbb{P}_y\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(t)\right)dt < t_0\right]\\
&\geqslant \mathbb{P}_y\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(t)\right)dt < t_0,\ t_0 < h+\tau_A\circ\vartheta_h\right]\\
&= \mathbb{E}_y\left[\omega\mapsto \mathbb{P}_{X(t_0)(\omega)}\left[\int_0^{t_0} \mathbf{1}_{K_\varepsilon}\left(X(t)(\omega)\right)dt\right.\right.\\
&\qquad\left.\left. + \int_0^{h+\tau_A\circ\vartheta_h(\omega)-t_0} \mathbf{1}_{K_\varepsilon}\left(X(t)\right)dt < t_0\right]\mathbf{1}_{\{t_0<h+\tau_A\circ\vartheta_h\}}(\omega)\right]\\
&\geqslant \mathbb{E}_y\left[\omega\mapsto \mathbb{P}_{X(t_0)(\omega)}\left[\int_0^{t_0} \mathbf{1}_{K_\varepsilon}\left(X(t)(\omega)\right)dt < t_0,\right.\right.\\
&\qquad\left.\left. X(t)\notin K_\varepsilon \ \text{ for all } t\in\left[0,h+\tau_A\circ\vartheta_h(\omega)-t_0\right)\right]\mathbf{1}_{\{t_0<h+\tau_A\circ\vartheta_h\}}(\omega)\right]\\
&\geqslant \mathbb{P}_y\left[\int_0^{t_0} \mathbf{1}_{K_\varepsilon}\left(X(t)\right)dt < t_0,\right.\\
&\qquad\left. X(t)\notin K_\varepsilon \ \text{ for all } t\in\left[t_0,h+\tau_A\circ\vartheta_h\right),\ h+\tau_A\circ\vartheta_h > t_0\right]\\
&\geqslant \mathbb{P}_y\left[\int_0^{t_0} \mathbf{1}_{K_\varepsilon}\left(X(t)\right)dt < t_0,\right.\\
&\qquad\left. X\left(t_0\right)\in V_\varepsilon,\ X(t)\notin K_\varepsilon \ \text{ for all } t\in\left[t_0,h+\tau_A\circ\vartheta_h\right),\ h+\tau_A\circ\vartheta_h > t_0\right]\\
&\geqslant \mathbb{E}_y\left[\mathbb{P}_{X(t_0)}\left[\tau_\varepsilon\circ\vartheta_h \geqslant \tau_A\circ\vartheta_h\right],\ X\left(t_0\right)\in V_\varepsilon\right]
\end{aligned}$$

(apply (10.224), the definition of α_ε)

$$\geqslant \alpha_\varepsilon P\left(t_0,y,V_\varepsilon\right) \geqslant \alpha_\varepsilon \inf_{x\in K_\varepsilon} P\left(t_0,x,V_\varepsilon\right) =: q > 0, \tag{10.226}$$

where we used the irreducibility of our Markov process, and the continuity of the function $x \mapsto P\left(t_0,x,V_\varepsilon\right)$. Hence we infer

$$\sup_{y\in K_\varepsilon} \mathbb{P}_y\left[\int_0^{h+\tau_h\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(t)\right)dt \geqslant t_0\right] \leqslant 1-q. \tag{10.227}$$

Put

$$\kappa_\varepsilon = \inf\left\{t > h : \int_0^t \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds \geqslant t_0\right\}$$
$$= \inf\left\{t > h : \int_0^t \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds = t_0\right\}. \tag{10.228}$$

Then κ_ε is a stopping time relative to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, because $X(s)$ is $\mathcal{F}_t$-measurable for all $0 \leqslant s \leqslant t$. Moreover, by right-continuity of the process $t \mapsto X(t)$ it follows that $X(\kappa_\varepsilon) \in K_\varepsilon$ on the event $\{\tau_\varepsilon < \infty\}$. Let $y \in A$. By induction we shall prove that

$$\mathbb{P}_y\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(t)\right) dt > kt_0\right] \leqslant (1-q)^{k-1}, \ k \in \mathbb{N}, \ k \geqslant 1. \tag{10.229}$$

To this end we put

$$\alpha_k = \sup_{x\in K_\varepsilon} \mathbb{E}_x\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds \geqslant kt_0\right]. \tag{10.230}$$

If x belongs to K_ε, then by the Markov property we have:

$$\mathbb{P}_x\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds > (k+1)t_0\right]$$
$$= \mathbf{P}_x\left[\int_{\kappa_\varepsilon}^{h+\tau_A\circ_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds > kt_0, \ \kappa_\varepsilon \leqslant h + \tau_A\circ\vartheta_h\right]$$
$$= \mathbb{E}_x\left[\mathbb{P}_{X(\kappa_\varepsilon)}\left[\int_0^\infty \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds > kt_0\right], \kappa_\varepsilon \leqslant h + \tau_A\circ\vartheta_h\right]$$
$$= \mathbb{E}_x\left[\mathbb{P}_{X(\kappa_\varepsilon)}\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds > kt_0\right],\right.$$
$$\left.\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds \geqslant t_0\right]$$
$$\leqslant \alpha_1\alpha_k. \tag{10.231}$$

From (10.231) and induction we infer

$$\sup_{x\in K_\varepsilon} \mathbb{P}_x\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds \geqslant kt_0\right]$$
$$\leqslant \alpha_1^k = \left(\sup_{x\in K_\varepsilon}\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}\left(X(s)\right) ds \geqslant t_0\right]\right)^k \leqslant (1-q)^k, \tag{10.232}$$

where in the final step of (10.232) we employed (9.58). If $y \in E$ is arbitrary, then we proceed as follows:

$$\begin{aligned}
&\mathbb{P}_y\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(s))\,ds > (k+1)t_0\right] \\
&= \mathbb{P}_y\left[\int_{\kappa_\varepsilon}^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(s))\,ds > kt_0,\ \kappa_\varepsilon \leqslant h+\tau_A\circ\vartheta_h\right] \\
&= \mathbb{E}_y\left[\mathbb{P}_{X(\kappa_\varepsilon)}\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(s))\,ds > kt_0\right],\ \kappa_\varepsilon \leqslant h+\tau_A\circ\vartheta_h\right] \\
&\leqslant (1-q)^k\mathbb{P}_y\left[\kappa_\varepsilon \leqslant h+\tau_A\circ\vartheta_h\right] \leqslant (1-q)^k. \qquad (10.233)
\end{aligned}$$

The inequality in (10.233) implies the one in (10.229). To show the first part of (10.222) with $f = \mathbf{1}_{K_\varepsilon}$, for $\varepsilon > 0$ small enough, we observe that for $x \in E$ we have

$$\begin{aligned}
&\mathbb{E}_x\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(s))\,ds\right] \\
&\leqslant \sum_{k=1}^{\infty} kt_0\mathbb{P}_x\left[(k-1)t_0 < \int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(s))\,ds \leqslant kt_0\right] \\
&\leqslant t_0 + \sum_{k=2}^{\infty} \mathbb{P}_x\left[\int_0^{h+\tau_A\circ\vartheta_h} \mathbf{1}_{K_\varepsilon}(X(s))\,ds > (k-1)t_0\right] \\
&\leqslant t_0 + t_0\sum_{k=2}^{\infty} k(1-q)^{k-2} = t_0\left(1+\frac{1}{q}+\frac{1}{q^2}\right) < \infty. \qquad (10.234)
\end{aligned}$$

The inequality in (10.222) is a consequence of (10.234) indeed with $f = \mathbf{1}_{K_\varepsilon}$. In other words for every compact subset K of E there exists an ε-neighborhood $K_\varepsilon \supset K$ such that (10.222) is satisfied. It follows that the functional $f \mapsto \int f\,d\pi$, $f \in C_b(E)$, $f \geqslant 0$, can be represented as a Radon measure. Since E is a Polish space, it also follows that the measure π is σ-finite.

In order to complete the proof of (10.220) we have to verify the inequality in (10.221). By assuming that

$$\begin{aligned}
&\sup_{y\in A}\mathbb{P}_y\left[X(t)\notin K_\varepsilon \text{ for all } t\in[h, h+\tau_A\circ\vartheta_h)\right] \\
&= \sup_{y\in A}\mathbb{P}_y\left[\tau_\varepsilon\circ\vartheta_h \geqslant \tau_A\circ\vartheta_h\right] = 0 \qquad (10.235)
\end{aligned}$$

we will arrive at a contradiction. If (10.235) holds, then for all $y \in A$ we have

$$0 = \mathbb{P}_y\left[\tau_\varepsilon\circ\vartheta_h \geqslant \tau_A\circ\vartheta_h\right] = \mathbb{E}_y\left[\mathbb{P}_{X(h)}\left[\tau_\varepsilon \geqslant \tau_A\right]\right], \qquad (10.236)$$

and hence since all measure $B \mapsto P(h', y, B)$, $B \in \mathcal{E}$, $h' > 0$, are equivalent we infer from (10.236) that

$$\mathbb{P}_y\left[h' + \tau_\varepsilon \circ \vartheta_{h'} \geqslant h' + \tau_A \circ \vartheta_{h'}\right] = 0 \tag{10.237}$$

for all $h' > 0$. In (10.237) we let $h' \downarrow 0$ to obtain

$$\mathbb{P}_y\left[\tau_\varepsilon \geqslant \tau_A\right] = 0 \tag{10.238}$$

for all $y \in A$. Choose $y \in A^r$: since $X(\tau_A) \in A^r$ $\mathbb{P}_x$-almost surely on $\{\tau_A < \infty\}$ and A is recurrent such points y exist. Then $\tau_A = 0$ $\mathbb{P}_y$-almost surely. From (10.238) we get $\mathbb{P}_y[\tau_\varepsilon = 0] = 0$ which is manifestly a contradiction, because y is an interior point of K_ε.

The proof of (10.222) follows the same pattern as the corresponding proof by Seidler in [Seidler (1997)], who in turn follows Khasminskii [Has′minskiĭ (1960)]. Let τ be the first hitting time of K. Since K is non-recurrent there exists $y_0 \notin K$ such that

$$\mathbb{P}_{y_0}\left[\tau = \infty\right] = \mathbb{P}_{y_0}\left[X(t) \notin K \quad \text{for all } t \geqslant 0\right] > 0.$$

There is one other issue to be settled, i.e. is the subspace $R(L) + \mathbb{R}\mathbf{1}$ $\mathcal{T}_\beta$-dense in $C_b(E)$? Therefore we consider a $\mathcal{T}_\beta$-continuous linear functional $\Lambda : C_b(E) \to \mathbb{R}$ which annihilates the subspaces $R(L) + \mathbb{R}\mathbf{1}$. Suppose that $\Lambda \neq 0$. Then Λ can be represented as a measure on $\mathcal{E}$, and since $\Lambda(\mathbf{1}) = 0$ by scaling we may and will assume that $\Lambda(f)$ can be written as $\Lambda(f) = \int f d\mu_1 - \int f d\mu_2$, $f \in C_b(E)$, where μ_1 and μ_2 are probability measures on $\mathcal{E}$. Then, since $\int Lf\, d\mu_1 - \int Lf\, d\mu_2 = 0$, it follows that

$$\begin{aligned}\int_E f(x)\, d\mu_1(x) - \int_E f(x)\, d\mu_2(x) &= \int_E e^{nL} f(x)\, d\mu_1(x) - \int_E e^{nL} f(x)\, d\mu_2(x) \\ &= \iint\limits_{E\times E} \left(e^{nL} f(x) - e^{nL} f(y)\right) d\mu_1(x)\, d\mu_2(x),\ n \in \mathbb{N},\ f \in C_b(E). \end{aligned} \tag{10.239}$$

In (10.239) we let $n \to \infty$, and we use Orey's theorem to conclude that $\int_E f(x)\, d\mu_1(x) - \int_E f(x)\, d\mu_2(x) = 0$, $f \in C_b(E)$. It follows that $\Lambda(f) = 0$, $f \in C_b(E)$. Consequently, by the Hahn-Banach theorem we infer that the subspace $R(L) + \mathbb{R}\mathbf{1}$ is $\mathcal{T}_\beta$-dense in $C_b(E)$.

By construction and (10.222) it follows that for every compact subset K of E there exists a function $f_K \in C_b(E)$ such that $\mathbf{1}_K \leqslant f_K \leqslant \mathbf{1}_{K_\varepsilon}$ and $\int f_K d\pi < \infty$. Hence, the open subset $\{f_\alpha > 0\}$ has σ-finite π-measure. Let the sequence of open subsets $(U_n)_{n\in\mathbb{N}}$ be as in Remark 10.8. Consequently, each open subset U_n for which there exists a compact subset K_n with $U_n \subset \{f_{K_n} > 0\}$ has σ-finite π-measure. Since by Remark 10.8 such open subsets cover E, it follows that the measure π is σ-finite. This is another argument to show that the invariant e^{tL}-measure π is σ-finite. A previous argument was based on Lemma 10.11.

Altogether this completes the proof of Theorem 10.12. □

In the proof of Proposition 10.8 below we need the following lemma. The proof requires the equalities in (10.272) which are the same as those in (10.205) and (10.206).

Lemma 10.11. *Let A be a compact subset which is recurrent with first hitting time τ_A. Let π_E be any non-negative invariant Radon measure on $\mathcal{E}$. Then $\int_E \mathbb{P}_x\left[\tau_A \leqslant m\right] d\pi_E(x) < \infty$ for every $m \in \mathbb{R}$. Put*

$$C\left(\frac{h}{2}, \pi_E\right) = \frac{2}{h}\int_E \mathbb{P}_x\left[\tau_A \leqslant \frac{h}{2}\right] d\pi_E(x). \tag{10.240}$$

Moreover, for $0 < m < \infty$, and $\alpha > 0$ the following inequalities hold:

$$0 < \int_E \mathbb{P}_x\left[\tau_A \leqslant m\right] d\pi_E(x) \leqslant (m+h)C\left(\frac{h}{2}, \pi_E\right), \quad \text{and} \tag{10.241}$$

$$\alpha\int_E \mathbb{E}_x\left[e^{-\alpha\tau_A}\right] d\pi_E(x) \leqslant (\alpha h + 1)\, C\left(\frac{h}{2}, \pi_E\right). \tag{10.242}$$

Proof. Since A is compact and π_E is a Radon measure there exists a bounded continuous function f such that $\mathbf{1}_A \leqslant f \leqslant \mathbf{1}$, and such that $\int_E f\, d\pi_E < \infty$. The first equality in (10.272) yields:

$$\begin{aligned} &\left(e^{-h\lambda'}e^{hL} - I\right) R_A\left(\lambda'\right) f(x) + \int_0^h e^{-\lambda' s} e^{sL} f(x)\, ds \\ &= \mathbb{E}_x\left[e^{-\lambda'\tau_A}\mathbb{E}_{X(\tau_A)}\left[\int_0^{h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho\right], \tau_A \leqslant h\right], \end{aligned} \tag{10.243}$$

and so we get by invariance of the measure π_E:

$$\begin{aligned} &\int_0^h e^{-\lambda' s} ds \int_E f(x)\, d\pi_E(x) = \int_0^h e^{-\lambda' s}\int_E e^{sL} f(x)\, d\pi_E(x)\, ds \\ &= \int_E \left(I - e^{-h\lambda'}e^{hL}\right) R_A\left(\lambda'\right) f(x)\, d\pi_E(x) \\ &\quad + \int_E \mathbb{E}_x\left[e^{-\lambda'\tau_A}\mathbb{E}_{X(\tau_A)}\left[\int_0^{\tau_A^{0,h}} e^{-\lambda'\rho} f\left(X\left(\rho\right)\right)\, d\rho\right], \tau_A \leqslant h\right] d\pi_E(x) \\ &= \left(1 - e^{-\lambda' h}\right)\int_E R_A\left(\lambda'\right) f(x)\, d\pi_E(x) \end{aligned}$$

$$+ \int_E \mathbb{E}_x \left[e^{-\lambda' \tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{\tau_A^{0,h}} e^{-\lambda' \rho} f\left(X\left(\rho\right)\right)\, d\rho \right], \tau_A \leqslant h \right] d\pi_E(x)$$
$$\geqslant \int_E \mathbb{E}_x \left[e^{-\lambda' \tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{\frac{1}{2}h} e^{-\lambda' \rho} f\left(X\left(\rho\right)\right)\, d\rho \right], \tau_A \leqslant \frac{1}{2}h \right] d\pi_E(x) \tag{10.244}$$

where for brevity we wrote

$$\tau_A^{0,h}\left(\omega, \omega'\right) = h - \tau_A(\omega) + \tau_A \circ \vartheta_{h-\tau_A(\omega)}\left(\omega'\right) \tag{10.245}$$

which indicates the first time of hitting A strictly after $h - \tau_A(\omega)$. Notice that on the event $\left\{\tau_A \leqslant \frac{1}{2}h\right\}$ the inequalities $\tau_A^{0,h} \geqslant \frac{1}{2}h + \tau_A \circ \vartheta_{\frac{1}{2}h} \geqslant \frac{1}{2}h$ hold. In (10.244) we let $\lambda \uparrow 0$ to get:

$$h \int_E f(x)\, d\pi_E(x) \geqslant \int_E \mathbb{E}_x \left[\mathbb{E}_{X(\tau_A)} \left[\int_0^{\frac{1}{2}h} f\left(X\left(\rho\right)\right)\, d\rho \right], \tau_A \leqslant \frac{1}{2}h \right] d\pi_E(x)$$
$$\geqslant \inf_{y \in A} \mathbb{E}_y \left[\int_0^{\frac{1}{2}h} f\left(X\left(\rho\right)\right)\, d\rho \right] \cdot \int_E \mathbb{P}_x \left[\tau_A \leqslant \frac{1}{2}h \right] d\pi_E(x). \tag{10.246}$$

Assuming that $\inf_{y \in A} \mathbb{E}_y \left[\int_0^{\frac{1}{2}h} f\left(X\left(\rho\right)\right)\, d\rho \right] = 0$ leads to a contradiction, as we shall see momentarily. Since the function $y \mapsto \mathbb{E}_y \left[\int_0^{\frac{1}{2}h} f\left(X\left(\rho\right)\right)\, d\rho \right]$ and A is compact is continuous our assumption implies that for some $y_0 \in A$ the following inequality holds for all $0 < h' < \frac{1}{2}h$:

$$\mathbb{E}_{y_0} \left[\int_0^{h'} f\left(X\left(\rho\right)\right)\, d\rho \right] = \int_0^{h'} e^{\rho L} f\left(y_0\right) d\rho = 0. \tag{10.247}$$

Dividing all members of (10.247) by $h' > 0$, letting h' to 0, we obtain $f\left(y_0\right) = 0$. Here we employ the $\mathcal{T}_\beta$-continuity of the function $t \mapsto e^{tL} f\left(y_0\right)$ which follows from the $\mathcal{T}_\beta$-continuity of the semigroup $t \mapsto e^{tL}$. Since $\mathbf{1}_A \leqslant f \leqslant \mathbf{1}$ and $y_0 \in A$ we have a contradiction. Thus we have $\inf_{y \in A} \mathbb{E}_y \left[\int_0^{\frac{1}{2}h} f\left(X\left(\rho\right)\right)\, d\rho \right] > 0$. In combination with (10.246) this yields that $\int_E \mathbb{P}_x \left[\tau_A \leqslant \frac{1}{2}h\right] < \infty$. By induction with respect to k it follows that

$$\int_E \mathbb{P}_x \left[0 < \tau_A \leqslant \frac{1}{2}kh \right] d\pi_E(x) \leqslant k \int_E \mathbb{P}_x \left[0 < \tau_A \leqslant \frac{1}{2}h \right] d\pi_E(x), \tag{10.248}$$

and hence

$$\int_E \mathbb{P}_x \left[\tau_A \leqslant \frac{1}{2}kh \right] d\pi_E(x) \leqslant (k+1) \int_E \mathbb{P}_x \left[\tau_A \leqslant \frac{1}{2}h \right] d\pi_E(x) < \infty. \tag{10.249}$$

Let us show (10.248). Since on events of the form $\{\tau_A > s\}$ we have $s+\tau\circ\vartheta_s$ $\mathbb{P}_x$-almost surely, we have

$$\begin{aligned}
&\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}(k+1)h\right]\\
&=\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]+\mathbb{P}_x\left[\frac{1}{2}kh<\tau_A\leqslant \frac{1}{2}(k+1)h\right]\\
&=\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]+\mathbb{P}_x\left[0<\tau_A\circ\vartheta_{\frac{1}{2}kh}\leqslant \frac{1}{2}h,\ \frac{1}{2}kh<\tau_A\right]
\end{aligned}$$

(Markov property)

$$\begin{aligned}
&=\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]+\mathbb{E}_x\left[\mathbb{P}_{X\left(\frac{1}{2}kh\right)}\left[0<\tau_A\leqslant \frac{1}{2}h\right],\ \frac{1}{2}kh<\tau_A\right]\\
&\leqslant\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]+\mathbb{E}_x\left[\mathbb{P}_{X\left(\frac{1}{2}kh\right)}\left[0<\tau_A\leqslant \frac{1}{2}h\right]\right]. \qquad (10.250)
\end{aligned}$$

Since the positive measure π_E is e^{tL}-invariant from (10.250) we deduce

$$\begin{aligned}
&\int_E\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}(k+1)h\right]d\pi_E(x)\\
&\leqslant\int_E\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]d\pi_E(x)+\int_E\mathbb{E}_x\left[\mathbb{P}_{X\left(\frac{1}{2}kh\right)}\left[0<\tau_A\leqslant \frac{1}{2}h\right]\right]d\pi_E(x)\\
&=\int_E\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]d\pi_E(x)+\int_E e^{\frac{1}{2}khL}\mathbb{P}_{(\cdot)}\left[0<\tau_A\leqslant \frac{1}{2}h\right](x)\,d\pi_E(x)
\end{aligned}$$

(e^{tL}-invariance for $t=\frac{1}{2}kh$)

$$=\int_E\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}kh\right]d\pi_E(x)+\int_E\mathbb{P}_x\left[0<\tau_A\leqslant \frac{1}{2}h\right]d\pi_E(x). \qquad (10.251)$$

Thus (10.248) follows by induction from (10.251). The inequality in (10.241) follows from (10.249). Since

$$\mathbb{E}_x\left[e^{-\alpha\tau_A}\right]=\alpha\int_0^\infty\mathbb{P}_x\left[\tau_A\leqslant s\right]e^{-\alpha s}ds$$

the equality in (10.242) follows from (10.241). Suppose that the invariant measure π_E is non-trivial. Then there remains to show that the quantity $\int_E\mathbb{P}_x\left[\tau_A\leqslant m\right]d\pi_E(x)$ is strictly positive for $0<m<\infty$. For $m\uparrow\infty$ the quantity $\int_E\mathbb{P}_x\left[\tau_A\leqslant m\right]d\pi_E(x)$ increases to

$$\int_E\mathbb{P}_x\left[\tau_A\leqslant m\right]d\pi_E(x)\uparrow\int_E\mathbb{P}_x\left[\tau_A\leqslant \infty\right]d\pi_E(x)=\int_E\mathbf{1}\,d\pi_E>0. \qquad (10.252)$$

Assume, to arrive at a contradiction that, for some $m \in (0, \infty)$ the integral $\int_E \mathbb{P}_x [\tau_A \leqslant m] \, d\pi_E(x)$ vanishes. Then by invariance we have

$$\begin{aligned}
&\int_E \mathbb{P}_x [m < \tau_A \leqslant 2m] \, d\pi_E(x) \\
&= \int_E \mathbb{P}_x [m + \tau_A \circ \vartheta_m \leqslant 2m, \, \tau_A > m] \, d\pi_E(x) \\
&= \int_E \mathbb{E}_x [\tau_A \circ \vartheta_m \leqslant m, \, \tau_A > m] \, d\pi_E(x) \\
&\leqslant \int_E \mathbb{E}_x \left[\mathbb{P}_{X(m)} [\tau_A \circ \vartheta_m \leqslant m]\right] d\pi_E(x)
\end{aligned}$$

(the measure π_E is e^{mL}-invariant)

$$= \int_E \mathbb{P}_x [\tau_A \circ \vartheta_m \leqslant m] \, d\pi_E(x) = 0. \tag{10.253}$$

Repeating the arguments in (10.253) then shows the equality

$$\begin{aligned}
&\int_E \mathbb{P}_x [\tau_A < \infty] \, d\pi_E(x) \qquad (10.254) \\
&\leqslant \int_E \mathbb{P}_x [\tau_A \leqslant m] \, d\pi_E(x) + \sum_{k=0}^{\infty} \int_E \mathbb{P}_x [km < \tau_A \leqslant (k+1)m] \, d\pi_E(x) = 0,
\end{aligned}$$

which contradicts the non-triviality of the measure π_E. Finally, the conclusion in (10.249) completes the proof of Lemma 10.11. □

Proposition 10.7. *Let π_E be an invariant Radon measure. If the function $f \geqslant 0$ belongs to $f \in L^1(E, \mathcal{E}, \pi_E)$, then*

$$\begin{aligned}
&h \int_E f(x) \, d\pi_E(x) \\
&= \int_E \mathbb{E}_x \left[\mathbb{E}_{X(\tau_A)} \left[\int_0^{h - \tau_A + \tau_A \circ \vartheta_{h-\tau_A}} f(X(\rho)) \, d\rho \right], \tau_A \leqslant h \right] d\pi_E(x) \\
&= \int_E \mathbb{E}_x \left[\int_{\tau_A}^{h + \tau_A \circ \vartheta_h} f(X(\rho)) \, d\rho, \, \tau_A \leqslant h \right] d\pi_E(x), \qquad (10.255)
\end{aligned}$$

and

$$\lim_{\lambda' \downarrow 0} \lambda' \int_E R_A(\lambda') f(x) \, d\pi_E(x) = \inf_{\lambda' > 0} \lambda' \int_E R_A(\lambda') f(x) \, d\pi_E(x) = 0. \tag{10.256}$$

First assume that the function f is such that the function $R_A(0)f$ is uniformly bounded. Since the Markov process is irreducible this is true whenever f is replaced by a function of the form $f\mathbf{1}_U$ whenever U is an appropriate open neighborhood of a given compact subset: see (9.119) in Corollary 9.5.

Proof. Let $f \in L^1(E,\mathcal{E},\pi_E)\bigcap C_b(E)$, and let π_E be an e^{tL}-invariant Radon measure. For the proof we need the equality in (10.272). From that equality in conjunction with the invariance property of the measure π_E we obtain:

$$\begin{aligned}
&\left(e^{-h\lambda'}-1\right)\int_E R_A\left(\lambda'\right)f(x)\,d\pi_E(x)+\frac{1-e^{-\lambda' h}}{\lambda'}\int_E f(x)\,d\pi_E(x)\\
&=\int_E \mathbb{E}_x\left[e^{-\lambda'\tau_A}\mathbb{E}_{X(\tau_A)}\left[\int_0^{\tau_A^{0,h}} e^{-\lambda'\rho}f\left(X\left(\rho\right)\right)\,d\rho\right],\ \tau_A\leqslant h\right]d\pi_E(x)\\
&=\int_E \mathbb{E}_x\left[\int_{\tau_A}^{h+\tau_A\circ\vartheta_h} e^{-\lambda'\rho}f\left(X\left(\rho\right)\right)\,d\rho,\ \tau_A\leqslant h\right]d\pi_A(x), \qquad (10.257)
\end{aligned}$$

where $\tau_A^{0,h}=h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}$: see (10.245). Upon letting $\lambda'\downarrow 0$ we get

$$\begin{aligned}
&h\int_E f(x)\,d\pi_E(x)\\
&=h\lim_{\lambda'\downarrow 0}\lambda'\int_E R_A\left(\lambda'\right)f(x)\,d\pi_E(x)\\
&\quad+\int_E \mathbb{E}_x\left[\mathbb{E}_{X(\tau_A)}\left[\int_0^{\tau_A^{0,h}} f\left(X\left(\rho\right)\right)\,d\rho\right],\ \tau_A\leqslant h\right]d\pi_E(x)\\
&=h\lim_{\lambda'\downarrow 0}\lambda'\int_E R_A\left(\lambda'\right)f(x)\,d\pi_E(x)\\
&\quad+\int_E \mathbb{E}_x\left[\int_{\tau_A}^{h+\tau_A\circ\vartheta_h} f\left(X\left(\rho\right)\right)\,d\rho,\ \tau_A\leqslant h\right]d\pi_A(x). \qquad (10.258)
\end{aligned}$$

Next in (10.272) we let λ' tend to zero to obtain the pointwise equality:

$$\begin{aligned}
&\left(e^{hL}-I\right)R_A\left(0\right)f(x)+\int_0^h e^{sL}f(x)\,ds\\
&=\mathbb{E}_x\left[\mathbb{E}_{X(\tau_A)}\left[\int_0^{h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}} f\left(X\left(\rho\right)\right)\,d\rho\right],\ \tau_A\leqslant h\right]\\
&=\mathbb{E}_x\left[\int_{\tau_A}^{h+\tau_A\circ\vartheta_h} f\left(X\left(\rho\right)\right)\,d\rho,\ \tau_A\leqslant h\right]. \qquad (10.259)
\end{aligned}$$

From (10.258) and (10.259) we see that the function $\left(I - e^{hL}\right) R_A(0)f$ belongs to $L^1\left(E, \mathcal{E}, \pi_E\right)$, and that

$$\begin{aligned}\int_E \left(I - e^{hL}\right) R_A(0)f(x)\, d\pi_E(x) &= \lim_{\lambda'\downarrow 0} \lambda' \int_E R_A\left(\lambda'\right)\, d\pi_E(x) \\ &= \inf_{\lambda'>0} \lambda' \int_E R_A\left(\lambda'\right)\, d\pi_E(x). \qquad (10.260)\end{aligned}$$

The fact that in (10.256) and in (10.260) we may replace the limit by an infimum is due to the fact that the function $\lambda' \mapsto \lambda' \int_E R_A\left(\lambda'\right)\, d\pi_E(x)$ is decreasing. This claim follows from the resolvent property of the family $\left\{R_A\left(\lambda\right) : \lambda > 0\right\}$ and the invariance of the measure π_E. The arguments read as follows. Let $\lambda' > \lambda'' > 0$. Then by the resolvent equation we have:

$$\lambda' R_A\left(\lambda'\right) - \lambda'' R_A\left(\lambda''\right) = \left(\lambda' - \lambda''\right)\left(I - \lambda' R_A\left(\lambda'\right)\right) R_A\left(\lambda''\right). \qquad (10.261)$$

For $g \in L^1\left(E, \mathcal{E}, \pi_E\right)$, $g \geqslant 0$, we also have

$$\lambda' \int_E R_A\left(\lambda'\right) g(x)\, d\pi_E(x) \leqslant \lambda' \int_E R\left(\lambda'\right) g(x)\, d\pi_E(x) = \int_E g(x)\, d\pi_E(x). \qquad (10.262)$$

From (10.261) and (10.262) the monotonicity of the function

$$\lambda' \mapsto \lambda' \int_E R_A\left(\lambda'\right)\, d\pi_E(x)$$

easily follows. We shall prove that this limit vanishes, and consequently the result in (10.256) follows. Therefore, for $m > 0$ arbitrary, we consider the following decomposition of the function $\lambda R_A(\lambda)f(x)$:

$$\begin{aligned}\lambda R_A(\lambda)f(x) &= \lambda \mathbb{E}_x\left[\int_0^{\tau_A} e^{-\lambda\rho} f\left(X(\rho)\right)\, d\rho\right] \qquad (10.263) \\ &= \lambda \mathbb{E}_x\left[\int_0^{(\tau_A - m)\vee 0} e^{-\lambda\rho} f\left(X(\rho)\right)\, d\rho\right] + \lambda \mathbb{E}_x\left[\int_{(\tau_A - m)\vee 0}^{\tau_A} e^{-\lambda\rho} f\left(X(\rho)\right)\, d\rho\right].\end{aligned}$$

From the Markov property we infer

$$\begin{aligned}&\lambda \mathbb{E}_x\left[\int_0^{(\tau_A - m)\vee 0} e^{-\lambda\rho} f\left(X(\rho)\right)\, d\rho\right] \\ &= \lambda \mathbb{E}_x\left[\int_0^{(\tau_A - m)\vee 0} e^{-\lambda\rho} f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A > m\right]\, d\rho\right]. \qquad (10.264)\end{aligned}$$

We also infer, again using the Markov property,

$$\lambda \mathbb{E}_x\left[\int_{(\tau_A - m)\vee 0}^{\tau_A} e^{-\lambda\rho} f\left(X(\rho)\right)\, d\rho\right]$$

$$= \lambda \mathbb{E}_x \left[\int_{(\tau_A - m)\vee 0}^{\tau_A} e^{-\lambda\rho} f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A \leqslant m\right] d\rho \right]. \tag{10.265}$$

In both equalities (10.264) and (10.265) we used the $\mathbb{P}_x$-almost sure equality $\rho + \tau_A \circ \vartheta_\rho = \tau_A$ on the event $\{\tau_A > \rho\}$. Next we estimate the expression in (10.264):

$$\begin{aligned}
&\lambda \mathbb{E}_x \left[\int_0^{(\tau_A - m)\vee 0} e^{-\lambda\rho} f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A > m\right] d\rho \right] \\
&= \lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A > m\right], \tau_A > m \right] d\rho \\
&\leqslant \lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A > m\right] d\rho \right] \\
&= \lambda R(\lambda) \left(f(\cdot) \mathbb{P}_{(\cdot)}\left[\tau_A > m\right] \right)(x).
\end{aligned} \tag{10.266}$$

The expression in (10.265) can be rewritten and estimated as follows:

$$\begin{aligned}
&\lambda \mathbb{E}_x \left[\int_{(\tau_A - m)\vee 0}^{\tau_A} e^{-\lambda\rho} f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A \leqslant m\right] d\rho \right] \\
&= \lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A \leqslant m\right] \right] d\rho
\end{aligned}$$

(Markov property and $\rho + \tau_A \circ \vartheta_\rho = \tau_A$ on $\{\tau_A > \rho\}$ $\mathbb{P}_x$-almost surely)

$$\begin{aligned}
&= \lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[\mathbb{E}_{X(\rho)} \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) \right] f(X(\rho)) \mathbb{P}_{X(\rho)}[\tau_A \leqslant m], \tau_A > \rho \right] \\
&\qquad d\rho \\
&\leqslant \lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[\mathbb{E}_{X(\rho)} \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) \right] f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A \leqslant m\right] \right] d\rho.
\end{aligned} \tag{10.267}$$

Employing the invariance of the measure π_E in the inequality in (10.266) shows

$$\begin{aligned}
&\lambda \int_E \mathbb{E}_x \left[\int_0^{(\tau_A - m)\vee 0} e^{-\lambda\rho} f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A > m\right] d\rho \right] d\pi_E(x) \\
&= \lambda \int_E R(\lambda) \left(f(\cdot) \mathbb{P}_{(\cdot)}\left[\tau_A > m\right] \right)(x)\, d\pi_E(x) \\
&= \int_E f(x) \mathbb{P}_x\left[\tau_A > m\right] d\pi_E(x).
\end{aligned} \tag{10.268}$$

A similar estimate for the term in (10.265) is somewhat more involved, but it really uses the recurrence of the set A. Again using the invariance of the measure π_E for the expression in (10.267) yields:

$$\begin{aligned}
&\lambda \int_E \mathbb{E}_x \left[\int_{(\tau_A - m)\vee 0}^{\tau_A} e^{-\lambda\rho} f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A \leqslant m\right] d\rho \right] d\pi_E(x) \\
&\leqslant \lambda \int_0^\infty e^{-\lambda\rho} \int_E \mathbb{E}_x \left[\mathbb{E}_{X(\rho)} \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) \right] f\left(X(\rho)\right) \mathbb{P}_{X(\rho)}\left[\tau_A \leqslant m\right] \right] \\
&\qquad d\pi_E(x)\, d\rho \\
&= \lambda \int_0^\infty e^{-\lambda\rho} \int_E \mathbb{E}_x \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) \right] f\left(x\right) \mathbb{P}_x\left[\tau_A \leqslant m\right] d\pi_E(x)\, d\rho \\
&= \int_E \lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) \right] d\rho f\left(x\right) \mathbb{P}_x\left[\tau_A \leqslant m\right] d\pi_E(x) \\
&\leqslant \left(1 - e^{-\lambda m}\right) \int_E f\left(x\right) \mathbb{P}_x\left[\tau_A \leqslant m\right] d\pi_E(x).
\end{aligned} \tag{10.269}$$

In the final step of (10.269) we used the fact that $\tau_A < \infty$ $\mathbb{P}_x$-almost surely for all $x \in E$. As a consequence of this we have

$$\begin{aligned}
&\lambda \int_0^\infty e^{-\lambda\rho} \mathbb{E}_x \left[\mathbf{1}_{[(\tau_A - m)\vee 0, \tau_A)}(\rho) \right] d\rho \\
&= \mathbb{E}_x \left[\lambda \int_{(\tau_A - m)\vee 0}^{\tau_A} e^{-\lambda\rho} d\rho \right] \leqslant 1 - e^{-\lambda m}
\end{aligned}$$

showing the final step in (10.269). From (10.263), (10.268), and (10.269) we deduce:

$$\begin{aligned}
&\lambda \int_E R_A(\lambda) f(x)\, d\pi_E(x) \\
&\leqslant \int_E f(x) \mathbb{P}_x\left[\tau_A > m\right] d\pi_E(x) + \left(1 - e^{-\lambda m}\right) \int_E f\left(x\right) \mathbb{P}_x\left[\tau_A \leqslant m\right] d\pi_E(x).
\end{aligned} \tag{10.270}$$

Here $m > 0$ is arbitrary. Let $\varepsilon > 0$ be arbitrary. First we choose $m > 0$ so large that the first term in the right-hand side of (10.270) is $\leqslant \frac{1}{2}\varepsilon$. Then we choose $\lambda > 0$ so small that the second term in (10.270) is $\leqslant \frac{1}{2}\varepsilon$ as well. As a consequence we see that (10.256) in Proposition 10.7 follows. Together with (10.258), (10.259), and (10.259) this completes the proof of Proposition 10.7. □

In the following crucial proposition we establish a strong link between Q_A-invariant measures on A, and e^{tL}-invariant measures on E. In particular it follows that invariant measures on E are unique whenever this is the case on A.

Proposition 10.8. *Let the Borel probability measure π_A on A and the measure π_E on $\mathcal{E}$ be related as follows. For all functions $f \in L^1(E, \mathcal{E}, \pi_E)$ the equality*

$$\int_A \mathbb{E}_x \left[\int_0^{\tau_A^h} f(X(\rho))\, d\rho \right] d\pi_A(x) = h \int_E f\, d\pi_E \tag{10.271}$$

holds. Then the following assertions are true:

(a) Let $f \in C_b(E)$, and $\lambda' \geqslant 0$. The following equalities hold (see (10.205)):

$$\begin{aligned}&\left(e^{-h\lambda'} e^{hL} - I\right) R_A(\lambda') f(x) + \int_0^h e^{-\lambda' s} e^{sL} f(x)\, ds \\ &= \mathbb{E}_x \left[e^{-\lambda' \tau_A} \mathbb{E}_{X(\tau_A)} \left[\int_0^{h-\tau_A+\tau_A\circ\vartheta_{h-\tau_A}} e^{-\lambda'\rho} f(X(\rho))\, d\rho \right], \tau_A \leqslant h \right] \\ &= \mathbb{E}_x \left[\int_{\tau_A}^{h+\tau_A\circ\vartheta_h} e^{-\lambda'\rho} f(X(\rho))\, d\rho,\ \tau_A \leqslant h \right].\end{aligned} \tag{10.272}$$

(b) The measure π_A is Q_A-invariant if and only if π_E is e^{tL}-invariant for all $t \geqslant 0$.

(c) If the Q_A-invariant measure π_A on the Borel field of A is given, then (10.271) can be used to define the invariant measure π_E on $\mathcal{E}$.

(d) If the e^{tL}-invariant measure π_E on the Borel field of E is given, then (10.272) together with the equality (10.255) of Proposition 10.7 can be used to define the invariant measure π_A on the Borel field of A.

(e) If there exists only one Q_A-invariant probability measure π_A, then the e^{tL}-invariant measure π_E is unique up to multiplicative constants.

(f) If π_E is an invariant measure on E, and f belongs to $L^1(E, \mathcal{E}, \pi_E)$, then the following inequality holds:

$$h \left| \int_E f d\pi_E \right| \leqslant \sup_{y\in A} \mathbb{E}_y \left[\int_0^{h+\tau_A\circ\vartheta_h} |f(X(\rho))|\, d\rho \right] \int_E \mathbb{P}_x[\tau_A \leqslant h]\, d\pi_E(x). \tag{10.273}$$

Let π_E be a e^{tL}-invariant measure. Notice that, with $\lambda' = 0$, the equalities in (10.272) together with (10.271) entail the equality:

$$\begin{aligned}&\int_A \mathbb{E}_x \left[\int_0^{\tau_A^{1,h}} f(X(\rho))\, d\rho \right] d\pi_A(x) \\ &= \int_E \mathbb{E}_x \left[\int_{\tau_A}^{\tau_A^{1,h}} f(X(\rho))\, d\rho,\ \tau_A \leqslant h \right] d\pi_E(x) = h \int_E f\, d\pi_E.\end{aligned} \tag{10.274}$$

If $f \geqslant 0$ belongs to $C_b(E)$, and if π_E is a positive measure on $\mathcal{E}$, then we use the first equality in (10.274) to associate to π_E a Borel measure π_A on A. Since the invariant probability measures on A are unique, it follows that the invariant measures on E are unique as well.

Proof. [Proof of Proposition 10.8.] (a). The equalities in (10.272) follow from the equalities in (10.205) and (10.206).

(b). Let the measures π_A on A and π_E be related as in (10.271). Then for $t \geqslant 0$, and $f \in L^1(E, \mathcal{E}, \pi_E)$ we have

$$h \int_E e^{tL} f \, d\pi_E = \int_A \mathbb{E}_x \left[\int_0^{\tau_A^h} e^{tL} f(X(\rho)) \, d\rho \right] d\pi_A(x)$$
$$= \int_A \mathbb{E}_x \left[\int_0^{\tau_A^h} \mathbb{E}_{X(\rho)} \left[f(X(t)) \right] d\rho \right] d\pi_A(x)$$

(Markov property)

$$= \int_A \mathbb{E}_x \left[\int_0^{\tau_A^h} f(X(\rho + t)) \, d\rho \right] d\pi_A(x)$$
$$= \int_A \mathbb{E}_x \left[\int_t^{t+\tau_A^h} f(X(\rho)) \, d\rho \right] d\pi_A(x). \tag{10.275}$$

We differentiate both sides of (10.275) to obtain

$$h \int_E e^{tL} L f \, d\pi_E$$
$$= \int_A \mathbb{E}_x \left[f\left(X\left(t + \tau_A^h \right) \right) \right] d\pi_A(x) - \int_A \mathbb{E}_x \left[f(X(t)) \right] d\pi_A(x)$$
$$= \int_A Q_A e^{tL} f(x) \, d\pi_A(x) - \int_A e^{tL} f(x) \, d\pi_A(x). \tag{10.276}$$

By setting $t = 0$ in (10.276) we see that π_A is Q_A-invariant if and only if π_E is e^{tL}-invariant (or L-invariant). This proves assertion (a).

(c). Let π_A be a (finite) Borel measure on A, and define the measure π_E on $\mathcal{E}$ by the equality in (10.271). If π_A is an invariant measure on A, then by assertion (b) π_E is e^{tL}-invariant. This proves assertion (c).

(d). Let $M_E(\lambda')$ be the space of all continuous functions g on E for which there exists a function $f \in C_b(E)$ such that $g(x)$, $x \in E$, can be written as in (10.272). Let $M_A(\lambda')$ be the subspace of $C(A)$ consisting of functions $g \in M_E(\lambda')$ restricted to A. Let π_E be a Borel measure on

$\mathcal{E}$ which is a positive Radon measure with the property that for some finite constant C the inequality $\int_E g(x)\,d\pi_E(x) \leqslant C \sup_{x\in A} g(x)$ holds for all $g \in M_E(0)$. Notice that in case a function $g \in M_A(0)$ has two extensions g_1 and g_2 in $M_E(0)$, then $\int_E g_1(x)\,d\pi_E(x) = \int_E g_2(x)\,d\pi_E(x)$. Define the functional $\widetilde{\Lambda}_A : M_A(0) \to \mathbb{R}$ by $\widetilde{\Lambda}_A(g) = \int_E g(x)\,d\pi_E(x)$. Then by assumption $\widetilde{\Lambda}_A(g) \leqslant C \sup_{x\in A} g(x)$, $g \in M_E(0)$, and hence by the observation above $\widetilde{\Lambda}_A$ is well-defined. By the Hahn-Banach extension theorem in combination with the Riesz representation theorem there exists a measure π_A on the Borel field of A such that $\int_A g(x)\,d\pi_A(x) = \int_E g(x)\,d\pi_E(x)$, $g \in M_A(0)$, and $\int_A g(x)\,d\pi_A(x) \leqslant C \sup_{x\in A} g(x)$ for all $g \in C(A)$. Next, let π_E be any non-negative e^{tL}-invariant Radon measure on E. Then Lemma 10.11 implies $\int_E \mathbb{P}_x \left[\tau_A \leqslant h\right] d\pi_E(x) < \infty$.

(e). Let $\pi_E^{(1)}$ and $\pi_E^{(2)}$ be two Radon measures on $\mathcal{E}$ which are e^{tL}-invariant. Then the construction in (d) gives finite measures $\pi_A^{(1)}$ and $\pi_A^{(2)}$ on the Borel field of A such that the equality

$$\int_A \mathbb{E}_x \left[\int_0^{\tau_A^h} f\left(X(\rho)\right)\,d\rho \right] d\pi_A^{(j)}(x) = h \int_E f\,d\pi_E^{(j)} \tag{10.277}$$

is satisfied for all functions $f \in L^1\left(E, \mathcal{E}, \pi_E^{(j)}\right)$, $j = 1, 2$: see (10.271). Then (10.277) implies that the measures $\pi_A^{(1)}$ and $\pi_A^{(2)}$ are Q_A-invariant. By uniqueness, they are constant multiples of each other. It follows that the measures $\pi_E^{(1)}$ and $\pi_E^{(2)}$ are scalar multiples of each other.

This completes the proof of item (e).

(f). The inequality in (10.273) is a consequence of the first equality in (10.262) in Proposition 10.7, and the fact that $X(\tau_A) \in A$ $\mathbb{P}_x$-almost surely.

Altogether this completes the proof of Proposition 10.8. □

Let π_E be an invariant Borel measure on $\mathcal{E}$, let $f \geqslant 0$ be a function in $C_b(E)$, and introduce the functions f_α, $\alpha > 0$, by $f_\alpha(x) = f(x)\mathbb{E}_x\left[e^{-\alpha\tau_A}\right]$. In the following proposition we show that the functions $R_A(\alpha)f_\alpha$ are very appropriate to approximate functions of the form $R_A(0)f$. In many aspects they can be used to play the role of $R_A(\alpha)f$ for $\alpha > 0$ small. If f belongs to $C_b(E)$, $R_A(\alpha)f_\alpha$ is a member of $L^1\left(E, \mathcal{E}, \pi_E\right)$ where π_E is an invariant measure. This result follows from Lemma 10.11; in particular see (10.259).

Proposition 10.9. *In several aspects the functions $R_A(\alpha)f_\alpha$, $\alpha > 0$, $f \in C_b(E)$, have properties which are similar to those of the form $R_A(\alpha)f$, $f \in C_b(E)$, but with functions $f_\alpha \in L^1\left(E, \mathcal{E}, \pi_E\right) \bigcap C_b(E)$ if $f \in C_b(E)$. If $f \in L^1\left(E, \mathcal{E}, \pi_E\right)$, then the family $\{\alpha R_A(\alpha)f_\alpha : \alpha > 0\}$ is uniformly integrable.*

Proof. Let the functions f_α, $\alpha > 0$, be as above: $f_\alpha(x) = f(x)\mathbb{E}_x\left[e^{-\alpha\tau_A}\right]$. Observe that

$$\begin{aligned} R_A(\alpha)f_\alpha(x) &= \mathbb{E}_x\left[\int_0^{\tau_A} f\left(X(\rho)\right)e^{-\alpha\rho}\mathbb{E}_{X(\rho)}\left[e^{-\alpha\tau_A}\right]\,d\rho\right] \\ &= \int_0^\infty \mathbb{E}_x\left[f\left(X(\rho)\right)e^{-\alpha\rho}\mathbb{E}_{X(\rho)}\left[e^{-\alpha\tau_A}\right],\, \tau_A > \rho\right]\,d\rho \end{aligned}$$

(Markov property)

$$= \int_0^\infty \mathbb{E}_x\left[f\left(X(\rho)\right)e^{-\alpha\rho-\alpha\tau_A\circ\vartheta_\rho},\, \tau_A > \rho\right]\,d\rho$$

(τ_A is a terminal stopping time: $\rho + \tau_A \circ \vartheta_\rho = \tau_A$ on the event $\{\tau_A > \rho\}$)

$$= \mathbb{E}_x\left[\int_0^{\tau_A} f\left(X(\rho)\right)d\rho\, e^{-\alpha\tau_A}\right], \tag{10.278}$$

and consequently $\lim_{\alpha\downarrow 0} R_A(\alpha)f_\alpha(x) = R_A(0)f(x)$. Here we employed the recurrence of the set A. By Lemma 10.11 we also see that the functions $R_\alpha f_\alpha$ are members of $L^1\left(E,\mathcal{E},\pi_E\right)$:

$$\begin{aligned} &\alpha\int_E R_A(\alpha)f_\alpha(x)\,d\pi_E(x) \leqslant \alpha\int_E R(\alpha)f_\alpha(x)\,d\pi_E(x) \\ &= \int_E f_\alpha(x)\,d\pi_E(x) < \infty. \end{aligned} \tag{10.279}$$

Next suppose that $f \in L^1\left(E,\mathcal{E},\pi_E\right)$. In order to prove that the family $\{\alpha R_A(\alpha)f_\alpha : \alpha > 0\}$ is uniformly integrable, it suffices to take $f \geqslant 0$, and $f \in C_b(E)$. Then the result follows from (10.256) in Proposition 10.7, because for such functions f the function $\alpha \mapsto \alpha\int_E R_A(\alpha)f\,d\pi_E$ is monotone increasing. Moreover, the following pointwise limits are valid: $\lim_{\alpha\downarrow 0}\alpha\int_E R_A(\alpha)f = 0$, $\lim_{\alpha\to\infty}\alpha\int_E R_A(\alpha)f = \mathbf{1}_{E\setminus A^r}f$. In addition, by (10.256) in Proposition 10.7 we have

$$\lim_{\alpha\downarrow 0}\alpha\int_E R_A(\alpha)f\,d\pi = 0, \quad \text{and} \quad \lim_{\alpha\to\infty}\alpha\int_E R_A(\alpha)f\,d\pi = \int_{E\setminus A^r} f\,d\pi.$$

From Scheffé's theorem it then follows that the family $\{\alpha R_A(\alpha)f : \alpha > 0\}$ is uniformly integrable. Since $0 \leqslant f_\alpha \leqslant f$, it also follows that the family $\{\alpha R_A(\alpha)f_\alpha : \alpha > 0\}$ is uniformly π_E-integrable.

This completes the proof of Proposition 10.9. □

In an earlier version of the present work the following lemma was used in the proof of Theorem 10.12 which establishes the existence of an invariant measure.

Lemma 10.12. *Let A be a compact recurrent subset of E such that $A^r = A$, i.e. the collection of its regular points coincides with A itself. Put*

$$H_A g(x) = \mathbb{E}_x \left[g\left(\tau_A, X\left(\tau_A\right)\right)\right] = \mathbb{E}_x \left[g\left(\tau_A, X\left(\tau_A\right)\right), \tau_A < \infty\right]. \qquad (10.280)$$

Here $g : [0,\infty) \times E \to \mathbb{R}$ is any bounded continuous function. Then the following assertions hold true:

(a) *Suppose that for every such function g the limit $\lim_{\lambda\downarrow 0} \lambda R(\lambda) H_A g(x)$ exists uniformly on compact subsets of E. Then the family $\{\lambda R(\lambda) H_A : \lambda > 0\}$ is $\mathcal{T}_\beta$-equi-continuous. In particular, it follows that for every compact subset K in E there exists a function $v \in H\left([0,\infty)\times E\right)$ such that*

$$\sup_{x\in K}\sup_{\lambda>0} \left|\lambda R(\lambda) H_A g(x)\right| \leqslant \sup_{(s,x)\in[0,\infty)\times E} \left|v(s,x) g(s,x)\right|. \qquad (10.281)$$

(b) *Suppose that for every compact subset K of E the following equality holds:*

$$\inf_{u\in N(L), v\in D(L)} \sup_{x\in K}\sup_{\lambda>0} \left|\lambda R(\lambda)\left(H_A g - u - Lv\right)(x)\right| = 0. \qquad (10.282)$$

Let g be any function in $C_b\left([0,\infty)\times E\right)$. Then the limit

$$PH_A g(y) = \lim_{\lambda\downarrow 0} \lambda R(\lambda) H_A g(y) \qquad (10.283)$$

exists uniformly on compact subsets of E, and $PH_A g$ belongs to $N(L)$. Consequently $H_A g = PH_A g + (I-P) H_A g$ decomposes the function $H_A g$ into two functions $PH_A g \in N(L)$ and $(I-P) H_A g$ which belongs to $\mathcal{T}_\beta$-closure of $R(L)$.

(c) *Suppose that for every function $g \in C_b\left([0,\infty)\times E\right)$ the limit*

$$PH_A g(x) = \lim_{\lambda\downarrow 0} \lambda R(\lambda) H_A g(x) \quad \text{exists uniformly on compact subsets of } E. \qquad (10.284)$$

If $x_0 \in E$ and $h > 0$, then $\lim_{\lambda\to\infty} \lambda R(\lambda) \mathbb{P}_{(\cdot)}\left[\tau_A \leqslant h\right](x_0) > 0$.

Recall that a function v belongs to $H\left([0,\infty)\times E\right)$ provided that for every $\alpha > 0$ the subset $\{(s,x) \in [0,\infty)\times E : v(s,x) \geqslant \alpha\}$ is contained in a compact subset of $[0,\infty)\times$. In particular it follows that uniformly on any compact subset of E we have $\lim_{t\to\infty} v(t,x) = 0$.

Proof. (a). Let $(g_n : n \in \mathbb{N})$ be any sequence of functions in $C_b\left([0,\infty)\times E\right)$ which decreases to zero pointwise on $[0,\infty)\times E$. Then $H_A g_n$

decreases pointwise to 0 on E. By Dini's lemma it decreases to zero uniformly on compact subsets of E. Define the functions $G_n : [0,\infty] \times E \to \mathbb{R}$ by

$$G_n(\lambda, x) = \begin{cases} \lambda R(\lambda) H_A(x), & 0 < \lambda < \infty,\ x \in E; \\ \lim_{\alpha \downarrow 0} \alpha R(\alpha) H_A g_n(x), & \lambda = 0,\ x \in E; \\ \lim_{\alpha \to \infty} \alpha R(\alpha) H_A g_n(x) = H_A g_n(x), & \lambda = \infty,\ x \in E. \end{cases} \tag{10.285}$$

Then the sequence $(G_n : n \in \mathbb{N})$ defined in (10.285) consists of continuous functions which converges pointwise on $[0,\infty] \times E$ to zero. By Dini's lemma this convergence occurs uniformly on $[0,\infty] \times K$ where K is any compact subset of E. It follows that for all compact subsets K of E

$$\sup_{x \in K} \sup_{\lambda > 0} \lambda R(\lambda) H_A g_n(x) \quad \text{decreases to 0 as } n \text{ tends to } \infty.$$

By Corollary 2.3 in Chapter 2 it follows that such a family is $\mathcal{T}_\beta$-equicontinuous. The inequality in (10.281) is a consequence of this equicontinuity. For more details see §2.1.

(b). For $u \in N(L)$ and $\lambda > 0$ we have $\lambda R(\lambda) u = u$, and for $v \in D(L)$ we have $\lambda R(\lambda) L v = \lambda \left(\lambda R(\lambda) - I\right) v$. It follows that $\lim_{\lambda \downarrow 0} \lambda R(\lambda) \left(u + Lv\right) = u$ uniformly on E. By assumption (10.282) we see that $\lim_{\lambda \downarrow 0} \lambda R(\lambda) H_A g(x)$ exists uniformly on compact subsets of E. This shows assertion (b).

(c). Let K be a compact subset of E. By Assertion (a) there exists a function $v \in H\left([0,\infty) \times E\right)$ such that (10.281) is satisfied. In particular it follows that

$$\sup_{x \in K} \sup_{\lambda > 0} \left|\lambda R(\lambda) H_A g(x)\right| \leqslant \sup_{(s,x) \in [0,\infty) \times E} \left|v(s,x) g(s)\right| \tag{10.286}$$

for all functions $g \in C_b\left([0,\infty)\right)$. We may choose continuous functions g_m satisfying $\mathbf{1}_{[m,\infty)} \leqslant g_m \leqslant \mathbf{1}_{[m-1,\infty)}$. Then by (10.286) we have for $x \in K$ and $\lambda > 0$

$$\begin{aligned} \lambda R(\lambda) \mathbb{P}_{(\cdot)}\left[\tau_A > m\right](x) &\leqslant \lambda R(\lambda) \mathbb{E}_{(\cdot)}\left[g_m\left(\tau_A\right)\right](x) \\ &\leqslant \sup_{(s,y) \in [0,\infty) \times E} \left|v\left(s,y\right) g_m(s)\right|. \end{aligned} \tag{10.287}$$

From the properties of the functions v and g_m it follows that

$$\inf_{m \in \mathbb{N}} \sup_{x \in K} \sup_{\lambda > 0} \lambda R(\lambda) \mathbb{P}_{(\cdot)}\left[\tau_A > m\right](x) = 0. \tag{10.288}$$

As in (c) assume that for some $h > 0$ $\lim_{\lambda \downarrow 0} \lambda R(\lambda) \mathbb{P}_{(\cdot)}\left[\tau_A \leqslant h\right](x) = 0$. Then by the Markov property we also have

$$\lambda R(\lambda) \mathbb{P}_{(\cdot)}\left[h < \tau_A \leqslant 2h\right](x)$$

$$\begin{aligned}
&= \lambda \int_0^\infty e^{-\lambda s} e^{sL} \mathbb{P}_{(\cdot)} \left[\tau_A \circ \vartheta_h \leqslant h,\, \tau_A > h\right](x)\, ds \\
&= \lambda \int_0^\infty e^{-\lambda s} \mathbb{E}_x \left[\mathbb{E}_{X(s)} \left[\mathbb{P}_{X(h)} \left[\tau_A \leqslant h\right],\, \tau_A > h\right]\right] ds \\
&\leqslant \lambda \int_0^\infty e^{-\lambda s} \mathbb{E}_x \left[\mathbb{E}_{X(s)} \left[\mathbb{P}_{X(h)} \left[\tau_A \leqslant h\right]\right]\right] ds \\
&= \lambda \int_0^\infty e^{-\lambda s} \mathbb{E}_x \left[\mathbb{E}_{X(s+h)} \left[\tau_A \leqslant h\right]\right] ds \\
&= \lambda e^{\lambda h} \int_h^\infty e^{-\lambda s} \mathbb{E}_x \left[\mathbb{E}_{X(s)} \left[\tau_A \leqslant h\right]\right] ds. \qquad (10.289)
\end{aligned}$$

Hence by (10.289) and by assumption we see that

$$\lim_{\lambda \downarrow 0} \lambda R(\lambda) \mathbb{P}_{(\cdot)} \left[h < \tau_A \leqslant 2h\right](x) = 0.$$

Consequently, we obtain

$$\lim_{\lambda \downarrow 0} \lambda R(\lambda) \mathbb{P}_{(\cdot)} \left[\tau_A \leqslant m\right](x) = 0 \quad \text{for all } m > 0. \qquad (10.290)$$

Since the set A is recurrent we have for $m > 0$, $m \in \mathbb{N}$,

$$\begin{aligned}
1 &= \lambda R(\lambda) \mathbb{P}_{(\cdot)} \left[\tau_A < \infty\right](x) \\
&= \lambda R(\lambda) \mathbb{P}_{(\cdot)} \left[\tau_A \leqslant m\right](x) + \lambda R(\lambda) \mathbb{P}_{(\cdot)} \left[\infty > \tau_A > m\right](x). \qquad (10.291)
\end{aligned}$$

The second term in the right-hand side of (10.291) converges to 0 uniformly in $\lambda > 0$ when $m \to \infty$. For every fixed m the first term in the right-hind side of (10.291) converges to 0 when $\lambda \downarrow 0$: see (10.290). These two observations contradict the equality in (10.291). It follows that for every $x \in E$ and every $h > 0$ the limit $\lim_{\lambda \downarrow 0} \lambda R(\lambda) \mathbb{P}_{(\cdot)} \left[\tau_A \leqslant h\right](x) > 0$.

Altogether this completes the proof of Lemma 10.12. □

Corollary 10.5. *Let the hypotheses and notation be as in Theorem 10.12. Suppose that there exists a recurrent subset A subset such that*

$$\sup_{x \in A} \mathbb{E}_x \left[h + \tau_A \circ \vartheta_h\right] < \infty \quad \textit{for some } h > 0. \qquad (10.292)$$

Then the invariant measure constructed in the proof of Theorem 10.12 is finite. Here τ_A is the first hitting time of the subset A.

Proof. Corollary 10.5 follows from inequality (10.220) with $f = \mathbf{1}$ and the inequalities (10.207) and (10.208) with $2h$ instead of h in the proof of Theorem 10.12: see Definition 9.4 as well. □

Suppose that $x \in A^r$. Then the limits in (10.118) are in fact suprema, provided the numbers h are taken of the form $2^{-n}h'$, $h' > 0$ fixed, and $n \to \infty$. Moreover, the expression in (10.118) vanishes for $x \in A$. Notice that the $\mathcal{T}_\beta$-equi-continuity of the family $\{\lambda R(\lambda) : 0 < \lambda < 1\}$ is a consequence of the $\mathcal{T}_\beta$-equi-continuity of the family $\left\{t^{-1}\int_0^t e^{\rho L} d\rho : t > 0\right\}$. The latter is stronger than the standard condition in order that the Krylov-Bogoliubov theorem holds. More precisely, if there exists a probability measure ν such that some sequence $t_n^{-1}\int_0^{t_n}\left(e^{\rho L}\right)^* \nu\, d\rho$ weakly converges to a probability measure π, then π is L-invariant. For more details on the Krylov-Bogoliubov theorem see e.g. Theorem 2.1.1 in [Cerrai (2001)]; the reader might want to consult [Da Prato and Zabczyk (1996)] as well.

Proposition 10.10. *Let the (embedded) Markov chain $(X(n) : n \in \mathbb{N})$ be Harris recurrent. Then the strict closure, i.e. the $\mathcal{T}_\beta$-closure, of $R(L) + \mathbb{R}\mathbf{1}$ coincides with $C_b(E)$. If, in addition, the family $\{\lambda R(\lambda) : \lambda > 0\}$ is $\mathcal{T}_\beta$-equi-continuous, then the chain $(X(n) : n \in \mathbb{N})$ is positive Harris recurrent.*

Proof. Let $\mu = \mu_2 - \mu_1$ be a difference of positive Borel measures such that $\int Lf\, d\mu = 0$ for all $f \in D(L)$, and such that $\int \mathbf{1}\, d\mu = 0$. Then $\int \mathbf{1}\, d\mu_1 = \int \mathbf{1}\, d\mu_2$, and since the chain $(X(n) : n \in \mathbb{N})$ is Harris recurrent we know that

$$\lim_{\lambda \downarrow 0}\left(\int \lambda R(\lambda) f\, d\mu_2 - \int \lambda R(\lambda) f\, d\mu_1\right) = 0. \tag{10.293}$$

This is a consequence of Orey's theorem: see Theorem 10.2. Since $\int Lg\, d\mu = 0$ for all $g \in D(L)$, we have

$$\begin{aligned}\int f\, d\mu &= \int f\, d\mu_2 - \int f\, d\mu_1 \\ &= \lim_{\lambda \to 0}\left(\int \lambda R(\lambda) f\, d\mu_2 - \int \lambda R(\lambda) f d\mu_1\right) = 0. \end{aligned}\tag{10.294}$$

From (10.294) we conclude that $\mu = 0$. From the Hahn-Banach theorem it then follows that $R(L) + \mathbb{R}\mathbf{1}$ is $\mathcal{T}_\beta$-dense in $C_b(E)$.

If, in addition, the family $\{\lambda R(\lambda) : \lambda > 0\}$ is $\mathcal{T}_\beta$-equi-continuous, then we define the invariant measure π by

$$\int f\, d\pi = \lim_{\lambda \downarrow 0} \lambda R(\lambda) f\,(x_0)\,. \tag{10.295}$$

The limit in (10.295) exists for $f \in LD(L) + \mathbb{R}\mathbf{1}$, and for $f \in R(L)$ it vanishes. Since the chain $(X(n) : n \in \mathbb{N})$ is Harris recurrent we know that the limit in (10.295) does not depend on the choice of x_0. Since the family

$\{\lambda R(\lambda) : \lambda > 0\}$ is $\mathcal{T}_\beta$-equi-continuous, the limit in (10.295) also exists for f in the $C_b(E)$ which is the $\mathcal{T}_\beta$-closure of $R(L) + \mathbb{R}\mathbf{1}$. In addition this limit is a probability measure on E, again by this $\mathcal{T}_\beta$-equi-continuity. So the proof of Proposition 10.10 follows. □

Proposition 10.11. *Let the hypotheses and notation be as in Proposition 10.5. Let f belong to the domain of L, and suppose that*

$$Lf(x) = L\mathbb{E}_{(\cdot)}\left[f\left(X\left(\tau_A\right)\right)\right](x), \quad x \in A^r. \tag{10.296}$$

Then the function $x \mapsto \mathbb{E}_x\left[e^{-\lambda\tau_A} f\left(\tau_A\right)\right]$ belongs to the pointwise domain of L, and the following equalities hold:

$$\left(\lambda I - L\right)\mathbb{E}_{(\cdot)}\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right](x) = 0, \ x \in E\backslash A^r, \ \textit{and} \tag{10.297}$$

$$\lim_{\lambda\downarrow 0}\left\{\left(\lambda I - L\right)\mathbb{E}_{(\cdot)}\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right](x) - \left(\lambda I - L\right) f(x)\right\} = 0, \quad x \in A^r. \tag{10.298}$$

Note: neither equality (10.296) nor (10.298) is automatically satisfied. In fact condition (10.296) is in fact kind of Wentzell type boundary condition. Let $f \in C_b(E)$ be such that (10.296) is satisfied. Then the function $x \mapsto H_A(0)f(x) = \mathbb{E}_x\left[f\left(X\left(\tau_A\right)\right)\right]$ is a function which is L-harmonic on $E\backslash A^r$, it coincides with f on A^r, and in addition the functions Lf and $LH_A(0)f$ coincide on the same set. We introduce the Wentzell subspace $D\left(L_A^{\mathrm{W}}\right)$ of $D(L)$ by:

$$D\left(L_A^{\mathrm{W}}\right) = \left\{f \in D(L) : f \text{ satisfies equality (10.296)}\right\}. \tag{10.299}$$

Proof. [Proof of Proposition 10.11.] First we observe that for $x \in E\backslash A^r$ and g in the pointwise domain of L we have

$$\begin{aligned}\left(L - L_A\right) g(x) &= \lim_{t\downarrow 0}\left\{\frac{\mathbb{E}_x\left[g\left(X(t)\right)\right] - g(x)}{t} - \frac{\mathbb{E}_x\left[g\left(X(t)\right), \tau_A > t\right] - g(x)}{t}\right\} \\ &= \lim_{t\downarrow 0}\frac{\mathbb{E}_x\left[g\left(X(t)\right), \tau_A \leqslant t\right]}{t} = 0, \end{aligned} \tag{10.300}$$

where in the final step of (10.300) we employed Lemma 10.1. We also have

$$\begin{aligned}&\left(\lambda I - L_A\right) H_A(\lambda) f(x) = \left(\lambda I - L_A\right)\mathbb{E}_{(\cdot)}\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right](x) \\ &= \lim_{\delta\downarrow 0}\frac{\mathbb{E}_x\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right] - \mathbb{E}_x\left[e^{-\delta\lambda}\mathbb{E}_{X(\delta)}\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right], \tau_A > \delta\right]}{\delta}\end{aligned}$$

(employ the Markov property)

$$= \lim_{\delta\downarrow 0}\frac{\mathbb{E}_x\left[e^{-\lambda\tau_A} f\left(X\left(\tau_A\right)\right)\right] - \mathbb{E}_x\left[e^{-\delta\lambda - \lambda\tau_A\circ\vartheta_\delta} f\left(X\left(\delta + \tau_A\circ\vartheta_\delta\right)\right), \tau_A > \delta\right]}{\delta}$$

(on the event $\{\tau_A > \delta\}$ the equality $\delta + \tau_A \circ \vartheta_\delta = \tau_A$ holds)

$$= \lim_{\delta \downarrow 0} \frac{\mathbb{E}_x \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right), \tau_A \leqslant \delta \right]}{\delta} = 0, \quad \text{for } x \in E \backslash A^r. \tag{10.301}$$

In the final step of (10.301) we again used Lemma 10.1. An application of (10.300) and (10.301) to the function $g(x) = H_A(\lambda) f(x)$ shows the validity of (10.297) for $x \in E \backslash A^r$.

Next we treat the (important) case that $x \in A^r$. Since the process $t \mapsto e^{-\lambda t} f\left(X(t)\right) - f\left(X(0)\right) + \int_0^t e^{-\lambda s} f\left(X(s)\right) ds$ is $\mathbb{P}_x$-martingale, we get

$$\begin{aligned}
&\mathbb{E}_x \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right)\right] - f(x) \\
&= \mathbb{E}_x \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right) - f\left(X(0)\right)\right] \\
&= -\mathbb{E}_x \left[\int_0^{\tau_A} e^{-\lambda s} \left(\lambda I - L\right) f\left(X(s)\right) ds \right] \\
&= -\int_0^\infty e^{-\lambda s} \mathbb{E}_x \left[\left(\lambda I - L\right) f\left(X(s)\right), \tau_A > s \right] ds \\
&= -R_A(\lambda) \left(\lambda I - L\right) f(x).
\end{aligned} \tag{10.302}$$

We also have:

$$\begin{aligned}
&\left(\lambda I - L\right) \mathbb{E}_{(\cdot)} \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right) - f\left(X(0)\right)\right](x) \\
&= \left(Lf(x) - L\mathbb{E}_{(\cdot)} \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right)\right](x)\right) \mathbb{P}_x \left[\tau_A = 0\right].
\end{aligned} \tag{10.303}$$

In (10.303) we let λ tend to zero to obtain:

$$\begin{aligned}
&\lim_{\lambda \downarrow 0} \left(\lambda I - L\right) \mathbb{E}_{(\cdot)} \left[e^{-\lambda \tau_A} f\left(X\left(\tau_A\right)\right) - f\left(X(0)\right)\right](x) \\
&= \left(Lf(x) - L\mathbb{E}_{(\cdot)} \left[f\left(X\left(\tau_A\right)\right)\right](x)\right) \mathbb{P}_x \left[\tau_A = 0\right].
\end{aligned} \tag{10.304}$$

Here we use the fact that the subset A is recurrent, i.e. $\mathbb{P}_x\left[\tau_A < \infty\right] = 1$, $x \in E$. So the following equality remains to be shown:

$$Lf(x) = L\mathbb{E}_{(\cdot)} \left[f\left(X\left(\tau_A\right)\right)\right](x), \quad x \in A^r.$$

However, this is assumption (10.296), which completes the proof of Proposition 10.11. □

10.3 A proof of Orey's theorem

In this section we will prove Orey's convergence theorem as formulated in Theorem 10.2. We will employ the formulas (10.17) and (10.18). First we will define an accessible atom.

Definition 10.4. Let

$$\{(\Omega, \mathcal{F}, \mathbb{P}), (X(n), n \in \mathbb{N}), (\vartheta_n, n \in \mathbb{N}), (E, \mathcal{E})\}$$

be a time-homogeneous Markov process with a Polish state space E. Let $(x, B) \mapsto P(x, B)$ be the corresponding probability transition function. A Borel subset A is called an atom if $x \mapsto P(x, A)$, $x \in A$, does not depend on $x \in A$. It is called an accessible atom, if it is an atom such that $P(x, A) > 0$, $x \in A$.

Lemma 10.13. *Let A be an accessible atom and let x_1 and x_2 belong to A. Then the measures $\mathbb{P}_{x_1}$ and $\mathbb{P}_{x_2}$ coincide.*

Proof. Let F_n be a random variable of the form $F_n = \prod_{j=1}^n f_j(X(j))$ where the functions $f_j : E \to \mathbb{R}$, $1 \leqslant j \leqslant n$, are bounded non-negative Borel functions. By the monotone class theorem it suffices to prove the equality $\mathbb{E}_{x_1}[F_n] = \mathbb{E}_{x_2}[F_n]$. We will prove this equality by induction with respect to n. For $n = 1$, the equality $\mathbb{E}_{x_1}[F_1] = \mathbb{E}_{x_2}[F_1]$ follows from the definition of atom:

$$\mathbb{E}_{x_1}[F_1] = \int_0^\infty P(x_1, \{f_1 \geqslant \xi\})\, d\xi = \int_0^\infty P(x_2, \{f_1 \geqslant \xi\})\, d\xi = \mathbb{E}_{x_2}[F_1]. \tag{10.305}$$

Next we consider

$$\mathbb{E}_{x_1}[F_{n+1}] = \mathbb{E}_{x_1}\left[F_n \mathbb{E}_{x_1}\left[f_{n+1}(X(n+1)) \,\middle|\, \mathcal{F}_n\right]\right]$$

(Markov property)

$$= \mathbb{E}_{x_1}\left[F_n \mathbb{E}_{X(n)}\left[f_{n+1}(X(1))\right]\right]$$

(induction hypothesis)

$$= \mathbb{E}_{x_2}\left[F_n \mathbb{E}_{X(n)}\left[f_{n+1}(X(1))\right]\right]$$

(once again Markov property)

$$= \mathbb{E}_{x_2}[F_{n+1}]. \tag{10.306}$$

So from (10.305) and (10.306) the statement in Lemma 10.13 follows. □

Let A be an atom. Then we write $\mathbb{E}_A[F] = \mathbb{E}_x[F]$, $x \in A$. A similar notation is in vogue for $\mathbb{P}_A$. From (10.18) together with Lemma 10.13 we deduce the equality ($x \in E$, $f \in L^\infty(E, \mathcal{E})$)

$\mathbb{E}_x[f(X(n))]$

$$
\begin{aligned}
&= \mathbb{E}_x\left[f\left(X(n)\right), \tau_A^1 \geqslant n\right] \qquad (10.307)\\
&\quad + \sum_{j=1}^{n-1}\sum_{k=1}^{j} \mathbb{P}_x\left[\tau_A^1 = k\right]\mathbb{P}_A\left[X(j-k)\in A\right]\mathbb{E}_A\left[f\left(X(n-j)\right), \tau_A^1 \geqslant n-j\right].
\end{aligned}
$$

In addition we have

$$
\begin{aligned}
&\sum_{j=1}^{n-1}\mathbb{P}_A\left[X(j)\in A\right]\mathbb{E}_A\left[f\left(X(n-j)\right), \tau_A^1 \geqslant n-j\right]\\
&= \sum_{j=1}^{n-1}\mathbb{E}_A\left[\mathbb{E}_A\left[f\left(X(n-j)\right), \tau_A^1 \geqslant n-j\right], X(j)\in A\right]\\
&= \sum_{j=1}^{n-1}\mathbb{E}_A\left[\mathbb{E}_{X(j)}\left[f\left(X(n-j)\right), \tau_A^1 \geqslant n-j\right], X(j)\in A\right]
\end{aligned}
$$

(Markov property)

$$
\begin{aligned}
&= \sum_{j=1}^{n-1}\mathbb{E}_A\left[f\left(X(n)\right), j+\tau_A^1\circ\vartheta_j \geqslant n, X(j)\in A\right]\\
&= \mathbb{E}_A\left[f\left(X(n)\right)\right]. \qquad (10.308)
\end{aligned}
$$

Put

$$
\begin{aligned}
a_x(k) &= \mathbb{P}_x\left[\tau_A^1 = k\right], \quad u_A(k) = \mathbb{P}_A\left[X(k)\in A\right],\\
p_{A,f}(k) &= \mathbb{E}_A\left[f\left(X(k)\right), \tau_A^1 = k\right], \quad \text{and}\\
\overline{p}_{A,f}(k) &= \mathbb{E}_A\left[f\left(X(k)\right), \tau_A^1 \geqslant k\right]. \qquad (10.309)
\end{aligned}
$$

From (10.307), (10.308), and (10.309) we infer

$$
\begin{aligned}
&\mathbb{E}_x\left[f\left(X(n)\right)\right] - \mathbb{E}_A\left[f\left(X(n)\right)\right]\\
&= \mathbb{E}_x\left[f\left(X(n)\right), \tau_A^1 \geqslant n\right] + \left(a_x * u_A - u_A\right) * \overline{p}_{A,f}(n-1). \qquad (10.310)
\end{aligned}
$$

Definition 10.5. Let $n \mapsto p(n)$ be a probability distribution on $\mathbb{N}\backslash\{0\}$. Define the function $u : \mathbb{N}\bigcup\{-1\} \to [0,1]$ as in (10.322) in Theorem 10.14 below. Then the function u is called the renewal function of the distribution p.

The following proposition says that the function u_A is the renewal function corresponding to the distribution $p_{A,\mathbf{1}}$.

Proposition 10.12. *Let the functions $n \mapsto p_{A,\mathbf{1}}(n)$ and $n \mapsto u_A(n)$ be defined as in 10.309). Then the function u_A is the renewal function corresponding to the distribution $p_{A,\mathbf{1}}$.*

Proof. To see this we introduce the hitting times τ_A^k, k positive integer as follows: $\tau_A^{k+1} = \inf\left\{\ell > \tau_A^k : X(\ell) \in A\right\}$, with $\tau_A^0 = 0$. Then it is easy to show that $\tau_A^{k_1+k_2} = \tau_A^{k_1} + \tau_A^{k_2} \circ \vartheta_{\tau_{k_1}}$. Moreover, by the strong Markov property the variables $\tau_A^{k+1} - \tau_A^k = \tau_A^1 \circ \vartheta_A^k$ are identically $\mathbb{P}_A$-distributed, and $\mathbb{P}_A$-independent. Then the following identities hold:

$$\begin{aligned} u_A(n) &= \mathbb{P}_A\left[X(n) \in A\right] = \sum_{k=1}^{\infty} \mathbb{P}_A\left[\tau_A^k = n\right] \\ &= \sum_{k=1}^{\infty} \mathbb{P}_A\left[\sum_{j=1}^{k}\left(\tau_A^{j+1} - \tau_A^j\right) = n\right] = \sum_{k=1}^{\infty} \mathbb{P}_A\left[\sum_{j=1}^{k} \tau_A^1 \circ \vartheta_{\tau_A^j} = n\right] \\ &= \sum_{k=1}^{\infty} p_{A,\mathbf{1}}^{*k}. \end{aligned} \tag{10.311}$$

From (10.311) we see that the sequences $p_{A,\mathbf{1}}(n)$ and $u_A(n)$ are related as the sequences $p(n)$ and $u(n)$ in (10.322) of Theorem 10.14 below.

This completes the proof of Proposition 10.12. □

Then under appropriate conditions we will prove that every term in the right-hand side of (10.310) tends to 0 when $n \to \infty$. In order to obtain such a result we will use some renewal theory together with a coupling argument. Suppose that the atom A is recurrent and that the distribution $p(n) = p_{A,\mathbf{1}}(n) = \mathbb{P}_A\left[\tau_A^1 = k\right]$ is aperiodic, i.e. it satisfies (10.312). Then the right-hand side of (10.310) converges to zero when $n \to \infty$. This result is a consequence of Theorem 10.14 below.

We need the following lemma.

Lemma 10.14. *Let a, b and p be probability distributions on $\mathbb{N}$. Suppose that $p(0) = 0$ and p is aperiodic, i.e. suppose*

$$\text{g.c.d.}\left\{n \geqslant 1,\, n \in \mathbb{N},\, p(n) > 0\right\} = 1. \tag{10.312}$$

Let $\{S_0, S_1, S_2, \ldots\}$ and $\{S_0', S_1', S_2', \ldots\}$ be sequences of positive integer valued processes with the following properties:

(a). Each random variable S_j, and S_j', $j \geqslant 1$, has the same distribution $p(k)$.
(b). The variables S_0 and S_0' are independent: S_0 has distribution $a(k)$, and S_0' has distribution $b(k)$.
(c). The variables $\{S_0, S_1, S_2, \ldots\}$ are mutually independent, and the same is true for the sequence $\{S_0', S_1', S_2', \ldots\}$.
(d). The variables S_j and S_k' are independent for all j and $k \in \mathbb{N}$.

Let $\mathcal{G}_n$ be the σ-field generated by the couples $\{(S_0, S_0'), \dots, (S_n, S_n')\}$, and let $T(n)$ be the $\mathcal{G}_n$-stopping time defined

$$T(n) = \inf\left\{m \geqslant 0 : \sum_{j=0}^{m} S_j \geqslant n+1 \text{ and } \sum_{j=0}^{m} S_j' \geqslant n+1\right\}. \tag{10.313}$$

Let $n \mapsto V^+(n) = \left(V_a^+(n), V_b^+(n)\right)$ be the bivariate linked forward recurrence time chain which links the processes $n \mapsto V_a^+(n) = \sum_{j=0}^{T(n)} S_j - n$ and $n \mapsto V_b^+(n) = \sum_{j=0}^{T(n)} S_j - n$. Then the process $n \mapsto V^+(n)$ satisfies:

$$V^+(n+1) = \begin{cases} V^+(n) - (1,1), & \text{on } \left\{V_a^+(n) \geqslant 2\right\} \bigcap \left\{V_b^+(n) \geqslant 2\right\}, \\ V^+ n + \left(S_{1+T(n)} - 1, S_{1+T(n)}' - 1\right), & \text{on } \left\{V_a^+(n) = 1\right\} \bigcup \left\{V_b^+(n) = 1\right\}. \end{cases} \tag{10.314}$$

Let $P((i,j),(k,\ell))$, $((i,j),(k,\ell)) \in (\mathbb{N}\backslash\{0\})^2 \times (\mathbb{N}\backslash\{0\})^2$ be the probability transition function of the process $n \mapsto V^+(n)$. Then $P((i,j),(k,\ell))$ is given by

$$\begin{aligned} P((i,j),(i-1,j-1)) &= 1, & i > 1,\ j > 1; \\ P((1,j),(k,j-1)) &= p(k), & k \geqslant 1,\ j > 1; \\ P((i,1),(i-1,k)) &= p(k), & i > 1,\ k \geqslant 1; \\ P((1,1),(i,j)) &= p(i)p(j), & i > 1,\ j > 1, \end{aligned} \tag{10.315}$$

and the other transitions vanish. Put

$$\tau_{1,1} = \inf\left\{n \in \mathbb{N} : V^+(n) = (1,1)\right\}. \tag{10.316}$$

Then $\mathbb{P}\left[\tau_{1,1} < \infty\right] = 1$, and the following coupling equalities holds:

$$\sum_{j=0}^{T(\tau_{1,1})} S_j = \sum_{j=0}^{T(\tau_{1,1})} S_j' = \tau_{1,1} + 1. \tag{10.317}$$

As a consequence we have the following proposition.

Proposition 10.13. *Put $X^*(n) = \sum_{j=0}^{n} S_j - \sum_{j=0}^{n} S_j'$. The equality in (10.317) says that the process $n \mapsto X^*(n)$ returns to zero in a finite time $\tau^* = T(\tau_{1,1})$ with $\mathbb{P}$-probability 1, no matter what its initial distribution is. In other words the process $n \mapsto X^*(n)$ is recurrent.*

Proof. [Proof of Lemma 10.14.] Fix $(i,j) \in (\mathbb{N}\backslash\{0\}) \times (\mathbb{N}\backslash\{0\})$, and choose $M \in \mathbb{N}$ so large that

$$\text{g.c.d.}\,\{M \geqslant n \geqslant 1,\, n \in \mathbb{N},\, p(n) > 0\} = 1. \tag{10.318}$$

A number M for which (10.318) holds can be found using Bézout's identity.

Suppose that the distribution $n \mapsto p(n)$ has period d. Then there exist positive integers $s_j \geqslant 1$, $1 \leqslant j \leqslant N$, such that $p(s_j) > 0$, and such that g.c.d. $(s_1, \dots, s_N) = d$. Then for certain integers k_j, $1 \leqslant j \leqslant N$, we have $\sum_{j=1}^N k_j s_j = d$. By renumbering we may assume that $k_j \geqslant 1$ for $1 \leqslant j \leqslant N_1$, and $k_j \leqslant -1$ for $N_1 + 1 \leqslant j \leqslant N$. Then we choose $M \geqslant \sum_{j=1}^{N_1} s_j$. In fact one may consider the smallest integer $k \geqslant 1$ such that $k = \sum_{j=1}^N k_j s_j$, where $N \in \mathbb{N}$, $k_j \in \mathbb{Z}$, and $p(s_j) > 0$. Then one proves $k = d$, by using the fact that $\mathbb{Z}$ is a Euclidean domain. More precisely, let $k \geqslant 1$ be the smallest positive integer which can be written as $k = \sum_{j=1}^N k_j s_j$. Then we write $s_j = q_j k + r_j$ with $0 \leqslant r_j < k$ and $q_j \geqslant 0$. Then $r_j = s_j - q_j k \in R = \left\{\sum_{\ell=1}^N \ell_j s_j : \ell_j \in \mathbb{Z}\right\}$. Since $0 \leqslant r_j < k$ we infer $r_j = 0$. It follows that k is a divisor of s_j, $1 \leqslant j \leqslant N$. Since $d \in R$, d divides k. Since, in addition, g.c.d. $(s_1, \dots, s_N) = d$ we infer $k = d$. So we obtain Bézout's identity: $d = \sum_{j=1}^N k_j s_j$ for certain positive integers s_j with $p(s_j) > 0$ and certain integers k_j, $1 \leqslant j \leqslant N$.

If the sequence $\{s_j : p(s_j) > 0\}$ is aperiodic, then we choose $d = 1$ in the above remarks.

Fix $(i_0, j_0) \in (\mathbb{N}\backslash\{0\}) \times (\mathbb{N}\backslash\{0\})$, and choose M so large that (i_0, j_0) belongs to the square $\{1, \dots, M\} \times \{1, \dots, M\}$, and that $M \geqslant \sum_{j=1}^{N_1} k_j s_j$ where $1 = \sum_{j=1}^{N_1} k_j s_j - \sum_{j=N_1+1}^N (-k_j)\, s_j$ with $k_j \geqslant 1$, $1 \leqslant j \leqslant N_1$, and $-k_j \geqslant 1$, $N_1 + 1 \leqslant j \leqslant N$, in Bézout's identity. Then all paths in the square $\{1, \dots, M\} \times \{1, \dots, M\}$ along which each one-time transition is strictly positive, i.e. either 1 (along a diagonal from north-east to south-west) or $p(k) > 0$ from a point on one of the "edges" $\{(1,j) : 1 \leqslant j \leqslant M\}$ or $\{(i,1) : 1 \leqslant i \leqslant M\}$ of the square to the horizontal line $\{(k, j-1) : 1 \leqslant k \leqslant M\}$ or the vertical line $\{(i-1, k) : 1 \leqslant k \leqslant M\}$ respectively. Let $\tau_{1,1}$ be defined as is (10.316) with S_0 with distribution δ_i and S_0' with distribution j. By (10.318) $\mathbb{P}$-almost all paths pass through $(1,1)$ after a finite time passage, and consequently we obtain

$$\lim_{n\to\infty} \lim_{N'\to\infty} \mathbb{P}\left[\bigcup_{k=1}^{n} \{V^+(k) = (1,1)\} \;\Big|\; S_j \leqslant M,\, S_j' \leqslant M,\, 0 \leqslant j \leqslant N'\right]$$

$$= \lim_{n\to\infty} \lim_{N'\to\infty} \mathbb{P}\left[\bigcup_{k=1}^{n} \{\tau_{1,1} = k\} \;\Big|\; S_j \leqslant M,\, S_j' \leqslant M,\, 0 \leqslant j \leqslant N'\right] = 1. \tag{10.319}$$

Notice that the limit in (10.319), as N' tends to ∞, can be interpreted as the construction of the measure $\mathbb{P}$ conditioned on the event

$$\bigcap_{j=0}^{\infty} \left\{S_j \leqslant M,\, S_j' \leqslant M\right\}.$$

The existence of this "conditional probability" follows from Kolmogorov's extension theorem in conjunction with the assumption that for each $0 \leqslant j_1 < j_2$, $(j_1, j_2) \in \mathbb{N} \times \mathbb{N}$, the pairs $\left(S_{j_1}, S_{j_1}'\right)$ and $\left(S_{j_2}, S_{j_2}'\right)$ are $\mathbb{P}$-independent.

The collection of bounded paths along which the process $V^+(n)$ moves with strictly positive probability and which miss the diagonal throughout their life time eventually dy out, i.e. this event is negligible. The reason for this is that at each time step the transition probability of such a path is either 1 or else one of the quantities $p\left(s_j\right)$, $1 \leqslant j \leqslant N$, where N is the number occurring in Bézout's identity, and that the non-one transition probability occur infinitely many often. The $\mathbb{P}$-negligibility then follows from the theorem of dominated convergence. The other paths end up in $(1,1)$ in finite time. In (10.319) we let M tend to ∞ to obtain $\mathbb{P}\left[\tau_{1,1} < \infty\right] = 1$. But then we see

$$\mathbb{P}_{i_0,j_0}\left[\bigcup_{n=1}^{\infty} \left\{V^+(n) = (1,1)\right\}\right] = 1 \tag{10.320}$$

where $\mathbb{P}_{i_0,j_0}[A] = \mathbb{P}\left[A \;\big|\; (S_0, S_0') = (i_0, j_0)\right]$, $A \in \mathcal{F}$. Since the pair $(i_0, j_0) \in (\mathbb{N}\backslash\{0\}) \times (\mathbb{N}\backslash\{0\})$ is arbitrary from (10.320) we get

$$\sum_{i,j} a(i)b(j)\mathbb{P}_{i,j}\left[\bigcup_{n=1}^{\infty} \left\{V^+(n) = (1,1)\right\}\right] = 1. \tag{10.321}$$

If now $\tau_{1,1}$ is defined as in (10.316), then (10.321) implies $\mathbb{P}\left[\tau_{1,1} < \infty\right] = 1$.

This completes the proof of Lemma 10.14. □

Remark 10.9. For the equality in (10.319) see the argument in §10.3.1 of [Meyn and Tweedie (1993b)] as well. For Bézout's identity the reader is referred to e.g. [Tignol (2001)] or to Lemma D.7.3 in [Meyn and Tweedie (1993b)].

The following result appears as Theorem 18.1.1 in [Meyn and Tweedie (1993b)].

Theorem 10.14. *Let a, b and p be probability distributions on $\mathbb{N}$, and let $u : \mathbb{N}\bigcup\{-1\} \to [0,\infty]$ be the renewal function corresponding to $n \mapsto p(n)$, defined by $u(-1) = 0$, $u(0) = 1$, and for $n \geqslant 1$*

$$u(n) = \sum_{j=0}^{\infty} p^{j*}(n) = \delta_0(n)+p(n)+\sum_{j=2}^{\infty} \sum_{k_1,\dots,k_j;0\leqslant k_i\leqslant n;\sum_{i=1}^{j} k_i=n} p\left(k_1\right)\cdots p\left(k_j\right). \tag{10.322}$$

Suppose that p is aperiodic, i.e. suppose

$$\text{g.c.d.}\,\{n \geqslant 1,\, n \in \mathbb{N},\, p(n) > 0\} = 1. \tag{10.323}$$

Then

$$\lim_{n\to\infty} |a * u(n) - b * u(n)| = 0, \quad \textit{and} \tag{10.324}$$

$$\lim_{n\to\infty} |a * u(n) - b * u(n)| * \overline{p}(n) = 0, \tag{10.325}$$

where $\overline{p}(n) = \sum_{k\geqslant n+1} p(k)$.

In the proof of Theorem 10.2 the result in Theorem 10.14 will be applied with $a(k) = \mathbb{P}_x\left[\tau_A^1 = k\right]$, $p(k) = p_{A,\mathbf{1}}(k) = \mathbb{P}_A\left[\tau_A^1 = k\right]$, $b(k) = \delta_0(k)$, and, consequently, $u(k) = \mathbb{P}_A\left[X(k) \in A\right] = u_A(k)$. Notice that $k \mapsto u_A(k)$ is the renewal function of the distribution $p_{A,\mathbf{1}}(k)$.

We follow the proof Theorem 18.1.1 in [Meyn and Tweedie (1993b)].

Proof. Let $\{S_0, S_1, S_2, \dots\}$ and $\{S_0', S_1', S_2', \dots\}$ be sequences of positive integer valued processes with the properties (a), (b), (c) and (d) of Lemma 10.14:

(a). Each random variable S_j, and S_j', $j \geqslant 1$, has the same distribution $p(k)$.
(b). The variables S_0 and S_0' are independent: S_0 has distribution $a(k)$, and S_0' has distribution $b(k)$.
(c). The variables $\{S_0, S_1, S_2, \dots\}$ are mutually independent, and the same is true for the sequence $\{S_0', S_1', S_2', \dots\}$.
(d). The variables S_j and S_k' are $\mathbb{P}$-independent for all pairs $(j,k) \in \mathbb{N}\times\mathbb{N}$.

We put $W_j = S_j - S_j'$, and $X^*(n) = \sum_{j=0}^{n}\left(S_j - S_j'\right)$. Notice that the variables W_j and $-W_j$, $j \in \mathbb{N}$, $j \geqslant 1$, have the same distributions. The distribution of $W_0 = S_0 - S_0'$ is determined by the distributions a of S_0 and

b of S'_0, and the fact that S_0 and S'_0 are independent. We also introduce the indicator variables $Z_a(n)$ and $Z_b(n)$, $n \in \mathbb{N}$:

$$Z_a(n) = \begin{cases} 1 & \text{if } \sum_{i=0}^{j} S_i = n \text{ for some } j \geqslant 0; \\ 0 & \text{elsewhere.} \end{cases} \tag{10.326}$$

Hence $Z_a(n) = \mathbf{1}_{\bigcup_{j=0}^{\infty}\{\sum_{i=0}^{j} S_i = n\}}$. The indicator process $Z_b(n)$ is defined similarly, but with S'_j instead of S_j. Then $\mathbb{P}[Z_a(n) = 1] = a * u(n)$, and $\mathbb{P}[Z_b(n) = 1] = b * u(n)$. The coupling time of the renewal processes is defined by

$$T_{a,b} = \min\left\{ n = \sum_{i=0}^{j} S_i = \sum_{i=0}^{j} S'_i \in \mathbb{N} : n \geqslant 1, \text{ for some } j \in \mathbb{N} \right\}. \tag{10.327}$$

We also have

$$T_{a,b} = \min\left\{ \sum_{i=0}^{j} S_i : j \geqslant 1,\, X^*(j) = 0 \right\}. \tag{10.328}$$

Let $T^*_{a,b}$ be defined by $T^*_{a,b} = \inf\{j \geqslant 1 : X^*(j) = 0\}$. Then $T_{a,b} = \sum_{j=0}^{T^*_{a,b}} S_j = \sum_{j=0}^{T^*_{a,b}} S'_j$. From Proposition 10.13 it follows that the coupling time $T_{a,b}$ is finite $\mathbb{P}$-almost surely. Based on this property we will prove the equalities in (10.324) and (10.325). Therefore we put

$$Z_{a,b}(n) = \begin{cases} Z_a(n), & \text{if } n < T_{a,b}; \\ Z_b(n), & \text{if } n \geqslant T_{a,b}. \end{cases} \tag{10.329}$$

Then we have

$$\begin{aligned}
&|a * u(n) - b * u(n)| \\
&= |\mathbb{P}[Z_a(n) = 1] - \mathbb{P}[Z_b(n) = 1]| \\
&= |\mathbb{P}[Z_{a,b}(n) = 1] - \mathbb{P}[Z_b(n) = 1]| \\
&= |\mathbb{P}[Z_{a,b}(n) = 1,\, T_{a,b} > n] + \mathbb{P}[Z_{a,b}(n) = 1,\, T_{a,b} \leqslant n] \\
&\quad -\mathbb{P}[Z_b(n) = 1,\, T_{a,b} > n] - \mathbb{P}[Z_b(n) = 1,\, T_{a,b} \leqslant n]| \\
&= |\mathbb{P}[Z_a(n) = 1,\, T_{a,b} > n] + \mathbb{P}[Z_b(n) = 1,\, T_{a,b} \leqslant n] \\
&\quad -\mathbb{P}[Z_b(n) = 1,\, T_{a,b} > n] - \mathbb{P}[Z_b(n) = 1,\, T_{a,b} \leqslant n]| \\
&\leqslant \max\left(\mathbb{P}[Z_a(n) = 1,\, T_{a,b} > n], \mathbb{P}[Z_b(n) = 1,\, T_{a,b} > n]\right) \\
&\leqslant \mathbb{P}[T_{a,b} > n].
\end{aligned} \tag{10.330}$$

Since $\mathbb{P}[T_{a,b} < \infty] = 1$, the inequality in (10.330) yields the equality in (10.324). Next we consider the backward recurrence chains $V_a^-(n)$ and

$V_b^-(n)$ for the renewal processes of the sequences $\{S_0, S_1, S_2, \dots\}$ and $\{S'_0, S'_1, S'_2, \dots\}$ defined by respectively:

$$\begin{aligned} V_a^-(n) &= \min\left\{n - \sum_{j=0}^k S_j : \sum_{j=0}^k S_j \leqslant n\right\} \\ &= \min\left\{n - \sum_{j=0}^k S_j : \sum_{j=0}^k S_j \leqslant n < \sum_{j=0}^{k+1} S_j\right\}, \end{aligned}$$

and

$$\begin{aligned} V_b^-(n) &= \min\left\{n - \sum_{j=0}^k S'_j : \sum_{j=0}^k S'_j \leqslant n\right\} \\ &= \min\left\{n - \sum_{j=0}^k S'_j : \sum_{j=0}^k S'_j \leqslant n < \sum_{j=0}^{k+1} S'_j\right\}. \end{aligned} \tag{10.331}$$

It follows that there exists a random non-negative integer $K_a(n)$ which satisfies $\sum_{j=0}^{K_a(n)} S_j \leqslant n < n+1 \leqslant \sum_{j=0}^{K_a(n)+1} S_j$, and hence $V_a^-(n) = n - \sum_{j=0}^{K_a(n)} S_j$. For the moment fix $0 \leqslant m \leqslant n$. Since the variables $\{S_0, S_1, S_2, \dots\}$ are mutually independent, S_0 has distribution $a(k)$, and the others have distribution $p(k)$ we have

$$\begin{aligned} \mathbb{P}\left[V_a^-(n) = m\right] &= \sum_{k=0}^{\infty} \mathbb{P}\left[n - \sum_{j=0}^k S_j = m, \sum_{j=0}^{k+1} S_j \geqslant n+1\right] \\ &= \sum_{k=0}^{\infty} \mathbb{P}\left[\sum_{j=0}^k S_j = n-m, S_{k+1} \geqslant m+1\right] \\ &= \sum_{k=0}^{\infty} \mathbb{P}\left[\sum_{j=0}^k S_j = n-m\right] \mathbb{P}\left[S_{k+1} \geqslant m+1\right] \\ &= \sum_{k=0}^{\infty} a * p^{*k}(n-m)\overline{p}(m) \end{aligned} \tag{10.332}$$

where, with a notation we employed earlier, $\overline{p}(m) = \sum_{j=m+1}^{\infty} p(j)$. Of course, for the process $V_b^-(n)$ we have a similar distribution with b instead of a. From (10.332) and a similar expression for $\mathbb{P}\left[V_b^-(n) = m\right]$ we also infer

$$\begin{aligned} &\sup_{A \subset \mathbb{N}} \left|\mathbb{P}\left[V_a^-(n) \in A\right] - \mathbb{P}\left[V_a^-(n) \in A\right]\right| \\ &= \frac{1}{2} \sum_{m=0}^{\infty} \left|\mathbb{P}\left[V_a^-(n) = m\right] - \mathbb{P}\left[V_a^-(n) = m\right]\right| \end{aligned}$$

$$= \frac{1}{2} \sum_{m=0}^{n} |a * u(n-m)\overline{p}(m) - b * u(n-m)\overline{p}(m)|$$
$$= \frac{1}{2} |a * u - b * u| * \overline{p}(n). \tag{10.333}$$

It also follows that on the event $A_{a,b}(n)$ defined by

$$A_{a,b}(n) = \left\{ T_{a,b} = \sum_{j=0}^{T^*_{a,b}} S_j \leqslant n \right\}$$

the $\mathbb{P}$-distributions of $V_a^-(n)$ and $V_b^-(n)$ coincide. This is a consequence of the strong Markov property of the process $\left\{ \left(\sum_{i=0}^{j} S_j, \sum_{i=0}^{j} S'_j \right) : j \in \mathbb{N} \right\}$:

$$\begin{aligned}
&\mathbb{P}\left[V_a^-(n) \in A,\ A_{a,b}(n)\right] \\
&= \sum_{k=0}^{\infty} \mathbb{P}\left[n - \sum_{j=0}^{T^*_{a,b}+k} S_j \in A,\ \sum_{j=0}^{T_{a,b}+k} S_j \leqslant n < \sum_{j=0}^{T^*_{a,b}+k+1} S_j \right] \\
&= \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbb{P}\left[n - \sum_{j=0}^{T^*_{a,b}+k} S_j \in A,\ \sum_{j=0}^{T^*_{a,b}+k} S_j \leqslant n < \sum_{j=0}^{T^{|ast}_{a,b}+k+1} S_j \;\Big|\; \mathcal{G}_{T_{a,b}} \right]\right]
\end{aligned}$$

(strong Markov property together with the definition of $T_{a,b} = \sum_{i=0}^{T^*_{a,b}} S_j$, and the fact that the variables S_j and S'_j, $j \geqslant 1$, heve the same distribution)

$$\begin{aligned}
&= \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbb{P}\left[n - \sum_{j=0}^{T^*_{a,b}+k} S'_j \in A,\ \sum_{j=0}^{T^*_{a,b}+k} S'_j \leqslant n < \sum_{j=0}^{T^*_{a,b}+k+1} S'_j \;\Big|\; \mathcal{G}_{T^*_{a,b}} \right]\right] \\
&= \sum_{k=0}^{\infty} \mathbb{P}\left[n - \sum_{j=0}^{T^*_{a,b}+k} S'_j \in A,\ \sum_{j=0}^{T_{a,b}+k} S'_j \leqslant n < \sum_{j=0}^{T^*_{a,b}+k+1} S'_j \right] \\
&= \sum_{k=0}^{\infty} \mathbb{P}\left[n - \sum_{j=0}^{k} S'_j \in A,\ \sum_{j=0}^{k} S'_j \leqslant n < \sum_{j=0}^{k+1} S'_j,\ \sum_{j=0}^{T^*_{a,b}} S'_j \leqslant n \right] \\
&= \mathbb{P}\left[V_b^-(n) \in A,\ A_{a,b}(n)\right].
\end{aligned} \tag{10.334}$$

Here we wrote $\mathcal{G}_n = \sigma\left((S_j, S'_j) : 0 \leqslant j \leqslant n \right)$, and

$$\mathcal{G}_{T^*_{a,b}} = \bigcap_{n=0}^{\infty} \left\{ A \in \mathcal{G} : A \bigcap \left\{ T^*_{a,b} \leqslant n \right\} \in \mathcal{G}_n \right\}.$$

From (10.334) we infer

$$\begin{aligned}
&\left|\mathbb{P}\left[V_a^-(n)\in A\right]-\mathbb{P}\left[V_b^-(n)\in A\right]\right| \\
&=\left|\mathbb{P}\left[V_a^-(n)\in A, A_{a,b}(n)\right]+\mathbb{P}\left[V_a^-(n)\in A, \Omega\backslash A_{a,b}(n)\right]\right. \\
&\quad \left.-\mathbb{P}\left[V_b^-(n)\in A, A_{a,b}(n)\right]-\mathbb{P}\left[V_b^-(n)\in A, \Omega\backslash A_{a,b}(n)\right]\right| \\
&=\left|\mathbb{P}\left[V_a^-(n)\in A, \Omega\backslash A_{a,b}(n)\right]-\mathbb{P}\left[V_b^-(n)\in A, \Omega\backslash A_{a,b}(n)\right]\right| \\
&\leqslant \mathbb{P}\left[\Omega\backslash A_{a,b}(n)\right] =\leqslant \mathbb{P}\left[\sum_{j=0}^{T_{a,b}^*} S_j \geqslant n\right]. \qquad (10.335)
\end{aligned}$$

From (10.333), (10.334) and (10.335) we deduce

$$\left|a*u-b*u\right|*\overline{p}(n)\leqslant 2\mathbb{P}\left[\sum_{j=0}^{T_{a,b}^*} S_j\geqslant n\right]. \qquad (10.336)$$

Since by Proposition 10.13 the process $X^*(n)$ is recurrent, and hence

$$\mathbb{P}\left[T_{a,b}=\sum_{j=0}^{T_{a,b}^*} S_j<\infty\right]=1,$$

it follows from (10.336) that $\lim_{n\to\infty}\left|a*u-b*u\right|*\overline{p}(n)=0$. However, this is the same as equality (10.325).

This completes the proof of Theorem 10.14 □

Before we complete the proof of 10.2 we insert some definitions which are taken from [Meyn and Tweedie (1993b)]. Let $(X(n),\mathbb{P}_{x\in E})$ be a Markov chain with the property that all measures $B\mapsto P(1,x,B)=\mathbb{P}_x\left[X(n)\in B\right]$, $x\in E$, are equivalent. Fix $x_0\in E$. We say that the Markov chain is recurrent, if for all subsets $B\in\mathcal{E}$ with $P(1,x_0,B)>0$ we have $\mathbb{P}_x\left[\tau_B^1<\infty\right]>0$.

Definition 10.6. A subset $C\in\mathcal{E}$ is called small if there exists $m\in\mathbb{N}$, $m\geqslant 1$, and a non-trivial positive Borel measure ν_m such that the inequality

$$P(m,x,B)\geqslant\nu_m(B) \qquad (10.337)$$

holds for $x\in C$ and all $B\in\mathcal{E}$.

The following definition also occurs in formula (10.11) in Definition 10.1.

Definition 10.7. A Markov chain $\{X(n),\mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ is called aperiodic if there exists no partition of $E=(D_0,D_1,\dots,D_{p-1})$ for some $p\geqslant 2$ such that for all $i\in\mathbb{N}$

$$P\left[X(i)\in D_{i \bmod (p)}\middle|X(0)\in D_0\right]=\int_{D_0}\mathbb{P}_x\left[X(i)\in D_{i \bmod (p)}\right]d\mu_0(x)=1, \qquad (10.338)$$

for some initial probability distribution μ_0.

A Markov chain $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ having initial distribution μ_0 is called periodic if there exists $p \geqslant 2$ and a partition $E = (D_0, D_1, \dots, D_{p-1})$ such that (10.338) holds. The largest d for which (10.338) holds is called the period of the Markov chain.

Let $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ be an aperiodic $P((1, x_0, \cdot)$-irreducible Markov chain. If there exists a ν_1-small set A with $\nu_1(A) > 0$, then the Markov chain $\{X(n), \mathbb{P}_{x\in E}\}_{n\in\mathbb{N},x\in E}$ is called strongly aperiodic.

The following theorem is proved in [Meyn and Tweedie (1993b)]: see theorems 5.2.1 and 5.2.2.

Theorem 10.15. *Let $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ be a $P(1, x_0, \cdot)$-irreducible Markov chain. Then for any $A \in \mathcal{E}$ with $P(1, x_0, A) > 0$ there exists $m \in \mathbb{N}$, $m \geqslant 1$, together with a ν_m-small set $C \subset A$ with $P(1, x_0, C) > 0$ such that $\nu_m(C) > 0$.*

Remark 10.10. Suppose that all measures $B \mapsto P(1, x, B)$, $B \in \mathcal{E}$, $x \in E$, are equivalent, then analyzing the proof of Theorem 5.2.1 in [Meyn and Tweedie (1993b)] shows that in Theorem 10.15 we may choose $m = 3$.

The following corollary is an immediate consequence of Definition 10.7 and Theorem 10.15.

Corollary 10.6. *Let $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ be a $P(1, x_0, \cdot)$-irreducible aperiodic Markov chain. Then there exists $m \in \mathbb{N}$ such that the skeleton chain $\{X(mn), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ is strongly aperiodic, and $P(1, x_0, \cdot)$-irreducible.*

Remark 10.11. In fact the skeleton chain $\{X(mn), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ is $P(m, x_0, \cdot)$-irreducible, provided that the chain $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ is also $P(1, x_0, \cdot)$-irreducible, and all measures of the form $B \mapsto P(1, x_0, B) = 0$ are equivalent, i.e. have the same negligible sets. Suppose that $B \in \mathcal{E}$ is such that $P(m, x_0, B) = 0$. Then

$$0 = P(m, x_0, B) = \int P(m-1, x_0, dy)\, P(1, y, B)\,. \tag{10.339}$$

From (10.339) we see that $P(1, y, B) = 0$ for $P(m-1, x_0, \cdot)$-almost all $y \in E$. Since $P(m-1, x_0, E) = 1$, it follows that $P(1, y, B) = 0$ for at least one $y \in E$. But then $P(1, x_0, B) = 0$ because all measures of the form $B \mapsto P(1, x_0, B) = 0$ are equivalent.

The following theorem is a consequence of Proposition 9.5, Lemma 9.4, and Theorem 10.15.

Theorem 10.16. *Let*

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(n), n \in \mathbb{N}), (\vartheta_n, n \in \mathbb{N}), (E, \mathcal{N})\} \tag{10.340}$$

be a Markov chain with the property that all Borel measures $B \mapsto P(1, x, B) = \mathbb{P}_x[X(1) \in B]$, $x \in E$, *are equivalent. In addition suppose that for every Borel subset* B *the function* $x \mapsto P(1, x, B)$ *is continuous. Let there exist a point* $x_0 \in E$ *such that every open neighborhood of* x_0 *is recurrent. Then there exists a compact recurrent subset, and all Borel subsets* B *for which* $P(1, x_0, B) > 0$ *are recurrent in the sense that* $\mathbb{P}_x\left[\tau_B^1 < \infty\right] = 1$ *for all* $x \in B$. *If, moreover, the Markov chain in (10.340) is aperiodic, then there exists an integer* $m \in \mathbb{N}$, $m \geqslant 1$, *and a compact* m*-small set* A *such that* $\nu_m(A) > 0$ *which is compact. Here the measure* ν_m *satisfies* $P(m, x, B) \geqslant \nu_m(B)$ *for all* $B \in \mathcal{B}$ *and all* $x \in A$.

Proof. The first two assertions are consequences of respectively Proposition 9.5 and Lemma 9.4. The final assertion is a consequence of Theorem 10.15, and the fact that Borel measures on a Polish space are inner-regular. □

Among other things the following lemma reduces the proof of Orey's theorem for arbitrary irreducible aperiodic Markov chains to that for arbitrary irreducible strongly aperiodic Markov chains.

Lemma 10.15. *Let* μ_1 *and* μ_2 *be probability measures on* $\mathcal{E}$. *Then the sequence* $n \mapsto \iint \operatorname{Var}(P(n, x, \cdot) - P(n, y, \cdot))\, d\mu_1(x) d\mu_2(y)$ *is monotone decreasing.*

Proof. Fix $(x, y) \in E \times E$. The expression

$$\operatorname{Var}(P(n+1, x, \cdot) - P(n+1, y, \cdot))$$

can be rewritten as follows

$$\begin{aligned}
&\operatorname{Var}(P(n+1, x, \cdot) - P(n+1, y, \cdot)) \\
&= \sup\left\{\left|\int (P(n+1, x, dz) - P(n+1, y, dz))\, f(z)\, dz\right| : \|f\|_\infty \leqslant 1\right\} \\
&= \sup\left\{\left|\int (P(n, x, dw) - P(n, y, dw)) \int P(1, w, dz) f(z)\, dz\right| : \|f\|_\infty \leqslant 1\right\}
\end{aligned}$$

(notice that $\left|\int P(1, w, dz) f(z)\right| \leqslant \|f\|_\infty$, $w \in E$)

$$\leqslant \operatorname{Var}(P(n, x, \cdot) - P(n, y, \cdot)). \tag{10.341}$$

The inequality in (10.341) yields Lemma 10.15. □

We will also use the Nummelin splitting of general (Harris) recurrent chains. This splitting technique is taken from [Meyn and Tweedie (1993b)], §5.1 and §17.3.1. With a strongly aperiodic irreducible chain it associates a split chain with an accessible atom.

Let the Markov chain (10.340) have the properties described in Theorem 10.16. Then the Markov $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N}, x\in E}$ is aperiodic: see Proposition 10.1. From Corollary 10.6 it follows that there exists $m \in \mathbb{N}$, $m \geqslant 1$, such that the skeleton Markov chain $\{X(mn), \mathbb{P}_x\}_{n\in\mathbb{N}, x\in E}$ is strongly aperiodic. Definition 10.7 yields the existence of a compact recurrent subset C such that $P(1, x_0, C) > 0$ together with a probability measure ν on $\mathcal{E}$ such that $\nu(C) = 1$, and such that the following minorization condition is satisfied:

$$P(m, x, B) \geqslant \delta \mathbf{1}_C(x)\nu(B), \quad \text{for all } x \in X, \text{ and all } B \in \mathcal{E}. \tag{10.342}$$

In the presence of a subset C and a constant $m \in \mathbb{N}$ such that (10.342) holds for some probability measure ν with $\nu(C) = 1$ we will construct a split chain $\left\{\check{X}(n) = (X(mn), Y(n)), \check{\mathbb{P}}_{x,\varepsilon}\right\}_{n\in\mathbb{N}, x\in E, \varepsilon=0 \text{ or } 1}$. The m-step Markov chain $\{X(mn), \mathbb{P}_x\}_{n\in\mathbb{N}, x\in E}$ is strongly aperiodic, and it may be split to form a new chain with an accessible atom $C \times \{1\}$. Momentarily we will explain how the construction of this splitting can be performed.

In order to distinguish the new split Markov chain and the old skeleton chain we will introduce some new notation. We let the sequences of random variables $(Y(n), n \in \mathbb{N})$ attain the values zero and one. The value of $Y(n)$ indicates the level of the split m-skeleton at time mn. The split chain $\left\{\check{X}(n) = (X(mn), Y(n)), \check{\mathbb{P}}_{x,\varepsilon}\right\}_{n\in\mathbb{N}, x\in E, \varepsilon=0 \text{ or } 1}$ can be described in the following manner. Following Meyn and Tweedie [Meyn and Tweedie (1993b)] we write $\left\{\check{X}(n) = x_i\right\} = \{X(n) = x, Y(n) = i\}$, $x \in E$, $i = 0$ or $i = 1$. The new state space $\check{E}$ is given by $\check{E} = E \times \{0, 1\}$; $\check{\mathcal{E}}$ is the Borel field of $\check{E}$. The σ-field $\check{F}_k$ stands for

$$\check{F}_{k,\ell} = \sigma\left(X(j_1), Y(j_2) : 0 \leqslant j_1 \leqslant k, 0 \leqslant j_2 \leqslant \ell\right).$$

Let λ be any Borel measure on $\mathcal{E}$, then λ is split as a measure λ^* on $\check{E}$ in the following fashion. Let $A \in \mathcal{E}$ and put $A_0 = A \times \{0\}$, and $A_1 = A \times \{1\}$. Then the marginal measures of λ^* are given by

$$\left.\begin{aligned} \lambda^*(A_0) &= (1-\delta)\lambda\left(A \bigcap C\right) + \lambda\left(A \bigcap (E \backslash C)\right), \\ \lambda^*(A_1) &= \delta\lambda\left(A \bigcap C\right). \end{aligned}\right\} \tag{10.343}$$

Notice the equality $\lambda^*\left(A_0 \bigcup A_1\right) = \lambda(A)$, and $\lambda^*\left(A_0\right) = \lambda(A)$ when A is a subset of $E \backslash C$. In other words only subsets of C are split by this construction.

The splitting of the skeleton $\{X(nm), \mathbb{P}_x\}_{n\in\mathbb{N}, x\in E}$ is carried out as follows. Define the split kernel $\check{P}(m, x_i, A)$, $x_i \in \check{E}$, $A \in \check{\mathcal{E}}$ by

$$\left.\begin{aligned} \check{P}(m, x_0, \cdot) &= P(m, x, \cdot)^*, && x_0 \in E_0 \backslash C_0; \\ \check{P}(m, x_0, \cdot) &= \frac{P(m, x, \cdot)^* - \delta\nu^*(\cdot)}{1-\delta}, && x_0 \in C_0; \\ \check{P}(m, x_1, \cdot) &= \delta\nu^*(\cdot), && x_1 \in E_1. \end{aligned}\right\} \tag{10.344}$$

On $E_1 = E \times \{1\}$ the distribution of the split chain is also determined by prescribing the following conditional expectations:

$$\begin{aligned} &\check{\mathbb{E}}\left[\prod_{j=1}^m f_j\left(X(nm+j)\right), Y(n) = 1 \,\middle|\, \check{\mathcal{F}}_{nm,n-1}; X(nm) = x\right] \\ &= \check{\mathbb{E}}\left[\prod_{j=1}^m f_j\left(X(j)\right), Y(0) = 1 \,\middle|\, X(0) = x\right] \\ &= \delta\mathbb{E}_x\left[\prod_{j=1}^m f_j\left(X(j)\right) r\left(x, X(m)\right)\right], \end{aligned} \tag{10.345}$$

where the Borel measurable function $(x, y) \mapsto r(x, y)$ is the Radon-Nikodym derivative:

$$r(x, y) = \mathbf{1}_C(x)\frac{\nu(dy)}{P(m, x, dy)}. \tag{10.346}$$

By putting $f_j = \mathbf{1}$, $1 \leqslant j \leqslant m-1$, in (10.346) we see that

$$\begin{aligned} &\check{\mathbb{E}}\left[f_m\left(X\left((n+1)m\right)\right), Y(n) = 1 \,\middle|\, \check{\mathcal{F}}_{nm,n-1}; X(nm) = x\right] \\ &= \delta\mathbb{E}_x\left[f_m\left(X(m)\right) r\left(x, X(m)\right)\right] = \delta\mathbf{1}_C(x)\int f_m(y)d\nu(y). \end{aligned} \tag{10.347}$$

By taking $f_m = \mathbf{1}$ in (10.347) we get

$$\check{\mathbb{P}}\left[Y(n) = 1 \,\middle|\, \check{\mathcal{F}}_{nm,n-1}; X(nm) = x\right] = \delta\mathbf{1}_C(x). \tag{10.348}$$

By Bayes rule applied to (10.347) and (10.348) we obtain

$$\check{\mathbb{E}}\left[f\left(X\left((n+1)m\right)\right) \,\middle|\, \check{\mathcal{F}}_{nm,n}; X(nm) = x, Y(n) = 1\right] = \int f(y)d\nu(y). \tag{10.349}$$

Let f_j, $0 \leqslant j \leqslant N$, be bounded Borel functions on E, and let the numbers ε_j, $0 \leqslant j \leqslant N$, be equal to 0 or 1. From the tower property of conditional expectations, the Markov property of the process

$$\left\{\left(\breve{\Omega}, \breve{\mathcal{F}}, \breve{P}_{(x,i)}\right)_{(x,i)\in E\times\{0,1\}}, \left(\breve{X}(n), n \geqslant 0\right), \left(\breve{E}, \breve{\mathcal{E}}\right)\right\}, \tag{10.350}$$

and (10.349) we infer, with

$$F_{n+1} = \prod_{j=0}^{N} f_j\left(X((n+1)m+j)\right)\delta_{\varepsilon_j}\left(Y(j+n+1)\right),$$

that

$$\begin{aligned}
&\breve{\mathbb{E}}_{x,1}\left[F_{n+1} \bigm| \breve{\mathcal{F}}_{nm,n}\right] \\
&= \breve{\mathbb{E}}\left[F_{n+1} \bigm| \breve{\mathcal{F}}_{nm,n}; X(nm) = x,\, Y(n) = 1\right] \\
&= \breve{\mathbb{E}}\left[\breve{\mathbb{E}}\left[F_{n+1} \bigm| \breve{\mathcal{F}}_{(n+1)m,n+1}\right] \bigm| \breve{\mathcal{F}}_{nm,n}; X(nm) = x,\, Y(n) = 1\right] \\
&= \breve{\mathbb{E}}\Big[\breve{\mathbb{E}}\Big[F_{n+1} \bigm| \sigma(X((n+1)m), Y(n+1))\Big] \bigm| \breve{\mathcal{F}}_{nm,n}; X(nm) = x,\, Y(n) = 1\Big] \\
&= \int \breve{\mathbb{E}}\left[F_{n+1} \bigm| \sigma\left(Y(n+1)\right); X((n+1)m) = y\right] d\nu(y) \qquad (10.351) \\
&= \int \breve{\mathbb{E}}_{y,\varepsilon_0}\left[\prod_{j=0}^{N} f_j\left(X(j)\right)\delta_{\varepsilon_j}\left(Y(j)\right)\right] d\nu(y). \qquad (10.352)
\end{aligned}$$

The equality in (10.351) yields the $\breve{\mathbb{P}}_{x,1}$-independence of the following two σ-fields,
given that $Y(n) = 1$: $\breve{\mathcal{F}}_{nm,n} = \sigma\left(X(i), Y(j) : 0 \leqslant i \leqslant nm,\, 0 \leqslant j \leqslant n\right)$ and $\breve{\mathcal{F}}^{(n+1)m,n+1} = \sigma\left(X(i), Y(j) : i \geqslant (n+1)m,\, j \geqslant n+1\right)$.

From (10.351) it also follows that for $f \geqslant 0$ and Borel measurable, $k \in \mathbb{N}$, $k \geqslant 1$, and $\varepsilon = 0$ or 1,

$$\begin{aligned}
&\breve{\mathbb{E}}_{x,1}\left[f\left(X\left((n+1)m+k\right)\right)\delta_{\varepsilon}\left(Y\left((n+1)m+k\right)\right) \bigm| \breve{\mathcal{F}}_{nm,n}\right] \\
&\quad = \breve{\mathbb{E}}\Big[f\left(X\left((n+1)m+k\right)\right)\delta_{\varepsilon}\left(Y\left((n+1)m+k\right)\right) \bigm| \\
&\qquad\qquad \breve{\mathcal{F}}_{nm,n}; X(nm) = x,\, Y(n) = 1\Big] = \int \mathbb{E}_y\left[f\left(X(k)\right)\right] d\nu(y). \qquad (10.353)
\end{aligned}$$

From (10.353) we infer by taking expectation with respect to $\breve{\mathbb{E}}_{x,1}$ that

$$\breve{\mathbb{E}}_{x,1}\left[f\left(X\left((n+1)m+k\right)\right)\delta_{\varepsilon}\left(Y\left((n+1)m+k\right)\right)\right] = \int \mathbb{E}_y\left[f\left(X(k)\right)\right] d\nu(y),$$

and consequently, the subset $C \times \{1\}$ serves as an atom for the split Markov chain

$$\left\{ (X(mn), Y(n)), \check{\mathbb{P}}_{x,\varepsilon} \right\}_{(x,\varepsilon)\in E\times\{0,1\}}. \tag{10.354}$$

It is assumed that the process in (10.354) is a time-homogeneous Markov chain with transition function $\check{P}(nm, x_i, A)$, $n \in \mathbb{N}$, $x_i = (x,i) \in E \times \{0,1\}$. Here the "first-step" transition function $P(m, x_i, A)$ is given by (10.344). By the Markov property it then follows that the transition function $P(nm, x_i, A)$ satisfies the Chapman-Kolmogorov equation, i.e. the equality

$$\int_{\check{E}} \check{P}(jm, x_i, dy_j)\, \check{P}(km, y_j, A) = \check{P}((j+k)\,m, x_i, A) \tag{10.355}$$

holds for all $x_i \in \check{E}$ and $A \in \check{\mathcal{E}}$ and j, $k \in \mathbb{N}$. Compare all this with the Markov chain in (10.350).

The following theorem appears as Theorem 5.1.3 in Meyn and Tweedie.

Theorem 10.17. *Let $\delta > 0$, the probability measure ν, and $m \in \mathbb{N}$. $m \geqslant 1$ be as in (10.342). Let φ be a σ-finite measure on $\mathcal{E}$. Suppose that the function $P(nm, x_i, A)$ serves as a transition function for the Markov process in (10.354). In particular the Chapman-Kolmogorov identity (10.355) is satisfied. Then the following assertions hold:*

(a) The chain $\{X(nm), \mathcal{P}_x\}_{n\in\mathbb{N}, x\in E}$ is the marginal chain of

$$\left\{ \check{X}(nm) = (X(nm), Y(n)), \check{\mathcal{P}}_{x,i} \right\}_{n\in\mathbb{N},(x,i)\in E\times\{0,1\}}, \tag{10.356}$$

in the sense that the equality

$$\int_E P(km, x, A)\; d\lambda(x) = \int_{\check{E}} \check{P}\left(km, y_i, A_0 \bigcup A_1\right)\; d\lambda^*(y_i) \tag{10.357}$$

holds for all Borel measures λ, all $A \in \mathcal{E}$ and all $k \in \mathbb{N}$.

(b) If the Markov chain in (10.356) is φ^*-irreducible, then the Markov chain $\{X(nm), \mathcal{P}_x\}$ is φ-irreducible.

(c) If the chain $\{X(nm), \mathcal{P}_x\}_{n\in\mathbb{N}, x\in E}$ is φ-irreducible with $\varphi(C) > 0$, then the split chain in (10.356) is ν^*-irreducible, and $C \times \{1\}$ is an accessible atom for the split chain (10.356).

For the definition of accessible atom the reader is referred to Definition 10.4.

Proof. (a). It suffices to prove (10.357) with $\lambda = \delta_x$, the Dirac measure at $x \in E$. We will employ induction with respect to k. First assume that

$k = 1$. By (10.343), (10.344) and the equality $\nu(E \backslash C) = 0$ for $x \in E \backslash C$ we have

$$\begin{aligned}
&\int_{\check{E}} d\delta_x^*(y_i)\check{P}\Big(m, y_i, A_0 \bigcup A_1\Big) = \check{P}\Big(m, x_0, A_0 \bigcup A_1\Big) + \check{P}\Big(m, x_1, A_0 \bigcup A_1\Big) \\
&= P(m, x, \cdot)^*\Big(A_0 \bigcup A_1\Big) + \nu^*\Big(A_0 \bigcup A_1\Big) = P(m, x, A) + \nu\Big(A \bigcap (E \backslash C)\Big) \\
&= P(m, x, A). \qquad (10.358)
\end{aligned}$$

Next let $x \in C$. Again by employing (10.343) and (10.344) we infer

$$\begin{aligned}
&\int_{\check{E}} d\delta_x^*(y_i)\,\check{P}\left(m, y_i, A_0 \bigcup A_1\right) \\
&= \int_{E \times \{0\}} d\delta_x^*(y_i)\,\check{P}\left(m, y_i, A_0 \bigcup A_1\right) + \int_{E \times \{1\}} d\delta_x^*(y_i)\,\check{P}\left(m, y_i, A_0 \bigcup A_1\right) \\
&= (1-\delta)\check{P}\left(m, x_0, A_0 \bigcup A_1\right) + \delta\check{P}\left(m, x_1, A_0 \bigcup A_1\right) \\
&= (1-\delta)\frac{P(m, x, \cdot)^*(A_0 \bigcup A_1) - \delta\nu^*(A_0 \bigcup A_1)}{1-\delta} + \delta\nu^*\left(A_0 \bigcup A_1\right) \\
&= P(m, x, \cdot)^*\left(A_0 \bigcup A_1\right) = P(m, x, A). \qquad (10.359)
\end{aligned}$$

The equalities (10.358) (for $x \in E \backslash C$) and (10.359) (for $x \in C$) yield assertion (a) for $n = 1$ and $\lambda = \delta_x$. From Fubini's theorem assertions (a) is then also true for any bounded measure λ.

Next we assume that the equality in (10.357) holds for $1 \leqslant k \leqslant n$. First we notice that

$$\int_{\check{E}} \lambda^*(dx_i)\,\check{P}(m, x_i, \cdot) = \left(\int_E \lambda(dx) P(m, x, \cdot)\right)^*. \qquad (10.360)$$

Using the Chapman-Kolmogorov equation for the probability transition function $\check{P}(km, x_i, A)$ in combination with (10.360) and induction then shows

$$\int_{\check{E}} \lambda^*(dx_i)\,\check{P}(nm, x_i, \cdot) = \left(\int_E \lambda(dx) P(nm, x, \cdot)\right)^*. \qquad (10.361)$$

Here we need the Chapman-Kolmogorov identity (10.355) for A of the form $B_0 \bigcup B_1$ with $B \in \mathcal{E}$. For $k = n+1$ we then have

$$\begin{aligned}
&\int_E \lambda(dx)\,P((n+1)m, x, A) = \int_E \lambda(dx)\,P(nm, x, dy)\;P(m, y, A) \\
&= \int_{\check{E}} \left(\int_E \lambda(dx)\,P(nm, x, \cdot)\right)^*(dy_j)\,\check{P}\left(m, y_j, A_0 \bigcup A_1\right)
\end{aligned}$$

$$= \int_{\breve{E}} \left(\int_E \lambda(dx)\, P\left(nm, x, \cdot\right) \right)^* (dy_j)\, \breve{P}\left(m, y_j, A_0 \bigcup A_1\right)$$

(apply equality (10.361))

$$= \int_{\breve{E}} \lambda^*\left(dx_i\right) \int_{\breve{E}} \breve{P}\left(nm, x_i, dy_j\right) \breve{P}\left(m, y_j, A_0 \bigcup A_1\right)$$

(Chapman-Kolomogorov (10.355))

$$= \int_{\breve{E}} \lambda^*\left(dx_i\right)\, \breve{P}\left((n+1)m, x_i, A_0 \bigcup A_1\right). \tag{10.362}$$

The assertion in (a) follows from (10.362). Assertion (b) follows from (a) with φ instead of λ. In order to prove (c) we observe that $C \times \{1\}$ is an atom for the Markov chain in (10.356), which is a consequence of the ultimate equality in (10.344). If $\varphi(C) > 0$, then from the minorization property in (10.342) it follows that the split chain (10.356) is ν^*-irreducible, and that $C \times \{1\}$ is an accessible atom.

Altogether this completes the proof of Theorem 10.17. □

Next we prove Orey's theorem, i.e. we prove Theorem 10.2.

Proof. [Proof of Theorem 10.2.] We distinguish three cases:

(i) The irreducible recurrent chain $\{X(n), \mathbb{P}_x\}_{x \in E}$ contains an accessible atom.
(ii) The irreducible recurrent chain is strongly aperiodic.
(iii) The irreducible recurrent chain is aperiodic.

In case the irreducible recurrent chain contains an accessible atom A we use formula (10.310) to obtain:

$$\begin{aligned} &\left|\mathbb{E}_x\left[f\left(X(n)\right)\right] - \mathbb{E}_A\left[f\left(X(n)\right)\right]\right| \\ &\leqslant \|f\|_\infty \left(\mathbb{P}_x\left[\tau_A^1 \geqslant n\right] + \left(a_x * u_A - u_A\right) * \overline{p}_{A,\mathbf{1}}(n-1)\right). \end{aligned} \tag{10.363}$$

Here $f \in C_b(E)$ is arbitrary, and the sequences are $a_x(n) = \mathbb{P}_x\left[\tau_A^1 = n\right]$, $u_A(n) = \mathbb{P}_A\left[X(n) \in A\right]$, and $p_{A,f}(n)$ are chosen as in (10.309). In fact

$$\begin{aligned} p_{A,f}(k) &= \mathbb{E}_A\left[f\left(X(k)\right), \tau_A^1 = k\right], \quad \text{and} \\ \overline{p}_{A,f}(k) &= \mathbb{E}_A\left[f\left(X(k)\right), \tau_A^1 \geqslant k\right]. \end{aligned} \tag{10.364}$$

Let n tend to ∞ in (10.363). Since $\mathbb{P}_x\left[\tau_A < \infty\right] = 1$ the first term in the right-hand side of (10.363) tends to zero uniformly in f provided that $\|f\|_\infty \leqslant 1$. The equality (10.325) in Theorem 10.14 yield that the second term in the right-hand side of (10.363) tends to zero, again uniformly in f provided $\|f\|_\infty \leqslant 1$. As a consequence we see that

$$\lim_{n\to\infty} \operatorname{Var}\left(P(n,x,\cdot) - P(n,A,\cdot)\right)$$
$$= \lim_{n\to\infty} \sup\left\{\left|\mathbb{E}_x\left[f\left(X(n)\right)\right] - \mathbb{E}_A\left[f\left(X(n)\right)\right]\right| : \|f\|_\infty \leqslant 1\right\} = 0. \quad (10.365)$$

By the triangle inequality and the dominated convergence theorem the equality in (10.15) in Theorem 10.2 is a consequence of (10.365). This proves assertion (i) in the beginning of this proof.

Next we will prove (10.15) in Theorem 10.2 in case the recurrent Markov chain $\{X(n), \mathbb{P}_x\}_{x\in E}$ is strongly aperiodic. This will be a consequence of Nummelin's splitting technique, and the fact that for Markov chains with an accessible atom Orey's theorem holds: see the arguments following equality (10.365). If the chain $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ is strongly aperiodic, then we know that inequality (10.342) holds with $m = 1$ for some recurrent subset C, and a probability measure ν on $\mathcal{E}$ with $\nu(C) = 1$ (and $P(1,x_0,C) > 0$). Using this subset C and this measure ν we may construct the split chain in (10.356) with marginal chain $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ (i.e. (10.357) is satisfied), and for which $C \times \{1\}$ is an accessible atom. These claims follow from assertion (b) and (c) in Theorem 10.17. Since the subset $C \times \{1\}$ is an accessible atom for the split chain in (10.356), we know that Orey's theorem holds for the split chain. The latter is a consequence of assertion (i), which in turn is a consequence of (10.365). Let x and $y \in E$. Then we infer

$$\operatorname{Var}\left(P(n,x,\cdot) - P(n,y,\cdot)\right)$$
$$\leqslant 2\sup_{A\in\mathcal{E}}\left|P(n,x,A) - P(n,y,A)\right|$$
$$= 2\sup_{A\in\mathcal{E}}\left|\iint_{\check{E}\times\check{E}}\left(\check{P}\left(n,x_i,A_0\bigcup A_1\right) - P\left(n,y_j,A_0\bigcup A_1\right)\right) d\delta_x^*(x_i)\, d\delta_y^*(y_j)\right|$$
$$\leqslant 2\iint_{\check{E}\times\check{E}}\sup_{A\in\mathcal{E}}\left|\check{P}\left(n,x_i,A_0\bigcup A_1\right) - P\left(n,y_j,A_0\bigcup A_1\right)\right| d\delta_x^*(x_i)\, d\delta_y^*(y_j)$$
$$\leqslant 2\iint_{\check{E}\times\check{E}}\operatorname{Var}\left(\check{P}(n,x_i,\cdot) - P(n,y_j,\cdot)\right) d\delta_x^*(x_i)\, d\delta_y^*(y_j). \quad (10.366)$$

By assertion (i), applied to the split chain in (10.356) (with $m = 1$) the final term in (10.366) converges to zero. By dominated convergence and (10.366) we see that

$$\lim_{n\to\infty}\iint_{E\times E}\operatorname{Var}\left(P(n,x,\cdot) - P(n,y,\cdot)\right) d\lambda_1(x) d\lambda_2(y) = 0. \quad (10.367)$$

The equality in (10.367) shows that Orey's theorem holds for strongly aperiodic recurrent Markov chains.

To finish the proof of Theorem 10.2 we suppose that $\{X(n), \mathbb{P}_x\}_{n\in\mathbb{N},x\in E}$ is an aperiodic recurrent chain. By assertion (ii), which has been proved now, for irreducible recurrent strongly aperiodic chains Orey's theorem holds. By Corollary 10.6 there exists $m \in \mathbb{N}$ such that the skeleton

$$\{(X(mn), \mathbb{P}_x) : n \in \mathbb{N}, x \in E\}$$

is strongly aperiodic. Since Orey's theorem holds for such chains, an application of Lemma 10.15 yields the result that Orey's theorem holds for all irreducible, recurrent aperiodic Markov chains.

This completes the proof of Theorem 10.2. □

10.4 About invariant (or stationary) measures

In this section we collect some references to work related to the existence of invariant or stationary measures for Markov processes. In this context we have to mention Harris [Harris (1956)] who proved the existence of a σ-finite invariant measure for recurrent irreducible Markov chains. Let $P(x, B)$ be a probability transition function which preserves the bounded continuous functions on a Polish space E. Suppose that P is irreducible (i.e. for every $x \in E$, and for every non-void open subset O, $P^n(x, O) > 0$ for some $n \in \mathbb{N}$, $n \geqslant 1$), and topologically recurrent (i.e. for every $x \in E$ and every open neighborhood O of x the equality $\mathbb{P}_x\left[\bigcup_{n=1}\{X(n) \in O\}\right] = 1$ holds). Here

$$\{(\Omega, \mathcal{F}, \mathbb{P}_x)_{x\in E}, (X(n), n \in \mathbb{N}), (\vartheta_k, k \in \mathbb{N}), (E, \mathcal{E})\}$$

is the Markov chain with transition function $(x, B) \mapsto P(x, B)$ $(x, B) \in E \times \mathcal{E}$. Harris proved that for a discrete state space E there exists a σ-finite invariant measure, and Orey [Orey (1959, 1962, 1964)] was the first to prove that in the presence of a finite invariant $\lim_{n\to\infty} \int_E f(y) P^n(x, dy)\, d\mu(x) = 0$ for all finite real Borel measures μ on E such that $\mu(E) = 0$. The original result by Harris and Orey for discrete positive recurrent chains were improved and generalized by Jamison and Orey [Jamison and Orey (1967)], and Kingman and Orey [Kingman and Orey (1964)] to Markov chains with a more general state space, and for null-recurrent chains. In [Nummelin and Tuominen (1982, 1983)] Nummelin and Tuominen discuss geometric ergodicity properties, and so do Tuominen and Tweedie in [Tuominen and Tweedie (1994)]. This is also the case in Baxendale [Baxendale (2005)]. For a general discussion on Markov chains and their limit theorems see e.g. the books by Nummelin [Nummelin (1984)], Revuz [Revuz (1975)], and Orey [Orey (1971)]. The new version of Meyn and Tweedie [Meyn and Tweedie (1993b)] also contains a wealth of information. It explains splitting (due to Nummelin [Nummelin (1978)]) and (dependent) coupling techniques (due to Ornstein [Ornstein (1969)]), and several limit properties as well as asymptotic behavior of Markov chains. In addition, it discusses geometric ergodic chains, certain functional central limit theorems, and laws of large numbers. All these topics are explained for discrete time Markov processes with an arbitrary state space. Moreover, each of the 19 chapters of [Meyn and Tweedie (1993b)] is concluded with a section, entitled Commentary, which contains bibliographic notes and relevant observations. Azema, Duflo and Revuz apply skeleton techniques to pass from discrete time limit theorems to continuous time limits: see e.g. [Azéma *et al.* (1965/1966, 1966, 1967)]. In the proof of Proposition 9.7 we applied the same methods. Our approach uses the techniques of Seidler [Seidler (1997)] (propositions 5.7 and 5.9) in combination with Orey's theorem for Markov chains on a compact space. For more details the reader is referred to the comments following Theorem 10.1, and to the Notes, pp. 319–320, in Supplement, Harris processes, Special functions, Zero-two law, written by Antoine Brunel in [Krengel (1985)]. Orey's convergence theorem is based on renewal theory which uses a linked forward recurrence time chain, which also plays a central role in the book by Meyn and Tweedie [Meyn and Tweedie (1993b)]. For more historical and bibliographical notes the reader is also referred to Kallenberg [Kallenberg (2002)], pp 569–593. On the other hand the author likes to mention the following papers and books explicitly: Doeblin [Doeblin (1937, 1940)], Kolmogorov [Kolmogorov (1956, 1991, 1993)], Doob [Doob (1953)] Chapter V, §5, and Dobrushin [Dobrushin (1956a,b)]. For a historical survey of the life and the mathematical work by Doeblin see e.g. Lindvall [Lindvall (1991)], Bru and Yor [Bru and Yor (2002)], and Mazliak [Mazliak (2007)]. For a martingale approach of Dobrushin's theorem on Markov chains see [Sethuraman and Varadhan (2005)]. The history and uses of the Markov-Dobrushin coefficient of ergodicity are explained by Seneta in [Seneta (1993)], and also in [Seneta (1981)]. They are used to give the speed of convergence, which for application is quite important. For a result on mixing properties and the central limit theorem see e.g. Bolthausen [Bolthausen (1982)]. For a general account of ergodic theory we also refer to Chen [Chen (1999)]. For more details on Markov chains the reader should also consult Nummelin [Nummelin (1984)] and Meyn and Tweedie [Meyn and Tweedie (1993b)]. For a more operator theoretic approach to ergodic theory see e.g. Foguel

[Foguel (1969)] and Meyn [Meyn (2008)]. Remarks about Kolmogorov's example, and extensions of Kolmogorov's work on Markov chains can be found in Reuter [Reuter (1969)], and in earlier work by Kendall and Reuter [Kendall and Reuter (1956)], and Doob [Doob (1945)]. It is noticed that Kendall and Reuter apply semigroup methods to treat path regularity properties of the underlying Markov process. In [Stroock and Zegarliński (1992)] Stroock and Zegarlinski explain the relationship between the logarithmic Sobolev inequality and Dobrushin's mixing condition for ergodicity.

10.4.1 *Possible applications*

The material presented in this book finds its applications in several branches of the scientific world. Markov theory is relevant in mathematical models from economics (equilibrium in markets), finance (backward equations in hedging strategies), equilibrium states in statistical mechanics, mathematical physics (Feynman-Kac type formulas), biology (equilibrium states). In the context of population dynamics we mention two standard textbooks [Allen (2003)], [Allen (2007)]. The book [Bharucha-Reid (1997)] contains several interesting models and applications. The textbook [Mikosch (1998)] contains a rather elementary introduction to stochastic (i.e. Itô) calculus with applications in relatively simple models for trading strategies.

10.4.2 *Conclusion*

A great part of this chapter was devoted to the proof of the existence and uniqueness of σ-finite invariant Borel measures: see Theorem 10.12. The relevant conditions are presented (like irreducibility, and existence of recurrent compact subset, which is a consequence of the almost separability property of the generator L of the Markov process (9.14)). For these results the reader is referred to Definition 9.2, and Propositions 9.1, 9.2, and 9.4 in Chapter 9. Another feature of the present chapter is a discussion and proof of Orey's convergence theorem: see §10.3.

To conclude this section we insert some well-known results related to ergodicity properties of Markov chains. Let $P = (p_{i,j})_{I,j\in S}$ be a row-stochastic matrix with real entries which serves as a transition matrix for a Markov chain $(X(n), \mathbb{P}_j)$. We say that P is row-stochastic if $p_{i,j} \geqslant 0$ for every $i, \in S$ and $\sum_{j\in S} p_{i,j} = 1$ for every $i \in S$. Set

$$\alpha(P) = \min_{i\neq j} \sum_{k\in S} \min\left(p_{i,k}, p_{j,k}\right), \quad \widetilde{\alpha}(P) = 1 - \alpha(P). \tag{10.368}$$

The number $\alpha(P)$ is known today as the Dobrushin coefficient of ergodicity: see Cohen [Cohen *et al.* (1993)], Dobrushin [Dobrushin (1956a,b)]. The following result can be found in Zaharopol and Zbaganu [Zaharopol and Zbaganu (1999)]: see Zaharopol [Zaharopol (2005)] and [Del Moral *et al.* (2003)] as well. Another source of information is Stachurski [Stachurski (2009)], in which Doobrushin's coefficients play a dominant role. One of the standard results reads as follows.

Theorem 10.18. *Let the Markov transition function P have Dobrushin's coefficient $\alpha(P)$. Then the following inequality holds for all probability distributions φ and ψ on S:*

$$\|\varphi P - \psi P\|_1 \leqslant (1 - \alpha(P)) \|\varphi - \psi\|_1 .$$

A similar result is also true for transition densities and integrals instead of sums: see e.g. Chapter 8 in [Stachurski (2009)].

We begin with a classical theorem in which Doeblin's condition plays a central role.

Theorem 10.19. *Let $(X(n), \mathbb{P}_j)_{n\in\mathbb{N}, j\in S}$ be a Markov chain in a countable state space S with transition probabilities $p_{i,j}$ such that: There exists a state $a \in S$ and $\varepsilon > 0$, with the property: $p_{i,a} \geqslant \varepsilon > 0$, for all $i \in S$. Then there is a unique stationary (or invariant) distribution π such that*

$$\sum_{j\in S} \left|\mathbb{P}_{X(0)}\left[X(n) = j\right] - \pi(j)\right| \leqslant 2\left(1 - \varepsilon\right)^n , \tag{10.369}$$

regardless of the initial state $X(0)$.

An "analytic" proof runs as follows. Think in terms of of the one-step transition matrix $P = (p_{i,j})$ as a linear operator acting on $\mathbb{R}^S$. Equip $\mathbb{R}^S$ with the norm $\|x\| := \sum_{j\in S} |x_j|$. Stroock [Stroock (2000)], pg. 28–29, proves that, for any $\rho \in \mathbb{R}^S$, such that $\sum_{j\in S} \rho_j = 0$, we have $\|\rho P\| \leqslant (1-\varepsilon)\|\rho\|$. He then claims that this implies that $\|\rho P^n\| \leqslant (1-\varepsilon)^n |\rho\|$, $n \in \mathbb{N}$, and uses this to show that, for any $\mu \in \mathbb{R}^S$ with, $\mu_i \geqslant 0$ for all $i \in S$, and $\sum_{i\in S} \mu_i = 1$, it holds that $\|\mu P^n - \mu P^m\| \leqslant 2(1-\varepsilon)^m$, for $m \geqslant n$.

A "probabilistic" proof runs as follows. Consider the following experiment. Suppose the current state is i. Toss a coin with $\mathbb{P}(heads) = \varepsilon$. If heads show up then move to state a. If tails show up, then move to state j with probability $\widetilde{p}_{i,j} = \dfrac{p_{i,j} - \varepsilon\delta_{a,j}}{1-\varepsilon}$. (That this is a valid probability indeed is a consequence of the assumption!) In this manner, we obtain a process that has precisely transition probabilities $p_{i,j}$. Note that state a will be visited either because of heads in a coin toss or because it was chosen so by the alternative transition probability. So state a will be visited at least as many times as the number of heads in a coin toss. This means that state a is positive recurrent. And so a stationary probability π exists. We will show that this π is unique and that the distribution of the chain converges to it. To do this, consider two chains X, $\overline{X}$, both with transition probabilities $p_{i,j}$, and realize them as follows. The first one starts with $X(0)$ distributed according to an arbitrary law μ. The second one starts with $\overline{X}(0)$ distributed according to π. Now do this: Use the same coin for both. So, if heads show up then move both chains to a. If tails show up then realize each one according to $\widetilde{p}$, independently. Repeat this at the next step, by tossing a new coin, independently of the past. Thus, as long as heads have not come up yet, the chains are moving independently. Of course, sooner or later, heads will show up and the chains will be the same thereafter. Let T be the first time at which heads show up. We have:

$$\begin{aligned}\mathbb{P}\left[X(n) \in B\right] &= \mathbb{P}\left[X(n) \in B,\, T > n\right] + \mathbb{P}\left[X(n) \in B,\, T \leqslant n\right] \\ &= \mathbb{P}\left[X(n) \in B,\, T > n\right] + \mathbb{P}\left[X(n) \in B,\, T \leqslant n\right] \\ &\leqslant \mathbb{P}\left[T > n\right] + \mathbb{P}\left[X(n) \in B\right] = \mathbb{P}\left[T > n\right] + \pi(B).\end{aligned}$$

Similarly,

$$\begin{aligned}\pi(B) = \mathbb{P}\left[\overline{X}(n) \in B\right] &= \mathbb{P}\left[\overline{X}(n) \in B,\, T > n\right] + \mathbb{P}\left[X(n) \in B,\, T \leqslant n\right] \\ &= \mathbb{P}\left[X(n) \in B,\, T > n\right] + \mathbb{P}\left[\overline{X}(n) \in B,\, T \leqslant n\right] \\ &\leqslant \mathbb{P}\left[T > n\right] + \mathbb{P}\left[X(n) \in B\right].\end{aligned}$$

Hence $\left|\mathbb{P}\left[X(n) \in B\right] - \pi(B)\right| \leqslant \mathbb{P}\left[T > n\right] = (1-\varepsilon)^n$. Finally, check that

$$\sup_{B \subset S} \left|\mathbb{P}\left[X(n) \in B\right] - \pi(B)\right| = \frac{1}{2} \sum_{i \in S} \left|\mathbb{P}\left[X(n) = i\right] - \pi(i)\right|.$$

The following theorem of Kolmogorov on mean recurrence times is taken from [Kallenberg (2002)] Theorem 7.22.

Theorem 10.20. *For a Markov chain with state space S and for states i, $j \in S$ with j aperiodic, the following equality holds:*

$$\lim_{n \to \infty} p_{ij}^n = \frac{\mathbb{P}_i\left[\tau_j < \infty\right]}{\mathbb{E}_j\left[\tau_j\right]},$$

where τ_j is the first time visiting j: $\tau_j = \inf\{m \geqslant 1 : X(m) = j\}$.

In [Foss and Konstantopoulos (2004)] the authors describe a generalization of this result by introducing what is called an inverse Palm construction and using Palm stationarity. For more information on Palm distributions see e.g. Chapter 8 and 9 in [Thorisson (2000)], [Etheridge (2000)], [Kallenberg (2008)], and [Dawson and Perkins (1999); Dawson (1993)]. For a concise formulation of a result concerning recurrent Markov chains, which in a discrete state space dates back to Doeblin we insert a definition.

Definition 10.8. Let $(X(n), \mathbb{P}_x)$ be a Markov chain with a Polish state space E, and transition function $B \mapsto P(x,B)$, $B \in \mathcal{E}$, the Borel field of E.

(i) The chain $(X(n), \mathbb{P}_x)$ is called uniformly ergodic provided there exists an invariant measure π on $\mathcal{E}$ such that $\lim_{n\to\infty} \sup_{x\in E} \|P^n(x,\cdot) - \pi(\cdot)\|_{\mathrm{Var}} = 0$.
(ii) The chain is said to satisfy Doeblin's condition if there exist a probability measure φ on $\mathcal{E}$ and strictly positive numbers δ and ε and a strictly positive integer m such that $\varphi(A) > \varepsilon$ implies $\inf_{x\in X} P^m(x,A) \geqslant \delta$.
(iii) The chain $(X(n), \mathbb{P}_x)$ has uniform geometric speed (or rate) of convergence if there exist an invariant probability π on $\mathcal{E}$ and constants R and $0 < r < 1$ such that $\|P^n(x,\cdot) - \pi(\cdot)\|_{\mathrm{Var}} \leqslant Rr^n$ for all $n \in \mathbb{N}$ and $x \in E$.
(iv) The chain $(X(n), \mathbb{P}_x)$ is uniformly positive recurrent if there exists a compact subset K such $\sup_{x\in K} \mathbb{E}_x[\tau_K] < \infty$.

In (iv) τ_K is the hitting time of K: $\tau_K = \inf\{m \geqslant 1 : X(m) \in K\}$. If such a compact subset K exists, then for all subsets $A \in \mathcal{E}$ for which $\varphi(A) > 0$ the inequality $\sup_{x\in A} \mathbb{E}_x[\tau_A] < \infty$ holds. Here φ is as in item (ii) of Definition 10.8: the existence of such a probability measure φ is guaranteed by item (iii) in Theorem 10.21 below. Theorem 16.0.2 in [Meyn and Tweedie (1993b)], which is more general then Theorem 10.21, says among other things that a Markov chain is uniformly ergodic if and only if it is aperiodic and satisfies Doeblin's condition. For the notion of aperiodicity and related topics see Definitions 10.1, 10.6 and 10.7.

Theorem 10.21. *Let $(X(n), \mathbb{P}_x)$ be a Markov chain with a Polish state space E, and transition function $B \mapsto P(x,B)$, $B \in \mathcal{E}$. Then the following assertions are equivalent:*

(i) The Markov chain $(X(n), \mathbb{P}_x)$ is uniformly ergodic;
(ii) The Markov chain $(X(n), \mathbb{P}_x)$ has uniform geometric speed of convergence.

(iii) The Markov chain $(X(n), \mathbb{P}_x)$ *is aperiodic and satisfies Doeblin's condition.*

(iv) The Markov chain $(X(n), \mathbb{P}_x)$ *is aperiodic and uniformly positive recurrent.*

With this nice theorem we conclude this chapter and this book. For more information and related results in the time discrete case the reader is referred to Meyn and Tweedie [Meyn and Tweedie (1993b)]. For recent work on ergodicity of Markov chains see e.g. Hairer and Mattingly [Hairer and Mattingly (2008b)]. Application of ergodicity properties of Markov processes can be found in the theory of stochastic partial differential equations by Hairer and co-authors, see e.g. [Hairer *et al.* (2004); Hairer and Mattingly (2008a,b); Hairer (2009)]. For the use of the Foster-Lyapunov criterion in the study of the stability of Markov chains the reader is referred to e.g. Meyn and Tweedie [Meyn and Tweedie (1993b)], and also to a recent paper by Connor and Fort [Connor and Fort (2009)]. It is possible that this Foster-Lyapunov criterion is linked to the separation property of the domain of the generator of the underlying Markov property as mentioned in Corollary 9.2 and Theorem 9.4. A consequence of this separation hypothesis is that the existence of a recurrent compact subset is guaranteed provided that all open subsets are recurrent, and that all measures $B \mapsto P(t, x, B)$, $B \in \mathcal{E}$, $(t, x) \in (0, \infty) \times E$, are equivalent.

Bibliography

Aida, S. (1998). Uniform positivity improving property, Sobolev inequalities, and spectral gaps, *J. Funct. Anal.* **158**, 1, pp. 152–185.

Albeverio, S., Rezende, J. and Zambrini, J.-C. (2006a). Probability and quantum symmetries. II. The theorem of Noether in quantum mechanics, *J. Math. Phys.* **47**, 6, pp. 062107, 61, doi:10.1063/1.2199087, URL `http://dx.doi.org/10.1063/1.2199087`.

Albeverio, S., Rezende, J. and Zambrini, J.-C. (2006b). Probability and quantum symmetries. II. The theorem of Noether in quantum mechanics, *J. Math. Phys.* **47**, 6, pp. 062107, 61 pages.

Aliprantis, C. D. and Border, K. C. (1994). *Infinite-dimensional analysis*, *Studies in Economic Theory*, Vol. 4 (Springer-Verlag, Berlin), ISBN 3-540-58308-4, a hitchhiker's guide.

Allen, L. J. S. (2003). *An Introduction to Stochastic Processes with Applications to Biology* (Pearson Education, Inc., Upper Saddle River, NJ 07458), ISBN 13: 9780130352187, textbook, Hardcover.

Allen, L. J. S. (2007). *Introduction to Mathematical Biology* (Pearson Education, Inc., Upper Saddle River, NJ 07458), ISBN 10: 0130352160, 13: 9780130352163.

Altomare, F. and Attalienti, A. (2002a). Degenerate evolution equations in weighted continuous function spaces, Markov processes and the Black-Scholes equation. I, *Results Math.* **42**, 3-4, pp. 193–211.

Altomare, F. and Attalienti, A. (2002b). Degenerate evolution equations in weighted continuous function spaces, Markov processes and the Black-Scholes equation. II, *Results Math.* **42**, 3-4, pp. 212–228.

Altomare, F. and Campiti, M. (1994). *Korovkin-type approximation theory and its applications*, *de Gruyter Studies in Mathematics*, Vol. 17 (Walter de Gruyter & Co., Berlin), ISBN 3-11-014178-7, appendix A by Michael Pannenberg and Appendix B by Ferdinand Beckhoff.

Andersen, E. S. and Jessen, B. (1948). On the introduction of measures in infinite product sets, *Danske Vid. Selsk. Mat.-Fys. Medd.* **25**, 4, p. 8.

Andrieu, C. and Fort, G. (2005). Explicit control of subgeometric ergodicity, Web page: http://www.tsi.enst.fr/ gfort/biblio.html.

Arendt, W. (1987). Vector-valued Laplace transforms and Cauchy problems, *Israel J. Math.* **59**, 3, pp. 327–352.

Arendt, W., Batty, C. J. K., Hieber, M. and Neubrander, F. (2001). *Vector-valued Laplace transforms and Cauchy problems*, *Monographs in Mathematics*, Vol. 96 (Birkhäuser Verlag, Basel), ISBN 3-7643-6549-8.

Azéma, J., Kaplan-Duflo, M. and Revuz, D. (1965/1966). Récurrence fine des processus de Markov, *Ann. Inst. H. Poincaré Sect. B (N.S.)* **2**, pp. 185–220.

Azéma, J., Kaplan-Duflo, M. and Revuz, D. (1966). Mesure invariante et théorèmes ergodiques sur les classes récurrentes des processus de Markov, *C. R. Acad. Sci. Paris Sér. A-B* **262**, pp. A1247–A1249.

Azéma, J., Kaplan-Duflo, M. and Revuz, D. (1967). Mesure invariante sur les classes récurrentes des processus de Markov, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8**, 3, pp. 157–181.

Bakry, D. (1985a). Transformations de Riesz pour les semi-groupes symétriques. I. Étude de la dimension 1, in *Séminaire de probabilités, XIX, 1983/84*, *Lecture Notes in Math.*, Vol. 1123 (Springer, Berlin), pp. 130–144.

Bakry, D. (1985b). Transformations de Riesz pour les semi-groupes symétriques. II. Étude sous la condition $\Gamma_2 \geqslant 0$, in *Séminaire de probabilités, XIX, 1983/84*, *Lecture Notes in Math.*, Vol. 1123 (Springer, Berlin), pp. 145–174.

Bakry, D. (1991). Inégalités de Sobolev faibles: un critère Γ_2, in *Séminaire de Probabilités, XXV*, *Lecture Notes in Math.*, Vol. 1485 (Springer, Berlin), pp. 234–261.

Bakry, D. (1994). L'hypercontractivité et son utilisation en théorie des semi-groupes, in *Lectures on probability theory (Saint-Flour, 1992)*, *Lecture Notes in Math.*, Vol. 1581 (Springer, Berlin), pp. 1–114, doi:10.1007/BFb0073872, URL http://dx.doi.org/10.1007/BFb0073872.

Bakry, D. (2006). Functional inequalities for Markov semigroups, in *Probability measures on groups: recent directions and trends* (Tata Inst. Fund. Res., Mumbai), pp. 91–147.

Bakry, D. and Émery, M. (1985). Diffusions hypercontractives, in *Séminaire de probabilités, XIX, 1983/84*, *Lecture Notes in Math.*, Vol. 1123 (Springer, Berlin), pp. 177–206.

Bakry, D. and Ledoux, M. (2006). A logarithmic Sobolev form of the Li-Yau parabolic inequality, *Rev. Mat. Iberoam.* **22**, 2, pp. 683–702, URL http://projecteuclid.org/getRecord?id=euclid.rmi/1161871352.

Balakrishnan, A. V. (ed.) (2000). *Semigroups of operators: theory and applications*, Progress in Nonlinear Differential Equations and their Applications, 42 (Birkhäuser Verlag, Basel), ISBN 3-7643-6310-X.

Bally, V., Pardoux, E. and Stoica, L. (2005). Backward stochastic differential equations associated to a symmetric Markov process, *Potential Anal.* **22**, 1, pp. 17–60.

Barles, G. (1993). Discontinuous viscosity solutions of first-order Hamilton-Jacobi equations: a guided visit, *Nonlinear Anal.* **20**, 9, pp. 1123–1134.

Barles, G., Cardaliaguet, P., Ley, O. and Monneau, R. (2008). Global existence results and uniqueness for dislocation equations, *SIAM J. Math. Anal.* **40**, 1, pp. 44–69.

Barlow, M. T., Bass, R. F. and Kumagai, T. (2005). Note on the equivalence of parabolic Harnack inequalities and heat kernel estimates, http://www.math.ubc.ca/~barlow/preprints/.

Bass, R. F. and Kumagai, T. (2008). Symmetric Markov chains on $\mathbb{Z}^d$ with unbounded range, *Trans. Amer. Math. Soc.* **360**, 4, pp. 2041–2075, doi:10.1090/S0002-9947-07-04281-X, URL `http://dx.doi.org/10.1090/S0002-9947-07-04281-X`.

Bauer, H. (1981). *Probability theory and elements of measure theory* (Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London), ISBN 0-12-082820-0, second edition of the translation by R. B. Burckel from the third German edition, Probability and Mathematical Statistics.

Bäumer, B. and Neubrander, F. (1994). Laplace transform methods for evolution equations, *Confer. Sem. Mat. Univ. Bari* , 258-260, pp. 27–60, swabian-Apulian Meeting on Operator Semigroups and Evolution Equations (Italian) (Ruvo di Puglia, 1994).

Baxendale, P. H. (2005). Renewal theory and computable convergence rates for geometrically ergodic Markov chains, *Ann. Appl. Probab.* **15**, 1B, pp. 700–738.

Bell, D. R. (2006). *The Malliavin calculus* (Dover Publications Inc., Mineola, NY), ISBN 0-486-44994-7, reprint of the 1987 edition.

Benzécri, J.-P. (1995). Théorie des capacités (d'après G. Choquet), in *Séminaire Bourbaki, Vol. 3* (Soc. Math. France, Paris), ISBN 2-85629-036-1, pp. Exp. No. 120, 217–227.

Berg, C. and Forst, G. (1975). *Potential theory on locally compact abelian groups* (Springer-Verlag, New York), ergebnisse der Mathematik und ihrer Grenzgebiete, Band 87.

Bergelson, V., March, P. and Rosenblatt, J. (eds.) (1996). *Convergence in ergodic theory and probability*, Ohio State University Mathematical Research Institute Publications, 5 (Walter de Gruyter & Co., Berlin), ISBN 3-11-014219-8, papers from the Third Conference held at The Ohio State University, Columbus, Ohio, June 1993, Dedicated to Louis Sucheston.

Berlinet, A. and Thomas-Agnan, C. (2004). *Reproducing kernel Hilbert spaces in probability and statistics* (Kluwer Academic Publishers, Boston, MA), ISBN 1-4020-7679-7, with a preface by Persi Diaconis.

Bharucha-Reid, A. T. (1997). *Elements of the theory of Markov processes and their applications* (Dover Publications Inc., Mineola, NY), ISBN 0-486-69539-5, reprint of the 1960 original.

Bhatt, A. G. and Karandikar, R. L. (1993). Invariant measures and evolution equations for Markov processes characterized via martingale problems, *Ann. Probab.* **21**, 4, pp. 2246–2268, URL `http://links.jstor.org/sici?sici=0091-1798(199310)21:4<2246:IMAEEF>2.0.CO;2-V&origin=MSN`.

Bhattacharya, R. and Waymire, E. C. (2007). *A basic course in probability theory*, Universitext (Springer, New York), ISBN 978-0-387-71938-2.

Biane, P. (1995). Calcul stochastique non-commutatif, in *Lectures on probability theory (Saint-Flour, 1993)*, *Lecture Notes in Math.*, Vol. 1608 (Springer, Berlin), pp. 1–96, doi:10.1007/BFb0095746, URL `http://dx.doi.org/10.1007/BFb0095746`.

Biane, P. (2009). Matrix valued Brownian motion and a paper by Pólya, in *Séminaire de probabilités XLII*, *Lecture Notes in Math.*, Vol. 1979 (Springer, Berlin), pp. 171–185, doi:10.1007/978-3-642-01763-6_7, URL `http://dx.doi.org/10.1007/978-3-642-01763-6_7`.

Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. (1993). *Efficient and adaptive estimation for semiparametric models*, Johns Hopkins Series in the Mathematical Sciences (Johns Hopkins University Press, Baltimore, MD), ISBN 0-8018-4541-6.

Billingsley, P. (1999). *Convergence of probability measures*, 2nd edn., Wiley Series in Probability and Statistics: Probability and Statistics (John Wiley & Sons Inc., New York), ISBN 0-471-19745-9, a Wiley-Interscience Publication.

Bismut, J.-M. (1973). Conjugate convex functions in optimal stochastic control, *J. Math. Anal. Appl.* **44**, pp. 384–404.

Bismut, J.-M. (1978). Contrôle des systèmes linéaires quadratiques: applications de l'intégrale stochastique, in *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)*, *Lecture Notes in Math.*, Vol. 649 (Springer, Berlin), pp. 180–264.

Bismut, J.-M. (1981a). Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions, *Z. Wahrsch. Verw. Gebiete* **56**, 4, pp. 469–505.

Bismut, J.-M. (1981b). *Mécanique aléatoire*, *Lecture Notes in Mathematics*, Vol. 866 (Springer-Verlag, Berlin), ISBN 3-540-10840-8, with an English summary.

Bismut, J.-M. (1984). *Large deviations and the Malliavin calculus*, *Progress in Mathematics*, Vol. 45 (Birkhäuser Boston Inc., Boston, MA), ISBN 0-8176-3220-4.

Blackwell, D. and Freedman, D. (1964). The tail σ-field of a Markov chain and a theorem of Orey, *Ann. Math. Statist.* **35**, pp. 1291–1295.

Bloom, W. R. and Heyer, H. (1995). *Harmonic analysis of probability measures on hypergroups*, *de Gruyter Studies in Mathematics*, Vol. 20 (Walter de Gruyter & Co., Berlin), ISBN 3-11-012105-0.

Blumenthal, R. M. and Getoor, R. K. (1968). *Markov processes and potential theory*, Pure and Applied Mathematics, Vol. 29 (Academic Press, New York).

Blunck, S. (2001). Analyticity and discrete maximal regularity on L_p-spaces, *J. Funct. Anal.* **183**, 1, pp. 211–230.

Blunck, S. (2002). Some remarks on operator-norm convergence of the Trotter product formula on Banach spaces, *J. Funct. Anal.* **195**, 2, pp. 350–370.

Bobrowski, A. (1997). On the Yosida approximation and the Widder-Arendt representation theorem, *Studia Math.* **124**, 3, pp. 281–290.

Bogachev, V. I., Krylov, N. V. and Röckner, M. (2006). Elliptic equations for measures: regularity and global bounds of densities, *J. Math. Pures Appl. (9)* **85**, 6, pp. 743–757.

Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields, *Ann. Probab.* **10**, 4, pp. 1047–1050.

Border, K. C. (1998). Notes on the kolmogorov extension problem, URL `http://www.hss.caltech.edu/~kcb/Notes/Kolmogorov.pdf`.

Boufoussi, B. and van Casteren, J. (2004a). An approximation result for a nonlinear Neumann boundary value problem via BSDEs, *Stochastic Process. Appl.* **114**, 2, pp. 331–350.

Boufoussi, B. and van Casteren, J. (2004b). An approximation result for a nonlinear Neumann boundary value problem via BSDEs, *Stochastic Process. Appl.* **114**, 2, pp. 331–350, doi:10.1016/j.spa.2004.06.003, URL `http://dx.doi.org/10.1016/j.spa.2004.06.003`.

Boufoussi, B., Van Casteren, J. and Mrhardy, N. (2007). Generalized backward doubly stochastic differential equations and SPDEs with nonlinear Neumann boundary conditions, *Bernoulli* **13**, 2, pp. 423–446.

Bratteli, O. and Robinson, D. W. (1987). *Operator algebras and quantum statistical mechanics. 1*, 2nd edn., Texts and Monographs in Physics (Springer-Verlag, New York), ISBN 0-387-17093-6, C^*- and W^*-algebras, symmetry groups, decomposition of states.

Breiman, L. (1992). *Probability*, *Classics in Applied Mathematics*, Vol. 7 (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA), ISBN 0-89871-296-3, corrected reprint of the 1968 original.

Bressan, A. and Piccoli, B. (2007). *Introduction to the mathematical theory of control*, *AIMS Series on Applied Mathematics*, Vol. 2 (American Institute of Mathematical Sciences (AIMS), Springfield, MO), ISBN 978-1-60133-002-4; 1-60133-002-2.

Browder, F. E. (1963). Nonlinear elliptic boundary value problems, *Bull. Amer. Math. Soc.* **69**, pp. 862–874.

Browder, F. E. (1967). Existence and perturbation theorems for nonlinear maximal monotone operators in Banach spaces, *Bull. Amer. Math. Soc.* **73**, pp. 322–327.

Browder, F. E., de Figueiredo, D. G. and Gupta, C. P. (1970). Maximal monotone operators and nonlinear integral equations of Hammerstein type, *Bull. Amer. Math. Soc.* **76**, pp. 700–705.

Bru, B. and Yor, M. (2002). Comments on the life and mathematical legacy of Wolfgang Doeblin, *Finance Stoch.* **6**, 1, pp. 3–47.

Buck, R. C. (1958). Bounded continuous functions on a locally compact space, *Michigan Math. J.* **5**, pp. 95–104.

Buck, R. C. (1974). Approximation properties of vector valued functions, *Pacific J. Math.* **53**, pp. 85–94.

Buckdahn, R. and Ma, J. (2001a). Stochastic viscosity solutions for nonlinear stochastic partial differential equations. I, *Stochastic Process. Appl.* **93**, 2, pp. 181–204, doi:10.1016/S0304-4149(00)00093-4, URL `http://dx.doi.org/10.1016/S0304-4149(00)00093-4`.

Buckdahn, R. and Ma, J. (2001b). Stochastic viscosity solutions for nonlinear stochastic partial differential equations. II, *Stochastic Process. Appl.* **93**, 2, pp. 205–228, doi:10.1016/S0304-4149(00)00092-2, URL http://dx.doi.org/10.1016/S0304-4149(00)00092-2.

Budhiraja, A., Dupuis, P. and Maroulas, V. (2008). Large deviations for infinite dimensional stochastic dynamical systems, *Ann. Probab.* **36**, 4, pp. 1390–1420.

Bukhvalov, A. V. (1988). Order-bounded operators in vector lattices and spaces of measurable functions, in *Mathematical analysis, Vol. 26 (Russian)*, Itogi Nauki i Tekhniki (Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow), pp. 3–63, 148, translated in J. Soviet Math. **54** (1991), no. 5, 1131–1176.

Burkholder, D. L., Davis, B. J. and Gundy, R. F. (1972). Integral inequalities for convex functions of operators on martingales, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory* (Univ. California Press, Berkeley, Calif.), pp. 223–240.

Burrage, K. and Butcher, J. C. (1979). Stability criteria for implicit Runge-Kutta methods, *SIAM J. Numer. Anal.* **16**, 1, pp. 46–57, doi:10.1137/0716004, URL http://dx.doi.org/10.1137/0716004.

Carlen, E. A., Kusuoka, S. and Stroock, D. W. (1987). Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré Probab. Statist.* **23**, 2, suppl., pp. 245–287.

Carlen, E. A. and Stroock, D. W. (1986). An application of the Bakry-Emery criterion to infinite-dimensional diffusions, in *Séminaire de Probabilités, XX, 1984/85*, *Lecture Notes in Math.*, Vol. 1204 (Springer, Berlin), pp. 341–348.

Cerrai, S. (1994). A Hille-Yosida theorem for weakly continuous semigroups, *Semigroup Forum* **49**, 3, pp. 349–367.

Cerrai, S. (2001). *Second order PDE's in finite and infinite dimension*, *Lecture Notes in Mathematics*, Vol. 1762 (Springer-Verlag, Berlin), ISBN 3-540-42136-X, a probabilistic approach.

Chacon, R. V. and Ornstein, D. S. (1960). A general ergodic theorem, *Illinois J. Math.* **4**, pp. 153–160.

Chan, K. S. and Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving counts, *J. Amer. Statist. Assoc.* **90**, 429, pp. 242–252.

Chavel, I. (2001). *Isoperimetric inequalities*, *Cambridge Tracts in Mathematics*, Vol. 145 (Cambridge University Press, Cambridge), ISBN 0-521-80267-9, differential geometric and analytic perspectives.

Chavel, I. (2005). Topics in isoperimetric inequalities, in *Geometry, spectral theory, groups, and dynamics*, *Contemp. Math.*, Vol. 387 (Amer. Math. Soc., Providence, RI), pp. 89–110.

Chavel, I. and Feldman, E. A. (1991). Modified isoperimetric constants, and large time heat diffusion in Riemannian manifolds, *Duke Math. J.* **64**, 3, pp. 473–499.

Chen, M.-F. (2005). *Eigenvalues, inequalities, and ergodic theory*, Probability and its Applications (New York) (Springer-Verlag London Ltd., London), ISBN 1-85233-868-7.

Chen, M.-F. and Wang, F.-Y. (1997). Estimation of spectral gap for elliptic operators, *Trans. Amer. Math. Soc.* **349**, 3, pp. 1239–1267, doi:10.1090/S0002-9947-97-01812-6, URL `http://dx.doi.org/10.1090/S0002-9947-97-01812-6`.

Chen, M.-F. and Wang, F.-Y. (2000). Cheeger's inequalities for general symmetric forms and existence criteria for spectral gap, *Ann. Probab.* **28**, 1, pp. 235–257.

Chen, M.-F. and Wang, Y.-Z. (2003). Algebraic convergence of Markov chains, *Ann. Appl. Probab.* **13**, 2, pp. 604–627.

Chen, X. (1999). Limit theorems for functionals of ergodic Markov chains with general state space, *Mem. Amer. Math. Soc.* **139**, 664, pp. xiv+203.

Chib, S. (2004). Markov chain Monte Carlo technology, in *Handbook of computational statistics* (Springer, Berlin), pp. 71–102.

Chojnacki, W. (1998). On the equivalence of a theorem of Kisyński and the Hille-Yosida generation theorem, *Proc. Amer. Math. Soc.* **126**, 2, pp. 491–497.

Choquet, G. (1986). La naissance de la théorie des capacités: réflexion sur une expérience personnelle, *C. R. Acad. Sci. Sér. Gén. Vie Sci.* **3**, 4, pp. 385–397.

Chung, K. l. (1962). On the exponential formulas of semi-group theory, *Math. Scand* **10**, pp. 153–162.

Chung, K. L. and Zambrini, J.-C. (2001). *Introduction to random time and quantum randomness*, Monographs of the Portuguese Mathematical Society (McGraw-Hill, Lisbon), ISBN 972-773-114-7.

Cohen, J. E., Iwasa, Y., Răuţu, G., Ruskai, M. B., Seneta, E. and Zbăganu, G. (1993). Relative entropy under mappings by stochastic matrices, *Linear Algebra Appl.* **179**, pp. 211–235.

Connor, S. B. and Fort, G. (2009). State-dependent Foster-Lyapunov criteria for subgeometric convergence of Markov chains, *Stochastic Process. Appl.* **119**, 12, pp. 4176–4193, doi:10.1016/j.spa.2009.10.001, URL `http://dx.doi.org/10.1016/j.spa.2009.10.001`.

Conway, J. B. (1966). The strict topology and compactness in the space of measures, *Bull. Amer. Math. Soc.* **72**, pp. 75–78.

Conway, J. B. (1967). The strict topology and compactness in the space of measures. II, *Trans. Amer. Math. Soc.* **126**, pp. 474–486.

Costa, O. L. V. and Dufour, F. (2005). A sufficient condition for the existence of an invariant probability measure for Markov processes, *J. Appl. Probab.* **42**, 3, pp. 873–878, URL `http://projecteuclid.org/getRecord?id=euclid.jap/1127322035`.

Cox, D. C. (1984). Some sharp martingale inequalities related to Doob's inequality, in *Inequalities in statistics and probability (Lincoln, Neb., 1982)*, *IMS Lecture Notes Monogr. Ser.*, Vol. 5 (Inst. Math. Statist., Hayward, CA), pp. 78–83.

Cox, J. C., Ingersoll, J. E., Jr. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica* **53**, 2, pp. 385–407, doi:10.2307/1911242, URL `http://dx.doi.org/10.2307/1911242`.

Crandall, M. G., Evans, L. C. and Lions, P.-L. (1984). Some properties of viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.* **282**, 2, pp. 487–502, doi:10.2307/1999247, URL `http://dx.doi.org/10.2307/1999247`.

Crandall, M. G., Ishii, H. and Lions, P.-L. (1992a). User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (N.S.)* **27**, 1, pp. 1–67, doi:10.1090/S0273-0979-1992-00266-5, URL `http://dx.doi.org/10.1090/S0273-0979-1992-00266-5`.

Crandall, M. G., Ishii, H. and Lions, P.-L. (1992b). User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (N.S.)* **27**, 1, pp. 1–67.

Crouzeix, M. (1979). Sur la B-stabilité des méthodes de Runge-Kutta, *Numer. Math.* **32**, 1, pp. 75–82.

Crouzeix, M., Hundsdorfer, W. H. and Spijker, M. N. (1983). On the existence of solutions to the algebraic equations in implicit Runge-Kutta methods, *BIT* **23**, 1, pp. 84–91.

Cruzeiro, A. B. and Zambrini, J.-C. (1991). Malliavin calculus and Euclidean quantum mechanics. I. Functional calculus, *J. Funct. Anal.* **96**, 1, pp. 62–95.

Da Prato, G., Elworthy, K. D. and Zabczyk, J. (1995). Strong Feller property for stochastic semilinear equations, *Stochastic Anal. Appl.* **13**, 1, pp. 35–45.

Da Prato, G. and Zabczyk, J. (1992a). *Stochastic equations in infinite dimensions*, *Encyclopedia of Mathematics and its Applications*, Vol. 44 (Cambridge University Press, Cambridge), ISBN 0-521-38529-6.

Da Prato, G. and Zabczyk, J. (1992b). *Stochastic equations in infinite dimensions*, *Encyclopedia of Mathematics and its Applications*, Vol. 44 (Cambridge University Press, Cambridge), ISBN 0-521-38529-6.

Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for infinite-dimensional systems*, *London Mathematical Society Lecture Note Series*, Vol. 229 (Cambridge University Press, Cambridge), ISBN 0-521-57900-7.

Davies, E. B. (1990). *Heat kernels and spectral theory*, *Cambridge Tracts in Mathematics*, Vol. 92 (Cambridge University Press, Cambridge), ISBN 0-521-40997-7.

Dawson, D. A. (1993). Measure-valued Markov processes, in *École d'Été de Probabilités de Saint-Flour XXI—1991*, *Lecture Notes in Math.*, Vol. 1541 (Springer, Berlin), pp. 1–260.

Dawson, D. A. and Perkins, E. A. (1999). Measure-valued processes and renormalization of branching particle systems, in *Stochastic partial differential equations: six perspectives*, *Math. Surveys Monogr.*, Vol. 64 (Amer. Math. Soc., Providence, RI), pp. 45–106.

Del Moral, P., Ledoux, M. and Miclo, L. (2003). On contraction properties of Markov kernels, *Probab. Theory Related Fields* **126**, 3, pp. 395–420.

Delbaen, F., Hu, Y. and Richou, A. (2009). On the uniqueness of solutions to quadratic bsdes with convex generators and unbounded terminal conditions, URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0906.0752`.

Delbaen, F. and Schacher mayer, W. (2006). *The mathematics of arbitrage*, Springer Finance (Springer-Verlag, Berlin), ISBN 978-3-540-21992-7; 3-540-21992-7.

Dellacherie, C. and Meyer, P.-A. (1978). *Probabilities and potential*, *North-Holland Mathematics Studies*, Vol. 29 (North-Holland Publishing Co., Amsterdam), ISBN 0-7204-0701-X.

Demuth, M. and van Casteren, J. A. (2000). *Stochastic spectral theory for selfadjoint Feller operators*, Probability and its Applications (Birkhäuser Verlag, Basel), ISBN 3-7643-5887-4, a functional integration approach.

Deuschel, J.-D. and Stroock, D. W. (1989). *Large deviations*, *Pure and Applied Mathematics*, Vol. 137 (Academic Press Inc., Boston, MA), ISBN 0-12-213150-9.

Di Nunno, G., Øksendal, B. and Proske, F. (2009). *Malliavin calculus for Lévy processes with applications to finance*, Universitext (Springer-Verlag, Berlin), ISBN 978-3-540-78571-2, doi:10.1007/978-3-540-78572-9, URL `http://dx.doi.org/10.1007/978-3-540-78572-9`.

Diestel, J. and Uhl, J. J., Jr. (1977). *Vector measures* (American Mathematical Society, Providence, R.I.), with a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

Dobrushin, R. (1956a). Central limit theorem for non-stationary Markov chains. I, *Teor. Veroyatnost. i Primenen.* **1**, pp. 72–89.

Dobrushin, R. (1956b). Central limit theorem for non-stationary Markov chains. II, *Teor. Veroyatnost. i Primenen.* **1**, pp. 329–383.

Doeblin, W. (1937). Sur les propriétés asymptotiques de mouvements régis par certaines types des chaînes simples, *Bull. Math. Soc. Roumaine des Sciences.* **39**, 1, 2, pp. 57–115.

Doeblin, W. (1940). Eléments d'une théorie générale des chaînes simples constantes de markoff, *Ann. Sci. E.N.S.* **57**, pp. 61–111.

Donner, K. (1982). *Extension of positive operators and Korovkin theorems*, *Lecture Notes in Mathematics*, Vol. 904 (Springer-Verlag, Berlin), ISBN 3-540-11183-2.

Doob, J. L. (1945). Markoff chains—denumerable case, *Trans. Amer. Math. Soc.* **58**, pp. 455–473.

Doob, J. L. (1953). *Stochastic processes* (John Wiley & Sons Inc., New York).

Doob, J. L. (2001). *Classical potential theory and its probabilistic counterpart*, Classics in Mathematics (Springer-Verlag, Berlin), ISBN 3-540-41206-9, reprint of the 1984 edition.

Dorroh, J. R. and Neuberger, J. W. (1993). Lie generators for semigroups of transformations on a Polish space, *Electron. J. Differential Equations* , pp. No. 01, approx. 7 pp. (electronic only).

Driver, B. K. (1995; Last revised: January 29, 2003). A primer on riemannian geometry and stochastic analysis on path spaces, Preprint University of California, San Diego, La Jolla, CA 92093-0112.

Dudley, R. M. (2002). *Real analysis and probability*, *Cambridge Studies in Advanced Mathematics*, Vol. 74 (Cambridge University Press, Cambridge), ISBN 0-521-00754-2, revised reprint of the 1989 original.

Duflo, M. and Revuz, D. (1969). Propriétés asymptotiques des probabilités de transition des processus de Markov récurrents, *Ann. Inst. H. Poincaré Sect. B (N.S.)* **5**, pp. 233–244.

Durrett, R. (1984). *Brownian motion and martingales in analysis*, Wadsworth Mathematics Series (Wadsworth International Group, Belmont, CA), ISBN 0-534-03065-3.

Durrett, R. (1996). *Stochastic calculus*, Probability and Stochastics Series (CRC Press, Boca Raton, FL), ISBN 0-8493-8071-5, a practical introduction.

Dynkin, E. B. and Kuznetsov, S. E. (1996a). Linear additive functionals of superdiffusions and related nonlinear p.d.e, *Trans. Amer. Math. Soc.* **348**, 5, pp. 1959–1987.

Dynkin, E. B. and Kuznetsov, S. E. (1996b). Solutions of $Lu = u^\alpha$ dominated by L-harmonic functions, *J. Anal. Math.* **68**, pp. 15–37, doi:10.1007/BF02790202, URL `http://dx.doi.org/10.1007/BF02790202`.

Eberle, A. (1995). Weak Sobolev spaces and Markov uniqueness of operators, *C. R. Acad. Sci. Paris Sér. I Math.* **320**, 10, pp. 1249–1254.

Eberle, A. (1996). Girsanov-type transformations of local Dirichlet forms: an analytic approach, *Osaka J. Math.* **33**, 2, pp. 497–531, URL `http://projecteuclid.org/getRecord?id=euclid.ojm/1200786925`.

Eberle, A. (1999). *Uniqueness and non-uniqueness of semigroups generated by singular diffusion operators*, *Lecture Notes in Mathematics*, Vol. 1718 (Springer-Verlag, Berlin), ISBN 3-540-66628-1.

Eckmann, J.-P. and Hairer, M. (2001). Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise, *Comm. Math. Phys.* **219**, 3, pp. 523–565.

Eisner, T. (2005). Polynomially bounded C_0-semigroups, *Semigroup Forum* **70**, 1, pp. 118–126, doi:10.1007/s00233-004-0151-z, URL `http://dx.doi.org/10.1007/s00233-004-0151-z`.

Eisner, T. (2010). *Stability of Operators and Operator Semigroups*, *Operator Theory: Advances and Applications*, Vol. 209 (Springer Verlag, Birkhäuser, Boston), ISBN 978-3-0346-0194-8, vIII, 204 p. 1 illus. in color., Hardcover.

Eisner, T. and Zwart, H. (2007). A note on polynomially growing C_0-semigroups, *Semigroup Forum* **75**, 2, pp. 438–445, doi:10.1007/s00233-007-0723-9, URL `http://dx.doi.org/10.1007/s00233-007-0723-9`.

El Karoui, N. and Mazliak, L. (eds.) (1997). *Backward stochastic differential equations*, *Pitman Research Notes in Mathematics Series*, Vol. 364 (Longman, Harlow), ISBN 0-582-30733-3, papers from the study group held at the University of Paris VI, Paris, 1995–1996.

El Karoui, N., Pardoux, E. and Quenez, M. C. (1997). Reflected backward SDEs and American options, in *Numerical methods in finance*, *Publ. Newton Inst.*, Vol. 13 (Cambridge Univ. Press, Cambridge), pp. 215–231.

El Karoui, N. and Quenez, M. C. (1997). Imperfect markets and backward stochastic differential equations, in *Numerical methods in finance*, *Publ. Newton Inst.*, Vol. 13 (Cambridge Univ. Press, Cambridge), pp. 181–214.

Elworthy, K. D. (1982). *Stochastic differential equations on manifolds*, *London Mathematical Society Lecture Note Series*, Vol. 70 (Cambridge University Press, Cambridge), ISBN 0-521-28767-7.

Elworthy, K. D. and Li, X.-M. (1994). Formulae for the derivatives of heat semigroups, *J. Funct. Anal.* **125**, 1, pp. 252–286.

Engel, K.-J. and Nagel, R. (2000). *One-parameter semigroups for linear evolution equations*, *Graduate Texts in Mathematics*, Vol. 194 (Springer-Verlag, New York), ISBN 0-387-98463-1, with contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

Es-Sarhir, A. and Farkas, B. (2005). Positivity of perturbed Ornstein-Uhlenbeck semigroups on $C_b(H)$, *Semigroup Forum* **70**, 2, pp. 208–224.

Es-Sarhir, A. and Stannat, W. (2007). Invariant measures for semilinear SPDE's with local Lipschitz drift coefficients and applications, *Journal of Evolution Equations* **8**, 1, pp. 1424–3202.

Etheridge, A. M. (2000). *An introduction to superprocesses*, *University Lecture Series*, Vol. 20 (American Mathematical Society, Providence, RI), ISBN 0-8218-2706-5.

Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (John Wiley & Sons Inc., New York), ISBN 0-471-08186-8, doi:10.1002/9780470316658, URL `http://dx.doi.org/10.1002/9780470316658`, characterization and convergence.

Evans, L. C. (1998). *Partial differential equations*, *Graduate Studies in Mathematics*, Vol. 19 (American Mathematical Society, Providence, RI), ISBN 0-8218-0772-2.

Fang, S. (1993). Inégalité du type de Poincaré sur un espace de chemins, Univ. de Paris VI preprint.

Farmaki, V. (1995). A geometric characterization of the weak-Radon Nikodým property in dual Banach spaces, *Rocky Mountain J. Math.* **25**, 2, pp. 611–617.

Feng, J., Forde, M. and Fouque, J.-P. (2010). Short-maturity asymptotics for a fast mean-reverting Heston stochastic volatility model, *SIAM J. Financial Math.* **1**, pp. 126–141, doi:10.1137/090745465, URL `http://dx.doi.org/10.1137/090745465`.

Fleming, W. and Soner, H. M. (1993). *Controlled Markov processes and viscosity solutions*, *Applications of Mathematics*, Vol. 25 (Springer Verlag, Berlin).

Foguel, S. R. (1966). Limit theorems for Markov processes, *Trans. Amer. Math. Soc.* **121**, pp. 200–209.

Foguel, S. R. (1969). *The ergodic theory of Markov processes*, Van Nostrand Mathematical Studies, No. 21 (Van Nostrand Reinhold Co., New York).

Foguel, S. R. (1980). *Selected topics in the study of Markov operators*, *Carolina Lecture Series*, Vol. 9 (University of North Carolina Department of Mathematics, Chapel Hill, N.C.).

Fort, G. and Roberts, G. O. (2005). Subgeometric ergodicity of strong Markov processes, *Ann. Appl. Probab.* **15**, 2, pp. 1565–1589, doi:10.1214/105051605000000115, URL `http://dx.doi.org/10.1214/105051605000000115`.

Foss, S. and Konstantopoulos, T. (2004). An overview of some stochastic stability methods, *J. Oper. Res. Soc. Japan* **47**, 4, pp. 275–303.

Franz, U. and Schott, R. (1999). *Stochastic processes and operator calculus on* *quantum groups*, *Mathematics and its Applications*, Vol. 490 (Kluwer Academic Publishers, Dordrecht), ISBN 0-7923-5883-X.

Fuhrman, M. and Tessitore, G. (2002). The Bismut-Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces, *Stoch. Stoch. Rep.* **74**, 1-2, pp. 429–464, doi:10.1080/104S1120290024856, URL http://dx.doi.org/10.1080/104S1120290024856.

Fuhrman, M. and Tessitore, G. (2004). Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces, *Ann. Probab.* **32**, 1B, pp. 607–660, doi:10.1214/aop/1079021459, URL http://dx.doi.org/10.1214/aop/1079021459.

Gatarek, D. and Goldys, B. (1997). On invariant measures for diffusions on Banach spaces, *Potential Anal.* **7**, 2, pp. 539–553, doi:10.1023/A:1008663614438, URL http://dx.doi.org/10.1023/A:1008663614438.

Geissert, M., Lorenzi, L. and Schnaubelt, R. (2009). L^p–regularity for parabolic operators with unbounded time–dependent coefficients, URL http://www.citebase.org/abstract?id=oai:arXiv.org:0903.3117.

Giles, R. (1971). A generalization of the strict topology, *Trans. Amer. Math. Soc.* **161**, pp. 467–474.

Gillespie, D. T. (1992). *Markov processes* (Academic Press Inc., Boston, MA), ISBN 0-12-283955-2, an introduction for physical scientists.

Girardi, M. and Weis, L. (2003). Operator-valued Fourier multiplier theorems on $L_p(X)$ and geometry of Banach spaces, *J. Funct. Anal.* **204**, 2, pp. 320–354.

Goldstein, J. A. (1985). *Semigroups of linear operators and applications*, Oxford Mathematical Monographs (The Clarendon Press Oxford University Press, New York), ISBN 0-19-503540-2.

Goldys, B. and Maslowski, B. (2001). Uniform exponential ergodicity of stochastic dissipative systems, *Czechoslovak Math. J.* **51(126)**, 4, pp. 745–762.

Goldys, B. and Maslowski, B. (2006a). Exponential ergodicity for stochastic reaction-diffusion equations, in *Stochastic partial differential equations and applications—VII*, *Lect. Notes Pure Appl. Math.*, Vol. 245 (Chapman & Hall/CRC, Boca Raton, FL), pp. 115–131.

Goldys, B. and Maslowski, B. (2006b). Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDE's, *Ann. Probab.* **34**, 4, pp. 1451–1496, doi:10.1214/009117905000000800, URL http://dx.doi.org/10.1214/009117905000000800.

Goldys, B. and van Neerven, J. M. A. M. (2003). Transition semigroups of Banach space-valued Ornstein-Uhlenbeck processes, *Acta Appl. Math.* **76**, 3, pp. 283–330.

Gozlan, N. (2007). Characterization of Talagrand's like transportation-cost inequalities on the real line, *J. Funct. Anal.* **250**, 2, pp. 400–425.

Gozlan, N. (2008). A characterization of dimension free concentration in terms of transportation inequalities, URL http://www.citebase.org/abstract?id=oai:arXiv.org:0804.3089, the Annals of Probability 2009, to appear.

Gozlan, N. and Léonard, C. (2007). A large deviation approach to some transportation cost inequalities, *Probab. Theory Related Fields* **139**, 1-2, pp. 235–283.

Gulisashvili, A. and van Casteren, J. A. (2006). *Non-autonomous Kato classes and Feynman-Kac propagators* (World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ), ISBN 981-256-557-4.

Ha, D. M. (2007). *Functional Analysis Volume 1: A Gentle Introduction* (Matrix Editions), ISBN 978-0-9715766-1-2, student-friendly but rigorous book aimed primarily at third or fourth year undergraduates. The pace is slow but thorough, with an abundance of motivations, examples, and counterexamples. Arguments used within proofs are explicitly cited, with references to where they were first proved. Many examples have solutions that make use of several results or concepts, so that students can see how various techniques can blend together into one.

Hahn, W. (1958). Über die Anwendung der Methode von Ljapunov auf Differenzengleichungen, *Math. Ann.* **136**, pp. 430–441.

Hairer, E. and Wanner, G. (1991). *Solving ordinary differential equations. II*, *Springer Series in Computational Mathematics*, Vol. 14 (Springer-Verlag, Berlin), ISBN 3-540-53775-9, stiff and differential-algebraic problems.

Hairer, E. and Wanner, G. (1996). *Solving ordinary differential equations. II*, *Springer Series in Computational Mathematics*, Vol. 14, 2nd edn. (Springer-Verlag, Berlin), ISBN 3-540-60452-9, stiff and differential-algebraic problems.

Hairer, M. (2009). An Introduction to Stochastic PDEs, URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0907.4178`.

Hairer, M. and Mattingly, J. C. (2008a). Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations, *Ann. Probab.* **36**, 6, pp. 2050–2091.

Hairer, M. and Mattingly, J. C. (2008b). Yet another look at Harris' ergodic theorem for Markov chains, URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0810.2777`.

Hairer, M., Mattingly, J. C. and Pardoux, É. (2004). Malliavin calculus for highly degenerate 2D stochastic Navier-Stokes equations, *C. R. Math. Acad. Sci. Paris* **339**, 11, pp. 793–796.

Harris, T. E. (1956). The existence of stationary measures for certain Markov processes, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II* (University of California Press, Berkeley and Los Angeles), pp. 113–124.

Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading, *Stochastic Process. Appl.* **11**, 3, pp. 215–260, doi:10.1016/0304-4149(81)90026-0, URL `http://dx.doi.org/10.1016/0304-4149(81)90026-0`.

Has′minskiĭ, R. Z. (1960). Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations, *Teor. Verojatnost. i Primenen.* **5**, pp. 196–214.

Hazewinkel, M. (ed.) (2001). *Encyclopaedia of mathematics. Supplement. Vol. III* (Kluwer Academic Publishers, Dordrecht), ISBN 1-4020-0198-3.

He, P. and Ying, J. (2009). Irreducibility and uniqueness of stationary distribution, URL http://www.citebase.org/abstract?id=oai:arXiv.org:0902.3296.

Hernández-Lerma, O. and Lasserre, J. B. (2003). *Markov chains and invariant probabilities*, *Progress in Mathematics*, Vol. 211 (Birkhäuser Verlag, Basel), ISBN 3-7643-7000-9.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies* **6**, 2, pp. 327–343.

Hieber, M., Lorenzi, L., Prüss, J., Rhandi, A. and Schnaubelt, R. (2009). Global properties of generalized Ornstein-Uhlenbeck operators on $L^p(\mathbb{R}^N, \mathbb{R}^N)$ with more than linearly growing coefficients, *J. Math. Anal. Appl.* **350**, 1, pp. 100–121.

Hirschfeld, R. A. (1974). Riots, *Nieuw Arch. Wisk. (3)* **22**, pp. 1–43.

Hoh, W. (1994). The martingale problem for a class of pseudo-differential operators, *Math. Ann.* **300**, 1, pp. 121–147.

Hoh, W. (1995a). Feller semigroups generated by pseudo-differential operators, in *Dirichlet forms and stochastic processes (Beijing, 1993)* (de Gruyter, Berlin), pp. 199–206.

Hoh, W. (1995b). Pseudodifferential operators with negative definite symbols and the martingale problem, *Stochastics Stochastics Rep.* **55**, 3-4, pp. 225–252.

Hoh, W. (2000). Pseudo differential operators with negative definite symbols of variable order, *Rev. Mat. Iberoamericana* **16**, 2, pp. 219–241.

Holden, H., Øksendal, B., Ubøe, J. and Zhang, T. (2010). *Stochastic partial differential equations*, 2nd edn., Universitext (Springer, New York), ISBN 978-0-387-89487-4, a modeling, white noise functional approach.

Horowitz, S. (1974). Semi-groups of Markov operators, *Ann. Inst. H. Poincaré Sect. B (N.S.)* **10**, pp. 155–166.

Hsu, E. P. (2002). *Stochastic analysis on manifolds*, *Graduate Studies in Mathematics*, Vol. 38 (American Mathematical Society, Providence, RI), ISBN 0-8218-0802-8.

Hudson, R. L. and Lindsay, J. M. (eds.) (1998). *Quantum probability communications*, QP-PQ, X (World Scientific Publishing Co. Inc., River Edge, NJ), ISBN 981-02-3541-0.

Hwang, C.-R., Hwang-Ma, S.-Y. and Sheu, S.-J. (2005). Accelerating Diffusions, *Annals of Appl. Prob.* **15**, 2, pp. 1433–1444.

Ikeda, N. and Watanabe, S. (1998). *Stochastic differential equations and diffusion processes*, *North-Holland Mathematical Library*, Vol. 24, 2nd edn. (North-Holland, Amsterdam).

Imkeller, P., Reveillac, A. and Richter, A. (2009). Differentiability of quadratic bsde generated by continuous martingales and hedging in incomplete markets, URL http://www.citebase.org/abstract?id=oai:arXiv.org:0907.0941.

In 't Hout, K. J. and Foulon, S. (2010). ADI finite difference schemes for option pricing in the Heston model with correlation, *Int. J. Numer. Anal. Model.* **7**, 2, pp. 303–320.

Jacob, N. (2001). *Pseudo differential operators and Markov processes. Vol. I* (Imperial College Press, London), ISBN 1-86094-293-8, fourier analysis and semigroups.

Jacob, N. (2002). *Pseudo differential operators and Markov processes. Vol. II* (Imperial College Press, London), markov processes.

Jacob, N. (2005). *Pseudo differential operators and Markov processes. Vol. III* (Imperial College Press, London), ISBN 1-86094-568-6, markov processes and applications.

Jakubowski, A. (1997). A non-Skorohod topology on the Skorohod space, *Electron. J. Probab.* **2**, pp. no. 4, 21 pp. (electronic).

Jakubowski, A. (2000). *Skorokhod's Ideas in Probability Theory*, chap. From convergence of functions to convergence of stochastic processes. On Skorokhod's sequential approach to convergence in distribution (Insitute of Mathematics, National Academy of Sciences of Ukraine, Kiev), pp. 179–194.

Jamison, B. (1974). Reciprocal processes, *Z. Wahrscheinlichkeitstheor. Verw. Gebiete* **30**, pp. 65–86.

Jamison, B. and Orey, S. (1967). Markov chains recurrent in the sense of Harris, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8**, 1, pp. 41–48.

Jamison, B., Orey, S. and Pruitt, W. (1965). Convergence of weighted averages of independent random variables, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **4**, pp. 40–44.

Jensen, R. (1989). Uniqueness criteria for viscosity solutions of fully nonlinear elliptic partial differential equations, *Indiana Univ. Math. J.* **38**, 3, pp. 629–667.

Kallenberg, O. (2002). *Foundations of modern probability*, 2nd edn., Probability and its Applications (New York) (Springer-Verlag, New York), ISBN 0-387-95313-2.

Kallenberg, O. (2008). Some local approximations of Dawson-Watanabe superprocesses, *Ann. Probab.* **36**, 6, pp. 2176–2214, doi:10.1214/07-AOP386, URL `http://dx.doi.org/10.1214/07-AOP386`.

Kanovei, V. (2008). *Borel Equivalence Relations: Structure and Classification*, University Lecture Series (American Mathematical Society, Providence, RI), ISBN 10 0821844539, 13 978-0821844533.

Kantorovitz, S. (2010). *Topics in Operator Semigroups*, *Progress in Mathematics*, Vol. 281 (Springer Verlag, Birkhäuser, Boston), ISBN 978-0-8176-4931-9, xIV, 266 p., Hardcover.

Karatzas, I. (1997). *Lectures on the mathematics of finance* (American Mathematical Society, Providence, RI), ISBN 0-8218-0637-8.

Karatzas, I. and Shreve, S. (1991a). *Brownian motion and stochastic calculus*, Vol. 113 (Springer Verlag, Berlin), graduate Texts in Mathematics.

Karatzas, I. and Shreve, S. E. (1991b). *Brownian motion and stochastic calculus*, *Graduate Texts in Mathematics*, Vol. 113, 2nd edn. (Springer-Verlag, New York), ISBN 0-387-97655-8.

Karatzas, I. and Shreve, S. E. (1998). *Methods of mathematical finance* (Springer-Verlag, New York).

Karlin, S. and Taylor, H. M. (1975). *A first course in stochastic processes*, 2nd edn.

(Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London).

Karoui, N. E., Kapoudjian, C., Pardoux, É., Peng, S. G. and Quenez, M. C. (1997). Reflected solutions of backward SDE's, and related obstacle problems for PDE's, *Ann. Probab.* **25**, 2, pp. 702–737.

Kaspi, H. and Mandelbaum, A. (1994). On Harris recurrence in continuous time, *Math. Oper. Res.* **19**, 1, pp. 211–222, doi:10.1287/moor.19.1.211, URL `http://dx.doi.org/10.1287/moor.19.1.211`.

Katilova, N. (2004). *On ergodicity and stability properties of finite Markov chains and their financial interpretations*, Mathematics, University of Antwerp, Middelheimlaan 1, 2020 Antwerp.

Katilova, N. (2005). On Markov and Kolmogorov matrices and their relationship with analytic operators, *New Zealand J. Math.* **34**, 1, pp. 43–60.

Katilova, N. (2008). Markov chains: ergodicity in time-discrete cases, *Rocky Mountain J. Math.* **38**, 3, pp. 941–954, doi:10.1216/RMJ-2008-38-3-941, URL `http://dx.doi.org/10.1216/RMJ-2008-38-3-941`.

Katsaras, A. K. (1983). On the strict topology in the nonlocally convex setting. II, *Acta Math. Hungar.* **41**, 1-2, pp. 77–88, doi:10.1007/BF01994064, URL `http://dx.doi.org/10.1007/BF01994064`.

Kazamaki, N. (1978). The equivalence of two conditions on weighted norm inequalities for martingales, in *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)* (Wiley, New York), pp. 141–152.

Kendall, D. G. and Reuter, G. E. H. (1956). Some pathological Markov processes with a denumerable infinity of states and the associated semigroups of operators on l, in *Proceedings of the International Congress of Mathematicians, 1954, Amsterdam, vol. III* (Erven P. Noordhoff N.V., Groningen), pp. 377–415.

Kingman, J. F. C. and Orey, S. (1964). Ratio limit theorems for Markov chains, *Proc. Amer. Math. Soc.* **15**, pp. 907–910.

Kiselman, C. O. (2000). Plurisubharmonic functions and potential theory in several complex variables, in *Development of mathematics 1950–2000* (Birkhäuser, Basel), pp. 655–714.

Kleinert, H. (2003). *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets* (World Scientific Publishing Co. Inc., River Edge, NJ), to appear.

Kolmogorov, A. N. (1956). *Foundations of the theory of probability* (Chelsea Publishing Co., New York), translation edited by Nathan Morrison, with an added bibliography by A. T. Bharucha-Reid.

Kolmogorov, A. N. (1991). *Selected works of A. N. Kolmogorov. Vol. I*, *Mathematics and its Applications (Soviet Series)*, Vol. 25 (Kluwer Academic Publishers Group, Dordrecht), ISBN 90-277-2796-1, mathematics and mechanics, With commentaries by V. I. Arnol′d, V. A. Skvortsov, P. L. Ul′yanov et al, Translated from the Russian original by V. M. Volosov, Edited and with a preface, foreword and brief biography by V. M. Tikhomirov.

Kolmogorov, A. N. (1993). *Selected works of A. N. Kolmogorov. Vol. III*, *Math-*

ematics and its Applications (Soviet Series), Vol. 27 (Kluwer Academic Publishers Group, Dordrecht), ISBN 90-277-2798-8, information theory and the theory of algorithms, With a biography of Kolmogorov by N. N. Bogolyubov, B. V. Gnedenko and S. L. Sobolev, With commentaries by R. L. Dobrushin, A. Kh. Shen′, V. M. Tikhomirov, Ya. M. Barzdin [Jānis Bārzdiņš], Ya. G. Sinaĭ, V. A. Uspenski [V. A. Uspenskiĭ] and A. L. Semyonov, Edited by A. N. Shiryayev [A. N. Shiryaev], Translated from the 1987 Russian original by A. B. Sossinsky [A. B. Sosinskiĭ].

Kolokoltsov, V. N. (2004a). Measure-valued limits of interacting particle systems with k-nary interactions. II. Finite-dimensional limits, *Stoch. Stoch. Rep.* **76**, 1, pp. 45–58.

Kolokoltsov, V. N. (2004b). On Markov processes with decomposable pseudo-differential generators, *Stoch. Stoch. Rep.* **76**, 1, pp. 1–44.

Korovkin, P. P. (1960). *Linear operators and approximation theory*, Translated from the Russian ed. (1959). Russian Monographs and Texts on Advanced Mathematics and Physics, Vol. III (Gordon and Breach Publishers, Inc., New York).

Kotelenez, P. M. and Kurtz, T. G. (2010). Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type, *Probab. Theory Related Fields* **146**, 1-2, pp. 189–222, doi:10.1007/s00440-008-0188-0, URL `http://dx.doi.org/10.1007/s00440-008-0188-0`.

Krengel, U. (1985). *Ergodic theorems*, *de Gruyter Studies in Mathematics*, Vol. 6 (Walter de Gruyter & Co., Berlin), ISBN 3-11-008478-3, with a supplement by Antoine Brunel.

Krylov, N. (2002). A simple proof of a result of A. Novikov, *arXiv:math.PR/ 0207013 v1* , 1, p. 3.

Krylov, N. V. and Rozovskii, B. L. (2007). Stochastic evolution equations [mr0570795], in *Stochastic differential equations: theory and applications*, *Interdiscip. Math. Sci.*, Vol. 2 (World Sci. Publ., Hackensack, NJ), pp. 1–69.

Kühnemund, F. (2003). A Hille-Yosida theorem for bi-continuous semigroups, *Semigroup Forum* **67**, 2, pp. 205–225, doi:10.1007/s00233-002-5000-3, URL `http://dx.doi.org/10.1007/s00233-002-5000-3`.

Kullback, S. (1997). *Information theory and statistics* (Dover Publications Inc., Mineola, NY), ISBN 0-486-69684-7, reprint of the second (1968) edition.

Kurtz, T. G. (1998). Martingale problems for conditional distributions of Markov processes, *Electron. J. Probab.* **3**, pp. no. 9, 29 pp. (electronic).

Kurtz, T. G. and Nappo, G. (2010). *The filtered martingale problem. Handbook on Nonlinear Filtering* (Oxford University Press), to appear.

Kurtz, T. G. and Ocone, D. L. (1988). Unique characterization of conditional distributions in nonlinear filtering, *Ann. Probab.* **16**, 1, pp. 80–107.

Kusuoka, S. and Stroock, D. (1984). Applications of the Malliavin calculus. I, in *Stochastic analysis (Katata/Kyoto, 1982)*, *North-Holland Math. Library*, Vol. 32 (North-Holland, Amsterdam), pp. 271–306.

Kusuoka, S. and Stroock, D. (1985). Applications of the Malliavin calculus. II, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**, 1, pp. 1–76.

Kusuoka, S. and Stroock, D. (1987). Applications of the Malliavin calculus. III, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34**, 2, pp. 391–442.

Labsker, L. G. (1982). Korovkin sets in a Banach space for sets of linear functionals, *Mat. Zametki* **31**, 1, pp. 93–112, 159.

Ledoux, M. (1992). On an integral criterion for hypercontractivity of diffusion semigroups and extremal functions, *J. Funct. Anal.* **105**, 2, pp. 444–465.

Ledoux, M. (2000). The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse Math. (6)* **9**, 2, pp. 305–366, probability theory.

Ledoux, M. (2004). Spectral gap, logarithmic Sobolev constant, and geometric bounds, in *Surveys in differential geometry. Vol. IX*, Surv. Differ. Geom., IX (Int. Press, Somerville, MA), pp. 219–240.

Lejay, A. (2002). BSDE driven by Dirichlet process and semi-linear parabolic PDE. Application to homogenization, *Stochastic Process. Appl.* **97**, 1, pp. 1–39.

Lejay, A. (2004). A probabilistic representation of the solution of some quasi-linear PDE with a divergence form operator. Application to existence of weak solutions of FBSDE, *Stochastic Process. Appl.* **110**, 1, pp. 145–176.

Lepeltier, J. P. and San Martin, J. (1997). Backward stochastic differential equations with continuous coefficient, *Statist. Probab. Lett.* **32**, 4, pp. 425–430, doi:10.1016/S0167-7152(96)00103-4, URL `http://dx.doi.org/10.1016/S0167-7152(96)00103-4`.

Lévy, P. (1937). *Théorie de l'addition des variables aléatoires* (Gauthier-Villars, Paris).

Li, X.-M. (1994). Strong p-completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds, *Probab. Theory Related Fields* **100**, 4, pp. 485–511.

Liggett, T. M. (2005). *Interacting particle systems*, Classics in Mathematics (Springer-Verlag, Berlin), reprint of the 1985 original.

Lin, M. (1972a). Semi-groups of Markov operators, *Boll. Un. Mat. Ital. (4)* **6**, pp. 20–44.

Lin, M. (1972b). Strong ratio limit theorems for Markov processes, *Ann. Math. Statist.* **43**, pp. 569–579.

Lin, M. (1974). On the uniform ergodic theorem. II, *Proc. Amer. Math. Soc.* **46**, pp. 217–225.

Lin, M. (1975). Quasi-compactness and uniform ergodicity of Markov operators, *Ann. Inst. H. Poincaré Sect. B (N.S.)* **11**, 4, pp. 345–354 (1976).

Lindvall, T. (1991). W. Doeblin 1915–1940, *Ann. Probab.* **19**, 3, pp. 929–934.

Lowen, R. (1997). *Approach spaces*, Oxford Mathematical Monographs (The Clarendon Press Oxford University Press, New York), ISBN 0-19-850030-0, the missing link in the topology-uniformity-metric triad, Oxford Science Publications.

Lumer, G. and Phillips, R. S. (1961). Dissipative operators in a Banach space, *Pacific J. Math.* **11**, pp. 679–698.

Lyubich, Y. (1999). Spectral localization, power boundedness and invariant subspaces under Ritt's type condition, *Studia Math.* **134**, 2, pp. 153–167.

Ma, J. and Yong, J. (1999). *Forward-backward stochastic differential equations and their applications*, *Lecture Notes in Mathematics*, Vol. 1702 (Springer-

Verlag, Berlin), ISBN 3-540-65960-9.

Ma, J., Zhang, J. and Zheng, Z. (2008). Weak solutions for forward-backward SDEs—a martingale problem approach, *Ann. Probab.* **36**, 6, pp. 2092–2125.

Malliavin, P. (1978). Stochastic calculus of variation and hypoelliptic operators, in *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)* (Wiley, New York), pp. 195–263.

Malliavin, P. and Thalmaier, A. (2006). *Stochastic calculus of variations in mathematical finance*, Springer Finance (Springer-Verlag, Berlin), ISBN 978-3-540-43431-3; 3-540-43431-3.

Mao, Y.-H. (2002). Strong ergodicity for Markov processes by coupling methods, *J. Appl. Probab.* **39**, 4, pp. 839–852.

Marsden, M. J. and Riemenschneider, S. D. (1974). Korovkin theorems for integral operators with kernels of finite oscillation, *Canad. J. Math.* **26**, pp. 1390–1404.

Maslowski, B. and Seidler, J. (1998). Invariant measures for nonlinear SPDE's: uniqueness and stability, *Arch. Math. (Brno)* **34**, 1, pp. 153–172, equadiff 9 (Brno, 1997).

Matsuda, M. (1985). On Saab's characterizations of weak Radon-Nikodým sets, *Publ. Res. Inst. Math. Sci.* **21**, 5, pp. 921–941.

Mazet, O. (2002). A characterization of Markov property for semigroups with invariant measure, *Potential Anal.* **16**, 3, pp. 279–287.

Mazliak, L. (2007). On the exchanges between Wolfgang Doeblin and Bohuslav Hostinský, *Rev. Histoire Math.* **13**, 1, pp. 155–180.

Metafune, G., Pallara, D. and Wacker, M. (2002a). Feller semigroups on $\mathbb{R}^N$, *Semigroup Forum* **65**, 2, pp. 159–205.

Metafune, G., Prüss, J., Rhandi, A. and Schnaubelt, R. (2002b). The domain of the Ornstein-Uhlenbeck operator on an L^p-space with invariant measure, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **1**, 2, pp. 471–485.

Meyer, P.-A. (1966). *Probability and potentials* (Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London).

Meyer, P.-A. (1993). *Quantum probability for probabilists*, Lecture Notes in Mathematics, Vol. 1538 (Springer-Verlag, Berlin), ISBN 3-540-56476-4.

Meyn, S. (2008). *Control techniques for complex networks* (Cambridge University Press, Cambridge), ISBN 978-0-521-88441-9.

Meyn, S. P. and Tweedie, R. L. (1993a). Generalized resolvents and Harris recurrence of Markov processes, in *Doeblin and modern probability (Blaubeuren, 1991)*, *Contemp. Math.*, Vol. 149 (Amer. Math. Soc., Providence, RI), pp. 227–250.

Meyn, S. P. and Tweedie, R. L. (1993b). *Markov chains and stochastic stability*, Communications and Control Engineering Series (Springer-Verlag London Ltd., London), ISBN 3-540-19832-6, (new version September 2005: http://probability.ca/MT/; Second edition, Cambridge University Press, 2009.).

Miana, P. J. (2005). Uniformly bounded limit of fractional homomorphisms, *Proc. Amer. Math. Soc.* **133**, 9, pp. 2569–2575 (electronic).

Mikosch, T. (1998). *Elementary stochastic calculus—with finance in view*, *Advanced Series on Statistical Science & Applied Probability*, Vol. 6 (World Scientific Publishing Co. Inc., River Edge, NJ), ISBN 981-02-3543-7.

Mininni, R. M. and Romanelli, S. (2003). Martingale estimating functions for Feller diffusion processes generated by degenerate elliptic operators, *J. Concr. Appl. Math.* **1**, 3, pp. 191–216.

Minty, G. J. (1963). on a “monotonicity” method for the solution of non-linear equations in Banach spaces, *Proc. Nat. Acad. Sci. U.S.A.* **50**, pp. 1038–1041.

Nadirashvili, N. (1997). Nonuniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **24**, 3, pp. 537–549.

Nagy, B. and Zemánek, J. (1999). A resolvent condition implying power boundedness, *Studia Math.* **134**, 2, pp. 143–151.

Nelson, E. (1967). *Dynamical theories of Brownian motion* (Princeton University Press, Princeton, N.J.).

Neveu, J. (1965). *Mathematical foundations of the calculus of probability*, Translated by Amiel Feinstein (Holden-Day Inc., San Francisco, Calif.).

Neveu, J. (1979). The filling scheme and the Chacon-Ornstein theorem, *Israel J. Math.* **33**, 3-4, pp. 368–377 (1980), doi:10.1007/BF02762170, URL `http://dx.doi.org/10.1007/BF02762170`, a collection of invited papers on ergodic theory.

Nishishiraho, T. (1998). Korovkin sets and mean ergodic theorems, *J. Convex Anal.* **5**, 1, pp. 147–151.

Norris, J. (1986). Simplified Malliavin calculus, in *Séminaire de Probabilités, XX, 1984/85*, *Lecture Notes in Math.*, Vol. 1204 (Springer, Berlin), pp. 101–130.

Norris, J. R. (1998). *Markov chains*, *Cambridge Series in Statistical and Probabilistic Mathematics*, Vol. 2 (Cambridge University Press, Cambridge), ISBN 0-521-48181-3, reprint of 1997 original.

Novikov, A. A. (1973). On moment inequalities and identities for stochastic integrals, in *Proceedings of the Second Japan-USSR Symposium on Probability Theory (Kyoto, 1972)* (Springer, Berlin), pp. 333–339. Lecture Notes in Math., Vol. 330.

Nualart, D. (1995). *The Malliavin calculus and related topics*, Probability and its Applications (New York) (Springer-Verlag, New York).

Nualart, D. (1998). Analysis on Wiener space and anticipating stochastic calculus, in *Lectures on probability theory and statistics (Saint-Flour, 1995)*, *Lecture Notes in Math.*, Vol. 1690 (Springer, Berlin), pp. 123–227, doi: 10.1007/BFb0092538, URL `http://dx.doi.org/10.1007/BFb0092538`.

Nualart, D. (2006). *The Malliavin calculus and related topics*, 2nd edn., Probability and its Applications (New York) (Springer-Verlag, Berlin), ISBN 978-3-540-28328-7; 3-540-28328-5.

Nualart, D. (2009). *Malliavin calculus and its applications*, *CBMS Regional Conference Series in Mathematics*, Vol. 110 (Published for the Conference Board of the Mathematical Sciences, Washington, DC), ISBN 978-0-8218-4779-4.

Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains, *Z. Wahrsch. Verw. Gebiete* **43**, 4, pp. 309–318.

Nummelin, E. (1984). *General irreducible Markov chains and nonnegative operators*, *Cambridge Tracts in Mathematics*, Vol. 83 (Cambridge University Press, Cambridge), ISBN 0-521-25005-6.

Nummelin, E. and Arjas, E. (1976). A direct construction of the R-invariant measure for a Markov chain on a general state space, *Ann. Probability* **4**, 4, pp. 674–679.

Nummelin, E. and Tuominen, P. (1982). Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory, *Stochastic Process. Appl.* **12**, 2, pp. 187–202.

Nummelin, E. and Tuominen, P. (1983). The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory, *Stochastic Process. Appl.* **15**, 3, pp. 295–311.

N'zi, M. and Owo, J.-M. (2009). Backward doubly stochastic differential equations with discontinuous coefficients, *Statist. Probab. Lett.* **79**, 7, pp. 920–926, doi:10.1016/j.spl.2008.11.011, URL `http://dx.doi.org/10.1016/j.spl.2008.11.011`.

Øksendal, B. and Reikvam, K. (1998). Viscosity solutions of optimal stopping problems, *Stochastics Stochastics Rep.* **62**, 3-4, pp. 285–301.

Orey, S. (1959). Recurrent Markov chains, *Pacific J. Math.* **9**, pp. 805–827.

Orey, S. (1962). An ergodic theorem for Markov chains, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **1**, pp. 174–176.

Orey, S. (1964). Potential kernels for recurrent Markov chains, *J. Math. Anal. Appl.* **8**, pp. 104–132.

Orey, S. (1971). *Lecture notes on limit theorems for Markov chain transition probabilities* (Van Nostrand Reinhold Co., London), van Nostrand Reinhold Mathematical Studies, No. 34.

Ornstein, D. S. (1969). Random walks. I, II, *Trans. Amer. Math. Soc. 138 (1969), 1-43; ibid.* **138**, pp. 45–60.

Özarslan, M. A. and Duman, O. (2007). MKZ type operators providing a better estimation on $[1/2, 1)$, *Canad. Math. Bull.* **50**, 3, pp. 434–439.

Pardoux, É. (1998a). Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, in *Stochastic analysis and related topics, VI (Geilo, 1996)*, *Progr. Probab.*, Vol. 42 (Birkhäuser Boston, Boston, MA), pp. 79–127.

Pardoux, É. (1998b). Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, in *Stochastic analysis and related topics, VI (Geilo, 1996)*, *Progr. Probab.*, Vol. 42 (Birkhäuser Boston, Boston, MA), pp. 79–127.

Pardoux, É. (1999). BSDEs, weak convergence and homogenization of semilinear PDEs, in *Nonlinear analysis, differential equations and control (Montreal, QC, 1998)* (Kluwer Acad. Publ., Dordrecht), pp. 503–549.

Pardoux, É. and Peng, S. G. (1990). Adapted solution of a backward stochastic differential equation, *Systems Control Lett.* **14**, 1, pp. 55–61, doi:10.1016/

0167-6911(90)90082-6, URL http://dx.doi.org/10.1016/0167-6911(90)90082-6.

Pardoux, É. and Peng, S. G. (1994). Backward doubly stochastic differential equations and systems of quasilinear SPDEs, *Probab. Theory Related Fields* **98**, 2, pp. 209–227, doi:10.1007/BF01192514, URL http://dx.doi.org/10.1007/BF01192514.

Pardoux, É. and Zhang, S. (1998). Generalized BSDEs and nonlinear Neumann boundary value problems, *Probab. Theory Related Fields* **110**, 4, pp. 535–558.

Parthasarathy, K. R. (2005). *Introduction to probability and measure*, *Texts and Readings in Mathematics*, Vol. 33 (Hindustan Book Agency, New Delhi), ISBN 81-85931-55-0, corrected reprint of the 1977 original.

Pazy, A. (1983a). *Semigroups of linear operators and applications to partial differential equations*, *Applied Mathematical Sciences*, Vol. 44 (Springer-Verlag, New York), ISBN 0-387-90845-5.

Pazy, A. (1983b). *Semigroups of linear operators and applications to partial differential equations*, *Applied Mathematical Sciences*, Vol. 44 (Springer-Verlag, New York), ISBN 0-387-90845-5.

Peng, S. G. (1991). A nonlinear Feynman-Kac formula and applications, in *Control theory, stochastic analysis and applications (Hangzhou, 1991)* (World Sci. Publ., River Edge, NJ), pp. 173–184.

Petersen, K. (1989). *Ergodic theory*, *Cambridge Studies in Advanced Mathematics*, Vol. 2 (Cambridge University Press, Cambridge), ISBN 0-521-38997-6, corrected reprint of the 1983 original.

Phelps, R. R. (1993). *Convex functions, monotone operators and differentiability*, *Lecture Notes in Mathematics*, Vol. 1364, 2nd edn. (Springer-Verlag, Berlin), ISBN 3-540-56715-1.

Pietsch, A. (2007). *History of Banach spaces and linear operators* (Birkhäuser Boston Inc., Boston, MA), ISBN 978-0-8176-4367-6; 0-8176-4367-2.

Prohorov, Y. V. (1956). Convergence of random processes and limit theorems in probability theory, *Teor. Veroyatnost. i Primenen.* **1**, pp. 177–238.

Prolla, J. B. (1977). *Approximation of vector valued functions* (North-Holland Publishing Co., Amsterdam), ISBN 0-444-85030-9, north-Holland Mathematics Studies, Vol. 25, Notas de Matemática, No. 61. [Notes on Mathematics, No. 61].

Prolla, J. B. (1993). The Weierstrass-Stone theorem in absolute valued division rings, *Indag. Math. (N.S.)* **4**, 1, pp. 71–78.

Prolla, J. B. and Navarro, S. (1997). Approximation results in the strict topology, *Ann. Math. Blaise Pascal* **4**, 2, pp. 61–82.

Protter, P. E. (2005). *Stochastic integration and differential equations*, *Stochastic Modelling and Applied Probability*, Vol. 21 (Springer-Verlag, Berlin), ISBN 3-540-00313-4, second edition. Version 2.1, Corrected third printing.

Prüss, J. (2002). Maximal regularity for evolution equations in L_p-spaces, *Conf. Semin. Mat. Univ. Bari* , 285, pp. 1–39 (2003).

Qian, Z. (1998). A comparison theorem for an elliptic operator, *Potential Anal.* **8**, 2, pp. 137–142.

Räbiger, F., Rhandi, A. and Schnaubelt, R. (1996). Perturbation and an abstract characterization of evolution semigroups, *J. Math. Anal. Appl.* **198**, 2, pp. 516–533.

Räbiger, F., Schnaubelt, R., Rhandi, A. and Voigt, J. (2000). Non-autonomous Miyadera perturbations, *Differential Integral Equations* **13**, 1-3, pp. 341–368.

Rachev, S. T. (1991). *Probability metrics and the stability of stochastic models*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics (John Wiley & Sons Ltd., Chichester), ISBN 0-471-92877-1.

Ramm, A. G. (2007). *Dynamical systems method for solving operator equations*, *Mathematics in Science and Engineering*, Vol. 208 (Elsevier B. V., Amsterdam), ISBN 978-0-444-52795-0; 0-444-52795-8.

Rao, K. M. (1969). On decomposition theorems of Meyer, *Math. Scand.* **24**, pp. 66–78.

Rebolledo, R., Rezende, J. and Zambrini, J.-C. (eds.) (2004). *Stochastic analysis and mathematical physics (SAMP/ANESTOC 2002)* (World Scientific Publishing Co. Inc., River Edge, NJ), ISBN 981-256-064-5.

Renardy, M. and Rogers, R. C. (2004). *An introduction to partial differential equations*, *Texts in Applied Mathematics*, Vol. 13, 2nd edn. (Springer-Verlag, New York), ISBN 0-387-00444-0.

Reuter, G. E. H. (1969). Remarks on a Markov chain example of Kolmogorov, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **13**, pp. 315–320.

Reveillac, A. (2009). On the orthogonal component of bsdes in a markovian setting, URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0907.1071`.

Revuz, D. (1975). *Markov chains* (North-Holland Publishing Co., Amsterdam), north-Holland Mathematical Library, Vol. 11.

Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Vol. 293, 3rd edn. (Springer-Verlag, Berlin), ISBN 3-540-64325-7.

Riddle, L. H. (1984). Dunford-Pettis operators and weak Radon-Nikodým sets, *Proc. Amer. Math. Soc.* **91**, 2, pp. 254–256.

Roberts, G. O. and Rosenthal, J. S. (1996). Quantitative bounds for convergence rates of continuous time Markov processes, *Electron. J. Probab.* **1**, pp. no. 9, approx. 21 pp. (electronic), URL `http://www.math.washington.edu/~ejpecp/EjpVol1/paper9.abs.html`.

Rockafellar, R. T. (1969). Local boundedness of nonlinear, monotone operators, *Michigan Math. J.* **16**, pp. 397–407.

Rockafellar, R. T. (1997). *Convex analysis*, Princeton Landmarks in Mathematics (Princeton University Press, Princeton, NJ), ISBN 0-691-01586-4, reprint of the 1970 original, Princeton Paperbacks.

Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge Mathematical Library (Cambridge University Press, Cambridge), ISBN 0-521-77593-0, itô calculus, Reprint of the second

(1994) edition.

Roth, J.-P. (1976). Formule de représentation et troncature des formes de Dirichlet sur $\mathbf{R}^m$, in *Séminaire de Théorie du Potentiel de Paris, No. 2 (Univ. Paris, Paris, 1975–1976)* (Springer, Berlin), pp. 260–274. Lecture Notes in Math., Vol. 563.

Rothaus, O. S. (1981a). Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities, *J. Funct. Anal.* **42**, 1, pp. 102–109.

Rothaus, O. S. (1981b). Logarithmic Sobolev inequalities and the spectrum of Schrödinger operators, *J. Funct. Anal.* **42**, 1, pp. 110–120.

Rothaus, O. S. (1986). Hypercontractivity and the Bakry-Emery criterion for compact Lie groups, *J. Funct. Anal.* **65**, 3, pp. 358–367.

Rudin, W. (1991). *Functional analysis*, 2nd edn., International Series in Pure and Applied Mathematics (McGraw-Hill Inc., New York), ISBN 0-07-054236-8.

Ruess, W. (1977). The strict topology and (DF) spaces, in *Functional analysis: surveys and recent results (Proc. Conf., Paderborn, 1976)* (North-Holland, Amsterdam), pp. 105–118. North–Holland Math. Studies, Vol. 27; Notas de Mat., No. 63.

Sanz-Solé, M. (2005). *Malliavin calculus*, Fundamental Sciences (EPFL Press, Lausanne), ISBN 2-940222-06-1; 0-8493-4030-6, with applications to stochastic partial differential equations.

Schaefer, H. H. (1971). *Topological vector spaces* (Springer-Verlag, New York), third printing corrected, Graduate Texts in Mathematics, Vol. 3.

Schwartz, L. (1973). *Radon measures on arbitrary topological spaces and cylindrical measures* (Published for the Tata Institute of Fundamental Research, Bombay by Oxford University Press, London), tata Institute of Fundamental Research Studies in Mathematics, No. 6.

Seidler, J. (1997). Ergodic behaviour of stochastic parabolic equations, *Czechoslovak Math. J.* **47(122)**, 2, pp. 277–316, doi:10.1023/A:1022821729545, URL `http://dx.doi.org/10.1023/A:1022821729545`.

Seneta, E. (1981). *Nonnegative matrices and Markov chains*, 2nd edn., Springer Series in Statistics (Springer-Verlag, New York), ISBN 0-387-90598-7.

Seneta, E. (1993). Applications of ergodicity coefficients to homogeneous Markov chains, in *Doeblin and modern probability (Blaubeuren, 1991)*, *Contemp. Math.*, Vol. 149 (Amer. Math. Soc., Providence, RI), pp. 189–199.

Sentilles, F. D. (1970). The strict topology on bounded sets, *Pacific J. Math.* **34**, pp. 529–540.

Sentilles, F. D. (1972). Bounded continuous functions on a completely regular space, *Trans. Amer. Math. Soc.* **168**, pp. 311–336.

Sethuraman, S. and Varadhan, S. R. S. (2005). A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains, *Electron. J. Probab.* **10**, pp. no. 36, 1221–1235 (electronic).

Shanbhag, D. N. and Rao, C. R. (eds.) (2001). *Stochastic processes: theory and methods*, *Handbook of Statistics*, Vol. 19 (North-Holland Publishing Co., Amsterdam), ISBN 0-444-50014-6.

Sharpe, M. (1988). *General theory of Markov processes*, *Pure and Applied Mathematics*, Vol. 133 (Academic Press Inc., Boston, MA), ISBN 0-12-639060-6.

Shepp, L. A. (1962). Symmetric random walk, *Trans. Amer. Math. Soc.* **104**, pp. 144–153.

Shepp, L. A. (1964). Recurrent random walks with arbitrarily large steps, *Bull. Amer. Math. Soc.* **70**, pp. 540–542.

Sheu, S. J. (1984). Stochastic control and principal eigenvalue, *Stochastics* **11**, 3-4, pp. 191–211.

Shiryayev, A. N. (1984). *Probability*, *Graduate Texts in Mathematics*, Vol. 95 (Springer-Verlag, New York), ISBN 0-387-90898-6, translated from the Russian by R. P. Boas.

Showalter, R. E. (1997). *Monotone operators in Banach space and nonlinear partial differential equations*, *Mathematical Surveys and Monographs*, Vol. 49 (American Mathematical Society, Providence, RI), ISBN 0-8218-0500-2.

Shreve, S. E. (2004). *Stochastic calculus for finance. II*, Springer Finance (Springer-Verlag, New York), continuous-time models.

Simon, B. (1979). *Functional integration and quantum physics*, *Pure and Applied Mathematics*, Vol. 86 (Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York), ISBN 0-12-644250-9.

Skorokhod, A. V. (1965). *Studies in the theory of random processes*, Translated from the Russian by Scripta Technica, Inc (Addison-Wesley Publishing Co., Inc., Reading, Mass.).

Skorokhod, A. V. (1989). *Asymptotic methods in the theory of stochastic differential equations*, *Translations of Mathematical Monographs*, Vol. 78 (American Mathematical Society, Providence, RI), ISBN 0-8218-4531-4, translated from the Russian by H. H. McFaden.

Soner, H. M. (1997). Controlled Markov processes, viscosity solutions and applications to mathematical finance, in *Viscosity solutions and applications (Montecatini Terme, 1995)*, *Lecture Notes in Math.*, Vol. 1660 (Springer, Berlin), pp. 134–185.

Spitzer, F. (1964). *Principles of random walk*, The University Series in Higher Mathematics (D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London).

Stachurski, J. (2009). *Economic dynamics* (MIT Press, Cambridge, MA), ISBN 978-0-262-01277-5, theory and computation.

Stannat, W. (2000). On the validity of the log-Sobolev inequality for symmetric Fleming-Viot operators, *Ann. Probab.* **28**, 2, pp. 667–684.

Stannat, W. (2005). On the Poincaré inequality for infinitely divisible measures, *Potential Anal.* **23**, 3, pp. 279–301.

Stannat, W. (2006). Stability of the optimal filter via pointwise gradient estimates, in *Stochastic partial differential equations and applications—VII*, *Lect. Notes Pure Appl. Math.*, Vol. 245 (Chapman & Hall/CRC, Boca Raton, FL), pp. 281–293.

Stein, E. M. and Shakarchi, R. (2005). *Real analysis*, Princeton Lectures in Analysis, III (Princeton University Press, Princeton, NJ), ISBN 0-691-11386-6, measure theory, integration, and Hilbert spaces.

Stettner, L. (1986). On the existence and uniqueness of invariant measure for continuous time Markov processes, Technical report lcds #86-18, Brown University, Lefeschetz Center for Dynamical Systems, Providence, RI.

The paper attempts to find fairly general conditions under which the existence and uniqueness of invariant measure is guaranteed. The obtained results are new or generalize at least slightly known theorems. The author introduces a terminology: weak, strong Harris, strong recurrence. Two Sections concern general standard processes. The other section restricts it to Feller or strong Feller standard processes. Three examples are considered to illustrate possible unpleasant situations one can meet in the general theory.

Stettner, Ł. (1994). Remarks on ergodic conditions for Markov processes on Polish spaces, *Bull. Polish Acad. Sci. Math.* **42**, 2, pp. 103–114.

Stroock, D. and Varadhan, S. S. (1979). *Multidimensional diffusion processes, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Vol. 233 (Springer-Verlag).

Stroock, D. W. (1981). The Malliavin calculus, a functional analytic approach, *J. Funct. Anal.* **44**, 2, pp. 212–257.

Stroock, D. W. (1999). *A concise introduction to the theory of integration*, 3rd edn. (Birkhäuser Boston Inc., Boston, MA), ISBN 0-8176-4073-8.

Stroock, D. W. (2000). *Probability theory, an analytic view* (Cambridge University Press, Cambridge), ISBN 0-521-66349-0/pbk.

Stroock, D. W. and Varadhan, S. R. S. (2006). *Multidimensional diffusion processes*, Classics in Mathematics (Springer-Verlag, Berlin), ISBN 978-3-540-28998-2; 3-540-28998-4, reprint of the 1997 edition.

Stroock, D. W. and Zegarliński, B. (1992). The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition, *Comm. Math. Phys.* **144**, 2, pp. 303–323.

Taira, K. (1992). On the existence of Feller semigroups with boundary conditions, *Mem. Amer. Math. Soc.* **99**, 475, pp. viii+65.

Taira, K. (1997). Analytic Feller semigroups, *Conf. Semin. Mat. Univ. Bari* , 267, pp. ii+29.

Tarafdar, E. U. and Chowdhury, M. S. R. (2008). *Topological methods for set-valued nonlinear analysis* (World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ), ISBN 978-981-270-467-2; 981-270-467-1.

Thieullen, M. (1993). Second order stochastic differential equations and non-Gaussian reciprocal diffusions, *Probab. Theory Related Fields* **97**, 1-2, pp. 231–257, doi:10.1007/BF01199322, URL `http://dx.doi.org/10.1007/BF01199322`.

Thieullen, M. (1998). Reciprocal diffusions and symmetries, *Stochastics Stochastics Rep.* **65**, 1-2, pp. 41–77.

Thieullen, M. and Zambrini, J. C. (1997a). Probability and quantum symmetries. I. The theorem of Noether in Schrödinger's Euclidean quantum mechanics, *Ann. Inst. H. Poincaré Phys. Théor.* **67**, 3, pp. 297–338.

Thieullen, M. and Zambrini, J. C. (1997b). Probability and quantum symmetries. I. The theorem of Noether in Schrödinger's Euclidean quantum mechanics, *Ann. Inst. H. Poincaré Phys. Théor.* **67**, 3, pp. 297–338, URL `http://www.numdam.org/item?id=AIHPA_1997__67_3_297_0`.

Thieullen, M. and Zambrini, J. C. (1997c). Symmetries in the stochastic calculus of variations, *Probab. Theory Related Fields* **107**, 3, pp. 401–427.

Thieullen, M. and Zambrini, J. C. (1997d). Symmetries in the stochastic calculus of variations, *Probab. Theory Related Fields* **107**, 3, pp. 401–427, doi:10.1007/s004400050091, URL `http://dx.doi.org/10.1007/s004400050091`.

Thomas, E. G. F. (1979). Integral representations of invariant reproducing kernels, in *Proceedings, Bicentennial Congress Wiskundig Genootschap (Vrije Univ., Amsterdam, 1978), Part II*, *Math. Centre Tracts*, Vol. 101 (Math. Centrum, Amsterdam), pp. 391–404.

Thomas, E. G. F. (1994). Integral representations in conuclear cones, *J. Convex Anal.* **1**, 2, pp. 225–258 (1995).

Thorisson, H. (2000). *Coupling, stationarity, and regeneration*, Probability and its Applications (New York) (Springer-Verlag, New York), ISBN 0-387-98779-7.

Tierney, L. (1994). Markov chains for exploring posterior distributions, *Ann. Statist.* **22**, 4, pp. 1701–1762, doi:10.1214/aos/1176325750, URL `http://dx.doi.org/10.1214/aos/1176325750`, with discussion and a rejoinder by the author.

Tignol, J.-P. (2001). *Galois' theory of algebraic equations* (World Scientific Publishing Co. Inc., River Edge, NJ), ISBN 981-02-4541-6.

Todd, C. (1965). Stone-Weierstrass theorems for the strict topology, *Proc. Amer. Math. Soc.* **16**, pp. 654–659.

Tuominen, P. and Tweedie, R. L. (1994). Subgeometric rates of convergence of f-ergodic Markov chains, *Adv. in Appl. Probab.* **26**, 3, pp. 775–798.

Üstünel, A. S. and Zakai, M. (2000a). *Transformation of measure on Wiener space*, Springer Monographs in Mathematics (Springer-Verlag, Berlin), ISBN 3-540-66455-6.

Üstünel, A. S. and Zakai, M. (2000b). *Transformation of measure on Wiener space*, Springer Monographs in Mathematics (Springer-Verlag, Berlin), ISBN 3-540-66455-6.

Van Casteren, J. A. (1985). *Generators of strongly continuous semigroups*, *Research Notes in Mathematics*, Vol. 115 (Pitman), pitman Advanced Publishing Program.

van Casteren, J. A. (1992). On martingales and Feller semigroups, *Results Math.* **21**, 3-4, pp. 274–288.

Van Casteren, J. A. (2000a). Some problems in stochastic analysis and semigroup theory, in *Semigroups of operators: theory and applications (Newport Beach, CA, 1998)*, *Progr. Nonlinear Differential Equations Appl.*, Vol. 42 (Birkhäuser, Basel), pp. 43–60.

Van Casteren, J. A. (2000b). Viscosity solutions and the Hamilton-Jacobi-Bellmann equation, in *Proceedings, Third International Conference on Applied Mathematics and Engineering Sciences, Casablanca, October 23, 24 and 25, CIMASI'2000*, CIMASI (Ecole Hassania des Travaux Publics, BP:8108, Oasis Route d'El Jadida, Km 7 Casablanca - Morocco), (on CD-rom).

Van Casteren, J. A. (2001). Feynman-Kac semigroups, martingales and wave operators, *J. Korean Math. Soc.* **38**, 2, pp. 227–274.

Van Casteren, J. A. (2002). Markov processes and Feller semigroups, *Conf.*

Semin. Mat. Univ. Bari , 286, pp. 1–75 (2003).

Van Casteren, J. A. (2003). The Hamilton-Jacobi-Bellman equation and the stochastic Noether theorem, in *Evolution equations: applications to physics, industry, life sciences and economics (Levico Terme, 2000)*, *Progr. Nonlinear Differential Equations Appl.*, Vol. 55 (Birkhäuser, Basel), pp. 375–401.

Van Casteren, J. A. (2005a). Analytic operators, in *in Liber Amicorum voor Nico Temme CWI* (CWI, Amsterdam, Netherlands), pp. 8–12.

Van Casteren, J. A. (2005b). Backward stochastic differential equations and Markov processes, in F. Brackx and H. D. Schepper (eds.), *Liber Amicorum, Richard Delanghe: een veelzijdig wiskundige* (Academia Press, University of Gent, Gent), pp. 199–239.

Van Casteren, J. A. (2008). Feynman-Kac formulas, backward stochastic differential equations and Markov processes, in *Functional analysis and evolution equations* (Birkhäuser, Basel), pp. 83–111.

Van Casteren, J. A. (2009). Viscosity solutions, backward stochastic differential equations and Markov processes, *Integration: Mathematical Theory and Applications* **1**, 4, pp. 95–142, conference "The Feynman Integral and Related Topics in Mathematics and Physics". Nova Publishers.

Van Casteren, J. A. (2010). On backward stochastic differential equations in infinite dimensions, Submitted for the Proceedings of the conference "EEMMAS - Evolution Equations and Mathematical Models in the Applied Sciences", Taranto (Italy), June 29–July 03, 2009.

van Neerven, J. (1996). *The asymptotic behaviour of semigroups of linear operators*, *Operator Theory: Advances and Applications*, Vol. 88 (Birkhäuser Verlag, Basel), ISBN 3-7643-5455-0.

van Neerven, J. (2004). The Doob-Meyer decomposition theorem, Electronically: `http://en.wikipedia.org/wiki/Doob-Meyer_decomposition_theorem`, seminar Lectures Technical University Delft: Lecture 3.

Varadarajan, V. S. (1961). Measures on topological spaces, *Mat. Sb. (N.S.)* **55 (97)**, pp. 35–100.

Varadarajan, V. S. (1999). *The selected works of V. S. Varadarajan* (American Mathematical Society, Providence, RI), ISBN 0-8218-1068-5.

Varadhan, S. R. S. (2007). *Stochastic processes*, *Courant Lecture Notes in Mathematics*, Vol. 16 (Courant Institute of Mathematical Sciences, New York), ISBN 978-0-8218-4085-6.

Varopoulos, N. T. (1985). Long range estimates for Markov chains, *Bull. Sci. Math. (2)* **109**, 3, pp. 225–252.

Varopoulos, N. T., Saloff-Coste, L. and Coulhon, T. (1992). *Analysis and geometry on groups*, *Cambridge Tracts in Mathematics*, Vol. 100 (Cambridge University Press, Cambridge), ISBN 0-521-35382-3.

Vestrup, E. M. (2003). *The theory of measures and integration*, Wiley Series in Probability and Statistics (Wiley-Interscience [John Wiley & Sons], Hoboken, NJ), ISBN 0-471-24977-7.

Villani, C. (2003). *Topics in optimal transportation*, *Graduate Studies in Mathematics*, Vol. 58 (American Mathematical Society, Providence, RI), ISBN 0-8218-3312-X.

Villani, C. (2009). *Optimal transport*, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Vol. 338 (Springer-Verlag, Berlin), ISBN 978-3-540-71049-3, old and new.

Wang, F.-Y. (2005). *Functional Inequalities Markov Semigroups and Spectral Theory*, *Mathematics Monograph Series*, Vol. 4 (Elsevier Science, Amsterdam, London, New York), ISBN 978-0-08-044942-5, part of: The Science Series of the Contemporary Elite Youth.

Watanabe, S. and Yamada, T. (1971). On the uniqueness of solutions of stochastic differential equations. II, *J. Math. Kyoto Univ.* **11**, pp. 553–563.

Wells, J. (1965). Bounded continuous vector-valued functions on a locally compact space, *Michigan Math. J.* **12**, pp. 119–126.

Wheeler, R. F. (1983). A survey of Baire measures and strict topologies, *Exposition. Math.* **1**, 2, pp. 97–190.

Widder, D. V. (1946). *The Laplace Transform*, Princeton Mathematical Series, v. 6 (Princeton University Press, Princeton, N. J.).

Williams, D. (1991). *Probability with martingales*, Cambridge Mathematical Textbooks (Cambridge University Press, Cambridge), ISBN 0-521-40455-X; 0-521-40605-6.

Wiweger, A. (1961). Linear spaces with mixed topology, *Studia Math.* **20**, pp. 47–68.

Wu, L. (2000). Uniformly integrable operators and large deviations for Markov processes, *J. Funct. Anal.* **172**, 2, pp. 301–376.

Yong, J. and Zhou, X. Y. (1999). *Stochastic controls*, *Applications of Mathematics (New York)*, Vol. 43 (Springer-Verlag, New York), ISBN 0-387-98723-1, hamiltonian systems and HJB equations.

Yosida, K. (1995). *Functional analysis*, Classics in Mathematics (Springer-Verlag, Berlin), ISBN 3-540-58654-7, reprint of the sixth (1980) edition.

Zaanen, A. C. (1997). *Introduction to operator theory in Riesz spaces* (Springer-Verlag, Berlin), ISBN 3-540-61989-5.

Zaharopol, R. (2005). *Invariant probabilities of Markov-Feller operators and their supports*, Frontiers in Mathematics (Birkhäuser Verlag, Basel), ISBN 3-7643-7134-X; 978-3-7643-7134-0.

Zaharopol, R. and Zbaganu, G. (1999). Dobrushin coefficients of ergodicity and asymptotically stable L^1-contractions, *J. Theoret. Probab.* **12**, 4, pp. 885–902.

Zambrini, J.-C. (1998a). A special time-dependent quantum invariant and a general theorem on quantum symmetries, in *Stochastic analysis and mathematical physics (Viña del Mar, 1996)* (World Sci. Publ., River Edge, NJ), pp. 197–210.

Zambrini, J.-C. (1998b). A special time-dependent quantum invariant and a general theorem on quantum symmetries, in *Stochastic analysis and mathematical physics (Viña del Mar, 1996)* (World Sci. Publ., River Edge, NJ), pp. 197–210.

Zhang, X. (2010). Stochastic Volterra equations in Banach spaces and stochastic partial differential equation, *J. Funct. Anal.* **258**, 4, pp. 1361–1425, doi:

10.1016/j.jfa.2009.11.006, URL http://dx.doi.org/10.1016/j.jfa.2009.11.006.

Zolotarev, V. M. (1983). Probability metrics, *Teor. Veroyatnost. i Primenen.* **28**, 2, pp. 264–287.

Index

Series on Concrete and Applicable Mathematics – Vol. 12

MARKOV PROCESSES, FELLER SEMIGROUPS AND EVOLUTION EQUATIONS

The book provides a systemic treatment of time-dependent strong Markov processes with values in a Polish space. It describes their generators and the link with stochastic differential equations in infinite dimensions. In a unifying way, where the square gradient operator is employed, new results for backward stochastic differential equations and long-time behavior are discussed in depth. The book also establishes a link between propagators or evolution families with the Feller property and time-inhomogeneous Markov processes. This mathematical material finds its applications in several branches of the scientific world, among which mathematical physics, hedging models in financial mathematics, and population or other models in which the Markov property plays a role.

World Scientific
www.worldscientific.com
7871 hc ISSN: 1793-1142

ISBN-13 978-981-4322-18-8
ISBN-10 981-4322-18-0
9 789814 322188