University of Alberta

Characterizations of reversibility for certain classes of finite and infinite dimensional diffusions

by

Omar Daniel Rivasplata ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematics

Department of Mathematical and Statistical Sciences

Edmonton, Alberta

Fall 2005

Your file Votre référence
ISBN: 0-494-09271-8
Our file Notre référence
ISBN: 0-494-09271-8

To my parents, Julia Rosa and Segundo German,

to my wife, Mariel,

and to our children, Gabriel Arturo and Sofia Isabel.

Abstract

In this thesis we study the problem of reversibility for certain classes of diffusion processes in finite and infinite dimensions. In the finite dimensional case we look at the generator of Brownian motion with drift, and we present two characterizations of reversibility: the criterion of Kolmogorov which establishes that reversibility is possible if and only if the drift is of gradient form, and a criterion proving that reversibility of a measure is equivalent to quasi-invariance under the group of all translations with a cocycle given in terms of the drift coefficients. Later we use the ideas from the second characterization in finite dimensions to explore the property of reversibility for an Ornstein-Uhlenbeck process with values in an infinite dimensional Hilbert space.

Acknowledgements

Many people have contributed to making my experience of the last three years interesting and enjoyable, either in the mathematical sense or in the more personal side. Both groups of people have their share in the successful conclusion of my program.

First of all, I sincerely thank my supervisor Dr. Byron Schmuland for his guidance and support. His generous funding made it possible to concentrate on my work without worries about financial survival issues. On the mathematical side, he patiently allowed me to explore avenues of possible research before choosing one to pursue, and all along he was there to provide valuable advice regarding the job of doing and writing mathematics.

I would like to thank Prof. Kenyi Handa. Although we have not met personally, he very kindly responded to my questions through the e-mail, and this communication helped clarify some ideas which were obscure to me at the time. He also suggested pertinent readings, thanks very much for this.

I owe words of gratitude to Jan Rychtar, friend and colleague, for many pleasant conversations about things mathematical and of other nature. In particular, I thank Jan for his explanations of some properties of function spaces, for suggesting the idea behind the proof of Proposition 2.1, and for many hints about how to write mathematical work in nice and convincing ways.

I would like to thank Dona Guelzow, graduate secretary of the Department of Mathematical and Statistical Sciences, for her limitless patience and always amazing willingness to help.

I also want to thank all my friends from Edmonton and from other warmer places in the world, their company and support has always made good times feel even better and difficult times feel a little less difficult.

Last, but by far not least, I want to express my heartfelt gratitude to my wife, Mariel, for her unconditional support all this time.

Table of Contents

Introduction

Intuitively speaking, given a stochastic process $X = (X_t)_{0 \leq t \leq T}$ one can define its time-reversal $\widetilde{X} = (\widetilde{X}_t)_{0 \leq t \leq T}$ as a new stochastic process whose value at time t is $\widetilde{X}_t := X_{T-t}$. This time-reversed process $\widetilde{X}$ runs the trajectories of X in the reverse direction, and it need not have the same probabilistic properties of the original process. In this setting, reversibility occurs when the process X and its time reversal $\widetilde{X}$ share the same finite-dimensional distributions, or in other words, when they do have the same probabilistic properties.

In this thesis we study characterizations of reversibility for certain classes of diffusion processes in finite and infinite dimensions. Our approach is to define the meaning of reversibility for a given diffusion in terms of its formal generator. This approach is also used in the papers [24], [25] of K.Handa, where reversibility of two particular kinds of diffusions is characterized in terms of a certain property called quasi-invariance. Motivated by these results, our goal is to characterize reversibility of an Ornstein-Uhlenbeck process with values in an infinite dimensional Hilbert space, a process of the kind described in the paper of A.Chojnowska-Michalik and B.Goldys [15]. Before dealing with infinite dimensions, we work out in detail two characterizations of reversibility for a class of diffusions on Euclidean space. This work in finite dimensions has mathematical interest in its own right, and also helps to gain insight into the infinite dimensional case. This explains the title of this thesis.

Let us now briefly describe the contents of this thesis. In Chapter 1 we provide a quick and rather informal introduction to Markov processes and, in particular, to diffusion processes. Here, we start from the basic definitions concerning stochastic processes and their finite dimensional distributions, and we present some examples of particularly important classes of processes, including Brownian motion. After a brief section intended to introduce the notions of semigroups of operators and their generators, we move in the last section to Markov processes and especially to Feller-Dynkin diffusions on Euclidean space. To close this chapter we quote a result establishing that the generator of such processes, when restricted to a certain space of functions, is a second-order differential operator of elliptic type. It is our intention that this introductory chapter provide some context for the ensuing discussion.

The next two chapters concentrate on reversibility for "Brownian motion with drift", a class of finite dimensional diffusions corresponding to a specific type of second-order differential operator. In Chapter 2 we focus on the first characterization of reversibility, the criterion of Kolmogorov. We introduce the notions of invariant and reversible measures for a given differential operator of second order, we discuss the literature concerning existence and regularity of invariant measures, and we prove Kolmogorov's criterion which establishes that reversibility is possible if and only if the drift is of gradient form. As a consequence we obtain that in one dimension reversibility is always possible.

Chapter 3 is devoted to the second characterization of reversibility in finite dimensions. Here we study quasi-invariance properties of measures under certain transformation groups, as well as the associated cocycles. The main result of this chapter, Theorem 3.1, establishes that reversibility of a measure is equivalent to quasi-invariance under all translations with cocycle given in terms of the drift coefficients. Closing this chapter we present a new result, Theorem 3.2, establishing that the cocycle identity implies that the drift is a conservative vector field.

The last chapter of the thesis, Chapter 4, explores the problem of reversibility for an Ornstein-Uhlenbeck process with values in an infinite dimensional Hilbert space. We introduce the formal generator of this process as an operator acting on cylindrical functions, and we define the meaning of invariant and reversible measures in this context. Here we rely on the ideas from the finite dimensional case to conjecture a characterization of reversibility in terms of quasi-invariance. Although we do not settle this conjecture completely, we make significant progress towards the proof. The solution of this problem, when finished, will constitute an original contribution to knowledge.

Throughout the thesis we provide indications as to which results were already known prior to our work and which are new. In some cases we provide our own proofs of previously known results.

Chapter 1

Basics about diffusion processes

1.1 Stochastic processes

Any probabilistic discussion starts by specifying a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Consider a measurable space $(E, \mathcal{B})$; a stochastic process X with index set I and state space E is a function $X : I \times \Omega \to E$ such that for each $t \in I$, the mapping $X(t, \cdot) : \Omega \to E$ is an E-valued random variable. The use of t for a generic element of the index set is not incidental since we are primarily interested in processes indexed by time. Consequently we will take $I = [0, \infty)$ for continuous time processes, or $I = \{0, 1, 2, \ldots\}$ for discrete time processes.

It is customary to use the notation X_t for the random variable $X(t, \cdot)$ described in the previous paragraph. This notation suggests that we may think of a stochastic process X as being a family of random variables indexed by time: $X = (X_t)_{t \geq 0}$. Stochastic processes provide the mathematical model for systems that evolve in time according to some random mechanism.

Given a stochastic process X, for any fixed $\omega \in \Omega$, there corresponds a function $t \mapsto X_t(\omega)$, called a trajectory of the process. In the literature trajectories are also called sample paths. The process $X = (X_t)_{t \geq 0}$ is called continuous (right continuous, left continuous) when it has continuous (resp. right continuous, left continuous) trajectories.

A collection $\{\mathcal{F}_t\}_{t \geq 0}$ of sub-σ-algebras of $\mathcal{F}$ is called a filtration if $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \leq t$. In particular, for a process X we define the natural filtration $\{\mathcal{F}_t^X\}_{t \geq 0}$ by setting $\mathcal{F}_t^X$ to be the smallest σ-algebra with respect to which all the functions X_s, $s \leq t$ are measurable. In the usual measure-theoretic notation, $\mathcal{F}_t^X := \sigma(X_s : s \leq t)$. Intuitively we think of $\mathcal{F}_t^X$ as the σ-algebra containing information about the history of the process X up to time t. When the probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is endowed with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$, we call the quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathsf{P})$ a filtered probability space.

A process X is said to be adapted to the filtration $\{\mathcal{F}_t\}$ when for each $t \geq 0$, the random variable X_t is $\mathcal{F}_t$-measurable. Note that X is $\{\mathcal{F}_t\}$-adapted if and only if $\mathcal{F}_t^X \subset \mathcal{F}_t$ for each $t \geq 0$.

Let the stochastic process $X = (X_t)_{t \geq 0}$ be given. For any $m \geq 1$ and any list of distinct numbers $t_1, t_2, \ldots, t_m \in [0, \infty)$ we define $\mu_{t_1, t_2, \ldots, t_m}$ to be the probability measure on $\mathcal{B} \otimes \cdots \otimes \mathcal{B}$ (m factors) induced by the mapping $(X_{t_1}, \ldots, X_{t_m}) \mapsto E^m$, that is,

$$\mu_{t_1, t_2, \ldots, t_m}(\Gamma) = \mathsf{P}[(X_{t_1}, \ldots, X_{t_m}) \in \Gamma] \quad \text{for } \Gamma \in \mathcal{B} \otimes \cdots \otimes \mathcal{B}.$$

The probability measures $\{\mu_{t_1, t_2, \ldots, t_m} : m \geq 1\}$ thus defined are called the finite dimensional distributions of X. In particular, the probability measure $\mu_0(\cdot) := \mathsf{P}(X_0 \in \cdot)$ is called the initial distribution of the process.

It is a fact of life that all the probabilistic properties of a given stochastic process X are determined by its finite dimensional distributions. To establish this fact rigorously one has to construct another process called the canonical version of X; the law of X is then defined as the law of its canonical version. The construction of the canonical version, as a process defined on a product space, makes it evident that the law of the process is determined by its finite dimensional distributions, see [35, Section I.3] for details on this.

On the other hand, if we are given a family $\{\mu_{t_1, t_2, \ldots, t_m} : m \geq 1\}$ of probability measures, then a question arises whether there is a process X having this family as its finite dimensional distributions. This question is answered by the following theorem due to Kolmogorov, its proof can be found in [40, p. 8].

Theorem 1.1. *Let E be a complete separable metric space. Assume that the family of probability measures $\{\mu_{t_1, t_2, \ldots, t_m} : m \geq 1\}$ satisfies the following consistency conditions:*

(i) $\mu_{t_1, t_2, \ldots, t_m}(B_1, \ldots, B_m) = \mu_{t_{i_1}, t_{i_2}, \ldots, t_{i_m}}(B_{i_1}, \ldots, B_{i_m})$ *for any permutation* $i_1, \ldots, i_m$ *of the numbers* $1, \ldots, m$*;*

(ii) $\mu_{t_1, t_2, \ldots, t_{m-1}, t_m}(B_1, \ldots, B_{m-1}, E) = \mu_{t_1, t_2, \ldots, t_{m-1}}(B_1, \ldots, B_{m-1})$.

Then there exists a stochastic process X whose finite dimensional distributions are the given family of measures.

Example 1.1 (Processes with independent increments). Assume that E is a linear space. A process $X = (X_t)_{t \geq 0}$ is said to be a process with independent increments if for any $n \geq 1$ and any choice of $t_0 < t_1 < \cdots < t_n$ the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.

Consider for instance $E = \mathbb{R}^d$, and let $\langle \cdot, \cdot \rangle$ denote the usual inner product in this space. The characteristic function $\varphi_{t_0,\dots,t_n}(z_0, \dots, z_n)$ of the random element $(X_{t_0}, \dots, X_{t_n})$ is given by

$$\varphi_{t_0,\dots,t_n}(z_0, \dots, z_n) = \mathsf{E} \exp \left\{ i \sum_{k=0}^{n} \langle z_k, X_{t_k} \rangle \right\}, \qquad \text{for} \quad z_0, \dots, z_n \in \mathbb{R}^d.$$

Note that we can write

$$\varphi_{t_0,\dots,t_n}(z_0, \dots, z_n) =$$
$$= \mathsf{E} \exp \left\{ i \langle \sum_{k=0}^{n} z_k, X_{t_0} \rangle + i \langle \sum_{k=1}^{n} z_k, X_{t_1} - X_{t_0} \rangle + \dots + i \langle z_n, X_{t_n} - X_{t_{n-1}} \rangle \right\}$$

Let us write $\psi_{t_k,t_l}(z) = \mathsf{E} \exp\{i \langle z, X_{t_k} - X_{t_l} \rangle\}$. Then using the independence of increments it follows that we can write

$$\varphi_{t_0,\dots,t_n}(z_0, \dots, z_n) = \varphi_{t_0}(z_0 + \dots + z_n) \prod_{k=1}^{n} \psi_{t_{k-1},t_k}(z_k + \dots + z_n).$$

Since the characteristic function of the random element $(X_{t_0}, \dots, X_{t_n})$ uniquely determines its distribution, we learn from this example that, in the case of a process with independent increments, to determine the finite dimensional distributions it suffices to know the one dimensional distributions of the process and the two-dimensional distributions of its increments. $\diamondsuit$

Example 1.2 (Gaussian processes). The real-valued process $X = (X_t)_{t \geq 0}$ is called a Gaussian process if the random vector $(X_{t_1}, \dots, X_{t_n})$ has the Gaussian distribution, for any choice of $t_1, \dots, t_n$. We call $a(t) := \mathsf{E}(X_t)$ the expectation function, and $c(t, s) := \mathsf{E}(X_t X_s) - a(t)a(s)$ the covariance function. The characteristic function of the random vector $(X_{t_1}, \dots, X_{t_n})$ is by definition

$$\varphi_{t_1,\dots,t_n}(z_1, \dots, z_n) = \mathsf{E} \exp \left\{ i \sum_{j=1}^{n} z_j X_{t_j} \right\} \qquad \text{for } (z_1, \dots, z_n) \in \mathbb{R}^n.$$

For fixed $(z_1, \dots, z_n) \in \mathbb{R}^n$ let us define the random variable $Y := \sum_{j=1}^{n} z_j X_{t_j}$. This is a Gaussian random variable (see for instance [20, Theorem 9.5.13] or [39, Theorem II.13.1]), and a few computations show that

$$\mathsf{E}(Y) = \sum_{j=1}^{n} z_j a(t_j), \qquad \mathrm{Var}(Y) = \sum_{j,k=1}^{n} z_j z_k c(t_j, t_k).$$

It is well known that if Z is a Gaussian random variable with mean m and variance σ^2 then its characteristic function is given as

$$\varphi_Z(t) = \exp \left\{ itm - \frac{1}{2} t^2 \sigma^2 \right\}, \quad \text{for } t \in \mathbb{R}.$$

Using this and the obvious fact that $\varphi_{t_1,\ldots,t_n}(z_1,\ldots,z_n) = \varphi_Y(1)$ we can write the characteristic function of the random vector $(X_{t_1},\ldots,X_{t_n})$ as follows

$$\varphi_{t_1,\ldots,t_n}(z_1,\ldots,z_n) = \exp\left\{ i\sum_{j=1}^{n} z_j a(t_j) - \frac{1}{2}\sum_{j,k=1}^{n} z_j z_k c(t_j,t_k) \right\}.$$

Thus we learn from this example that for a Gaussian process, the finite dimensional distributions are completely determined by its expectation function and its covariance function. ◇

To close this section we would like to present a very important example of stochastic process.

Example 1.3 (Brownian motion). We say that the process $W = (W_t)_{t\geq 0}$ is a one dimensional Brownian motion with respect to the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ if it is adapted to the given filtration and has the following properties:

(1) W starts at zero, *i.e.* $W_0(\omega) = 0$ for all ω;

(2) W has continuous trajectories;

(3) for $s < t$, the random variable $W_t - W_s$ has distribution $\mathcal{N}(0, t-s)$ and is independent of $\mathcal{F}_s$.

When we take $\mathcal{F}_t = \mathcal{F}_t^W$ to be the natural filtration generated by W, we simply say that W is a Brownian motion. The essential conditions here are (2) and (3); for if $W = (W_t)_{t\geq 0}$ is a Brownian motion (starting at zero), we can construct the process $(x + W_t)_{t\geq 0}$ which starts at the point x. Finally, let us remark that a d-dimensional Brownian motion is a d-dimensional process whose components are independent one-dimensional Brownian motions. ◇

We will take for granted the existence of Brownian motion, a proof of this fact can be found in [36, Theorem I.6.1]. The importance of the study of Brownian motion comes from the facts that it serves as a building block for other kinds of processes, and it can be taken as the first example of virtually every interesting class of processes. In particular, it is not difficult to see that Brownian motion is a Gaussian process with independent increments. The independence of the increments comes free of charge from the definition of Brownian motion, while its being a Gaussian process follows from the independence of increments and the fact that the individual variables are Gaussian, see [35, p. 17]. More details about Brownian motion and its many properties can be found in [36, Chapter I].

1.2 Semigroups of operators

This short section is an interlude on the theory of semigroups of bounded linear operators, and it has the sole purpose of presenting some definitions that may prove handy for what follows. Therefore no attempt will be made to make it comprehensive, let alone self contained. Very good references for this theory are [21] and [33], see also [22, Chapter 1].

Definition 1.1. Let B be a Banach space. A one parameter family $(T_t)_{t\geq 0}$ of bounded linear operators from B into B is called a semigroup if it satisfies the following equations:

(i) $T_0 = I$, the identity operator on B;

(ii) $T_{t+s} = T_t T_s$ for every $t, s \geq 0$ (the semigroup property).

We may think of semigroups of operators as being the analogs of exponential functions. It is well known that the exponential has the properties $e^{t+s} = e^t e^s$ and $e^0 = 1$, which resemble the equations defining a semigroup. Consider the real function $f(t) = e^{\alpha t}$, it is clear that this function is completely determined by the parameter α, and this parameter in turn can be determined by the behavior of the function near zero: $f'(0) = \alpha$. In this sense we can say that the parameter α "generates" the function $e^{\alpha t}$. These ideas lead to the following definition.

Definition 1.2. The infinitesimal generator of the semigroup $(T_t)_{t\geq 0}$ is the linear operator $G : \mathcal{D}(G) \subset B \to B$ defined by

$$\mathcal{D}(G) = \left\{ x \in B \, : \, \lim_{t\downarrow 0} \frac{T_t x - x}{t} \quad \text{exists} \right\} \tag{1.1}$$

$$Gx = \lim_{t\downarrow 0} \frac{T_t x - x}{t} = \left. \frac{d^+ T_t x}{dt} \right|_{t=0} \qquad \text{for } x \in \mathcal{D}(G) \tag{1.2}$$

Here, $\mathcal{D}(G)$ designates the domain of the operator G.

It should be pointed out that the domain of the generator of a semigroup is usually a proper subspace of the Banach space B, and it may be a hard problem to describe it. For our purposes the class of semigroups as presented above is too broad, and we will fix out attention in semigroups having the additional property of strong continuity.

Definition 1.3. A semigroup $(T_t)_{t\geq 0}$ of bounded linear operators on B is called strongly continuous if

$$\lim_{t\downarrow 0} T_t x = x \qquad \text{for every } x \in B.$$

For any strongly continuous semigroup $(T_t)_{t\geq 0}$ there exist constants $\eta \geq 0$ and $M \geq 1$ such that $\|T_t\| \leq Me^{\eta t}$ for $t \geq 0$ (see [33, Theorem 1.2.2]). If it turns out that $\eta = 0$, then reasonably enough we say that the semigroup is uniformly bounded. If in addition we have $M = 1$, then each operator T_t is a contraction; in this case we say that $(T_t)_{t\geq 0}$ is a strongly continuous semigroup of contractions.

The generator of a strongly continuous semigroup need not be a bounded operator. However, if G is the generator of a strongly continuous semigroup, then G is a closed operator, and is domain $\mathcal{D}(G)$ is dense, *i.e.* $\overline{\mathcal{D}(G)} = B$, (see [33, Corollary 1.2.5]). It can be proved that every bounded operator A generates a strongly continuous semigroup via the formula $T_t = e^{tA}$. On the other hand, if we are given an unbounded operator A on B, it is of particular interest for us to have conditions so that A generates a strongly continuous semigroup of contractions. This is done by the theorem of Hille-Yosida, the reader is referred to [33] for details on this.

1.3 Markov processes and their semigroups

Intuitively speaking, a Markov process is a stochastic process with the property that, conditional on the knowledge of its present state, its future values are independent of its past history. The purpose of this section is to present a mathematical framework for Markov processes in which the above intuitive description makes sense.

From now on we will take E to be a complete separable metric space, and we also assume E to be locally compact (see [20] for these definitions). For measurability purposes we consider E with its Borel σ-algebra $\mathcal{B}(E)$. We will use the following conventions:

$M(E)$ is the collection of all real-valued measurable functions on E

$M_b(E)$ denotes the Banach space of bounded measurable functions with the norm $\|f\| = \sup_{x\in E}|f(x)|$,

$C(E)$ is the space of all continuous functions on E,

$C_b(E)$ is the subspace of bounded continuous functions,

$C_0(E)$ is the subspace of (bounded) continuous functions on E which vanish at infinity,

$C_c(E)$ is the subspace of continuous functions with compact support.

Consider a stochastic process $X = (X_t)_{t\geq 0}$ with state space $(E, \mathcal{B}(E))$. We have noted that the history of the process X up to time s is modelled by the σ-algebra $\mathcal{F}_s^X := \sigma(X_u,\, u \leq s)$. If X is a Markov process (in the intuitive sense described at the beginning of this section), then for any $A \in \mathcal{B}(E)$ and $s < t$ the conditional probability

$$\mathsf{P}\left(X_t \in A \,|\, \mathcal{F}_s^X\right)$$

should be a function of the present state X_s, so it may be written in the form $g(X_s)$ for some measurable function $g : E \longrightarrow [0,1]$. But this function also depends on t and A, so it would be better to write $g_t(X_s, A)$ to indicate this dependence. Let us concentrate on a couple of the properties that this function ought to have. Clearly, as a function of A, it should be a probability measure describing the chance that the process be found in A at time t, knowing where it was at time s. On the other hand, for each A fixed, it is clear that the mapping $x \mapsto g_t(x, A)$ should be measurable. These ideas motivate the following definition.

Definition 1.4. A kernel on E is a map $\pi : E \times \mathcal{B}(E) \longrightarrow [0, +\infty]$ having the following properties

i) for every $x \in E$, the map $A \mapsto \pi(x, A)$ is a measure on $(E, \mathcal{B}(E))$;

ii) for every $A \in \mathcal{B}(E)$, the map $x \mapsto \pi(x, A)$ is measurable.

A kernel π is called a transition probability on E when $\pi(x, E) = 1$ for every $x \in E$.

A transition probability π provides the mathematical description of a random motion in E, as we exemplify now in a discrete-time setting. Assume that at time zero the motion starts at $x \in E$. Then at time 1 the position x_1 is chosen at random according to the probability $\pi(x, \cdot)$, at time 2 the position x_2 is chosen according to $\pi(x_1, \cdot)$, and so on and so forth.

Definition 1.5. A (homogeneous) transition function on $(E, \mathcal{B}(E))$ is a family P_t, $t \geq 0$ of transition probabilities on $(E, \mathcal{B}(E))$ such that for every pair of real numbers $s, t \geq 0$ we have

$$P_{t+s}(x, A) = \int P_t(x, dy) P_s(y, A)$$

for every $x \in E$ and $A \in \mathcal{B}(E)$. This relation is called the Chapman-Kolmogorov equation.

Assume that P_t is a transition function. From the definition above it follows that each P_t is a transition probability, so that $P_t(x, E) = 1$ for every x. In the literature it is common to refer to this kind of transition function as being markovian. However, there are cases where we may need to require that $P_t(x, E) < 1$ for some x's and t's, in this case we say that the transition function is submarkovian. If we think of the transition function as describing the random motion of a particle, the submarkovian case includes the possibility that the particle disappears or dies in a finite time.

It is possible to turn the submarkovian case into a markovian case by means of the following trick. We define a cemetery state Δ, and we adjoin this state to E as an isolated point, thus producing an extended state space $E_\Delta = E \cup \{\Delta\}$. In this case we also extend $\mathcal{B}(E)$ to $\mathcal{B}(E_\Delta) := \sigma(\mathcal{B}(E), \{\Delta\})$, the smallest σ-algebra on E_Δ containing $\mathcal{B}(E)$ and $\{\Delta\}$. The state Δ is considered to be absorbing, in the sense that once the particle enters this state, it stays there forever. The next definition was borrowed from [38].

Definition 1.6. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$ be a filtered probability space. We say that $X = ((X_t)_{t\geq 0}, (\mathsf{P}_z)_{z\in E_\Delta})$ is a Markov process with respect to $(\mathcal{F}_t)_{t\geq 0}$, with state space E and lifetime ζ if

1. For every $t \geq 0$, the map $X_t : \Omega \to E_\Delta$ is an E_Δ-valued random variable.
2. For each $t \geq 0$, X_t is $\mathcal{F}_t$-measurable.
3. For every $\omega \in \Omega$, $X_t(\omega) \in E$ for $t < \zeta(\omega)$, and $X_t(\omega) = \Delta$ for $t \geq \zeta(\omega)$.
4. $(\mathsf{P}_z)_{z\in E_\Delta}$ is a family of probability measures on $(\Omega, \mathcal{F})$ such that the map $z \mapsto \mathsf{P}_z(A)$ is $\mathcal{B}(E_\Delta)$-measurable for each $A \in \mathcal{B}(E_\Delta)$.
5. $\mathsf{P}_z[X_0 = z] = 1$ for all $z \in E_\Delta$.
6. (Markov property) For every $s, t \geq 0$ and every $z \in E_\Delta$,
$$\mathsf{P}_z[X_{t+s} \in A \mid \mathcal{F}_t] = \mathsf{P}_{X_t}[X_s \in A], \quad \mathsf{P}_z\text{-a.e.}, \qquad \text{for all } A \in \mathcal{B}(E_\Delta).$$

We will omit the specification of the filtration when we want to regard X as a Markov process with respect to its natural filtration $\{\mathcal{F}_t^X\}$. In this case we will simply say that X is a Markov processes.

Every Markov process X has an associated transition function P_t, which is defined by the expression

$$P_t(x, \Gamma) := \mathsf{P}_x[X_t \in \Gamma] = \mathsf{E}_x[\mathbf{1}_\Gamma(X_t)]. \tag{1.3}$$

In fact, there is a one-to-one correspondence between Markov processes and transition functions, as we now establish.

Recall that the initial distribution of X is the probability measure μ_0 on $(E, \mathcal{B}(E))$ defined by $\mu_0(B) := \mathsf{P}(X_0 \in B)$ for $B \in \mathcal{B}(E)$. Assuming that the initial distribution μ_0 of the Markov process X is known, its finite-dimensional distributions can be written as

$$\mathsf{P}(X_0 \in \Gamma_0, X_{t_1} \in \Gamma_1, \ldots, X_{t_n} \in \Gamma_n) = \tag{1.4}$$
$$\int_{\Gamma_0} \cdots \int_{\Gamma_n} P_{t_n - t_{n-1}}(y_{n-1}, dy_n) P_{t_{n-1}-t_{n-2}}(y_{n-2}, dy_{n-1}) \cdots P_{t_1}(y_0, dy_1)\mu_0(dy_0).$$

Thus it is clear that the finite-dimensional distributions of X are completely determined if we know the transition function corresponding to this process.

On the other hand, it is established in [22, Theorem 1.1.1] that for any time homogeneous transition function $P_t(x, \Gamma)$ and probability measure μ_0 on E there is a Markov process X with values in E whose finite-dimensional distributions are uniquely determined by (1.4).

Thus we see that, as far as probabilistic properties are concerned, the study of a Markov process can be reduced to the study of its corresponding transition function, provided that the initial distribution is known.

Example 1.4 (Brownian motion). The Brownian transition density is the function $p_t(x, y)$ defined as follows:

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left\{ -\frac{(y-x)^2}{2t} \right\}, \quad t > 0.$$

Using this density we can explicitly determine the transition function for Brownian motion:

$$P_t(x, \Gamma) = \begin{cases} \int_\Gamma p_t(x, y)dy, & \text{for } t > 0, \\ \delta_x(\Gamma), & \text{for } t = 0. \end{cases} \tag{1.5}$$

Note that this transition function is defined in terms of a Gaussian density, something to be expected since Brownian motion is a Gaussian process. ◇

It should be pointed out that usually transition functions are not given by explicit formulas, the previous example being the most notable exception. Consequently, it is not advisable to rely on explicit definitions of transition functions for constructing Markov processes. To go around this difficulty we use the one-to-one correspondence existing between transition functions and a certain kind of semigroups of operators, as we next explain. First we need a couple of definitions.

Definition 1.7 (Markov semigroup). Let B be a Banach space of real-valued functions on E. A semigroup $(T_t)_{t\geq 0}$ of bounded linear operators on B is called a Markov semigroup if $0 \leq f \leq 1$ implies $0 \leq T_t f \leq 1$ for $f \in B$, and also $T_t 1 = 1$ for all $t \geq 0$.

There is a particularly important class of Markov semigroups, the Feller-Dynkin semigroups. These are defined next.

Definition 1.8 (Feller-Dynkin semigroup). A Feller-Dynkin semigroup is a strongly continuous Markov semigroup $(T_t)_{t\geq 0}$ of bounded linear operators on the space $C_0(E)$.

Given a Markov process X with corresponding transition function P_t, the expression

$$T_t f(x) := \int f(y) P_t(x, dy) \tag{1.6}$$

defines a contraction semigroup on $M_b(E)$, as can be seen using the Chapman-Kolmogorov property. In fact, we can see without difficulty that equation (1.6) defines a Markov semigroup on $M_b(E)$. Thus it is clear that if we know the transition function $P_t(x, \Gamma)$ of the Markov process X, then its associated Markov semigroup (T_t) is determined.

Conversely, if we are given a Markov semigroup (T_t), we can recover the transition function applying formally the operators T_t on indicator functions of Borel sets:

$$P_t(x, \Gamma) = T_t \mathbf{1}_\Gamma(x).$$

This formal construction works well at least when (T_t) is a Feller-Dynkin semigroup. In fact, the importance of this kind of semigroups comes from the fact that any Feller-Dynkin semigroup has a corresponding transition function $P_t(x, \Gamma)$ on $(E, \mathcal{B}(E))$ such that equation (1.6) holds for all functions $f \in M_b(E)$; a proof of this can be found in [35, Proposition III.2.2].

Thus we learn that, in order to study the probabilistic properties of a Markov process, it suffices to concentrate on the study of its associated semigroup. This is a great gain since at this stage we can have at our disposal all the machinery developed for semigroups of operators on a Banach space. In particular, we can concentrate our efforts in studying the generator of the semigroup.

Example 1.5 (Brownian motion). Using the transition function presented in Example 1.4 we can construct explicitly the semigroup (T_t) corresponding to Brownian motion: for any bounded Borel function $f : \mathbb{R} \to \mathbb{R}$ and $t \geq 0$ we have

$$T_t f(x) := \begin{cases} \int_{-\infty}^{\infty} f(y) p_t(x, y) dy, & \text{for } t > 0, \\ f(x) & \text{for } t = 0. \end{cases}$$

Here, the semigroup property for (T_t) follows from the Chapman-Kolmogorov equations. Let us consider the generator G of this semigroup, *i.e.* the operator defined as

$$Gf = \lim_{t \downarrow 0} \frac{T_t f - f}{t}$$

for suitable functions f. Note that for $f \in C_c^2(\mathbb{R})$ we have

$$\begin{aligned} Gf(x) &= \lim_{t \downarrow 0} \frac{1}{t}(T_t f - f)(x) \\ &= \lim_{t \downarrow 0} \int_{-\infty}^{\infty} \frac{f(x + y\sqrt{t}) - f(x)}{t} \exp(-\frac{y^2}{2}) \frac{dy}{\sqrt{2\pi}} \\ &= \lim_{t \downarrow 0} \int_{-\infty}^{\infty} \frac{1}{t} \left\{ y\sqrt{t} f'(x) + \frac{1}{2} y^2 t f''(x + \theta y \sqrt{t}) \right\} \exp(-\frac{y^2}{2}) \frac{dy}{\sqrt{2\pi}}, \end{aligned}$$

where $\theta \in (0,1)$ depends on $y\sqrt{t}$. Calculating this limit we get

$$Gf(x) = \lim_{t \downarrow 0} \frac{1}{t}(T_t f - f)(x) = \frac{1}{2} f''(x)$$

Thus we see that the space $C_c^2(\mathbb{R})$ is contained in the domain of the generator of Brownian motion, and $Gf = \frac{1}{2} f''$ for $f \in C_c^2(\mathbb{R})$. ◇

Recall that for a twice-differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ we define the Laplacian of f, denoted Δf, by means of the expression

$$\Delta f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

It is established in [35, Proposition VII.1.10] that for one-dimensional Brownian motion the domain of the infinitesimal generator is $\mathcal{D}(G) = C_0^2(\mathbb{R})$ and we actually have $G = \frac{1}{2}\Delta$ on this space, see also [36, p. 243]. For n-dimensional Brownian motion with $n \geq 2$ this is no longer the case, but the domain $\mathcal{D}(G)$ of the generator is larger and contains $C_0^2(\mathbb{R})$ as a proper subspace. However, it is still true that the restriction of G to $C_0^2(\mathbb{R})$ coincides with the operator $\frac{1}{2}\Delta$ (see [35, Proposition VII.1.11] and also [36, p. 257]).

Let $E = \mathbb{R}^n$, and let $P_t(x, dy)$ denote the transition function of n-dimensional Brownian motion. For $f \in C_0(\mathbb{R}^n)$ we set

$$T_t f(x) := \int f(y) P_t(x, dy). \tag{1.7}$$

This definition is entirely analogous to that of the semigroup corresponding to one-dimensional Brownian motion, as presented in Example 1.5. Note, however, that now we are considering the operators T_t acting on the space $C_0(\mathbb{R}^n)$.

It can be proved that expression (1.7) defines a Feller-Dynkin semigroup on $C_0(\mathbb{R}^n)$, see [36, p. 242]. Once again we emphasize that, for the case $n = 1$ the generator of this semigroup is the operator $(\frac{1}{2}\Delta, C_0^2(\mathbb{R}))$; while for $n \geq 2$ the situation is more complicated, and the generator of the semigroup is a proper extension of $\frac{1}{2}\Delta$.

For the rest of this section we assume $E = \mathbb{R}^n$, but E could equally well be an n-dimensional smooth manifold. Recall that the lifetime ζ of the process X is defined as follows:

$$\zeta(\omega) := \inf\{t \geq 0 \ : \ X_t(\omega) = \triangle\}.$$

Also, we point out that the process associated to a Feller-Dynkin semigroup is called a Feller-Dynkin process.

Definition 1.9 (Feller-Dynkin diffusion). A Feller-Dynkin diffusion on $\mathbb{R}^n$ is a Feller-Dynkin process X with the additional properties that the trajectories $t \mapsto X_t(\omega)$ are continuous on $[0, \zeta)$, and the domain $\mathcal{D}(G)$ of the generator G of X contains $C_c^\infty(\mathbb{R}^n)$, the space of infinitely differentiable functions with compact support.

From now on the word diffusion will be reserved to mean a Feller-Dynkin diffusion. Let X be a diffusion, and denote by L the restriction of G to $C_c^\infty(\mathbb{R}^n)$. Then the operator L has the following properties:

(i) L is a linear map from $C_c^\infty(\mathbb{R}^n)$ to $C_0(\mathbb{R}^n)$.

(ii) L is local: if the functions $f, g \in C_c^\infty(\mathbb{R}^n)$ agree in some open neighborhood of a point x, then $Lf(x) = Lg(x)$ on that neighborhood.

(iii) L satisfies the maximum principle: if $f \in C_c^\infty(\mathbb{R}^n)$ attains its maximum at x and $f(x) \geq 0$, then $Lf(x) \leq 0$.

In addition, we have the following result, which we take as the basis for all our work in Chapter 2.

Theorem 1.2. *Let G be the generator of a Feller-Dynkin diffusion. The restriction L of G to $C_c^\infty(\mathbb{R}^n)$ is a second-order elliptic operator of the form*

$$Lf(x) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{n} b_i(x)\frac{\partial f}{\partial x_i}(x) - c(x)f(x),$$

where the functions $a_{ij}(\cdot)$, $b_i(\cdot)$ and $c(\cdot)$ are continuous, the matrix $(a_{ij}(x))$ is symmetric and nonnegative-definite for each x, and $c(x) \geq 0$ for each x.

The proof of this result can be found in [36, p. 258].

Chapter 2

Reversibility for Brownian motion with drift: first criterion

2.1 Invariant and reversible measures

In this chapter we will consider diffusion processes with values in $\mathbb{R}^n$, and we will describe them from the point of view of their infinitesimal generator. Recall from the previous chapter that the restriction of such generator to the space $C_c^\infty(\mathbb{R}^n)$ is a second-order differential operator of elliptic type. Our discussion will be restricted to operators of the form

$$Lf(x) = \frac{1}{2}\Delta f(x) + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i}f(x), \qquad \text{for } f \in C_c^\infty(\mathbb{R}^n). \tag{2.1}$$

We saw in the previous chapter, comments after Example 1.5, that the operator $(\frac{1}{2}\Delta, C_c^\infty(\mathbb{R}^n))$ corresponds to Brownian motion on $\mathbb{R}^n$. Thus it is natural to consider the operator L described above as corresponding to a "Brownian motion with drift", where the drift is determined by the first-order part.

We will assume from the outset that the drift coefficients $b_i(x)$ are smooth functions. Note that we can gather these coefficients as the components of a vector $b(x) = (b_1(x), \ldots, b_n(x))$ and thus obtain a smooth vector field (*i.e.* a smooth map from $\mathbb{R}^n$ to $\mathbb{R}^n$). Then the operator L can be expressed as

$$Lf = \frac{1}{2}\Delta f + \langle b, \nabla f\rangle, \qquad \text{for } f \in C_c^\infty(\mathbb{R}^n),$$

where $\nabla f = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})$ is the gradient of f, and $\langle\cdot,\cdot\rangle$ denotes the usual inner product in $\mathbb{R}^n$.

Next we start to explore the relation between the operator L and certain measures on $\mathbb{R}^n$. Recall that a Borel measure on $\mathbb{R}^n$ is a measure defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$. In all that follows whenever we say "measure" we will mean a regular Borel measure which is finite on compact sets, see [7] for precise definitions of these terms.

Definition 2.1. A measure μ on $\mathbb{R}^n$ is said to be an invariant measure for the operator L if

$$\int_{\mathbb{R}^n} Lf(x)\mu(dx) = 0 \qquad \forall f \in C_c^\infty(\mathbb{R}^n). \tag{2.2}$$

In case L has an invariant measure, we will say that L is well-behaved.

Remark. The term "well-behaved" that we have just defined is not standard, but we consider that giving a specific name to operators with an invariant measure will help improve the presentation. The definition of invariant measure as above does agree with the standard literature, see [10],[11],[24],[25].

Definition 2.2. A measure m on $\mathbb{R}^n$ is called a symmetrizing measure for L if

$$\int_{\mathbb{R}^n} Lf(x)\, g(x)\, m(dx) = \int_{\mathbb{R}^n} f(x)\, Lg(x)\, m(dx) \qquad \forall f, g \in C_c^\infty(\mathbb{R}^n). \tag{2.3}$$

Symmetrizing measures are also called reversible. In case such a measure exists, the operator L is said to be symmetrizable or also reversible.

Remark. If X is a diffusion process with semigroup $(T_t)_{t\geq 0}$, a (finite) measure μ is called invariant for $(T_t)_{t\geq 0}$ (or for the corresponding process X) if for all functions $f \in C_0(\mathbb{R}^n)$ and all $t \geq 0$,

$$\int_{\mathbb{R}^n} T_t f(x)\mu(dx) = \int_{\mathbb{R}^n} f(x)\mu(dx). \tag{2.2$'$}$$

In this case the measure μ can be normalized to become a probability measure. When there is an invariant probability measure μ for the process X, setting μ as the initial distribution makes the process become stationary (*i.e.* it has the same distribution at all times). For this reason invariant probability measures are also called stationary distributions, see [25]. The link with the definition of invariant measure we presented above is the following: If L denotes the generator of the semigroup $(T_t)_{t\geq 0}$ and the measure μ satisfies (2.2′), then it also satisfies (2.2), but the converse is not true (see [10, Remark 3.0]).

Similarly, a (finite) measure μ is called symmetrizing for the semigroup $(T_t)_{t\geq 0}$ if for all $f, g \in C_0(\mathbb{R}^n)$ and all $t \geq 0$,

$$\int_{\mathbb{R}^n} T_t f(x)\, g(x)\, \mu(dx) = \int_{\mathbb{R}^n} f(x)\, T_t g(x)\, \mu(dx). \tag{2.3$'$}$$

When the semigroup $(T_t)_{t\geq 0}$ has a symmetrizing measure μ, the corresponding process X becomes reversible, in the sense that if we set μ as the initial distribution, then both X and its time reversal have the same law. The technical definition of time reversal can be consulted in [35, Section VII.4]. Once again, condition (2.3′) is in general stronger than (2.3).

The first result that we present establishes that every reversible measure is also invariant. The papers of K. Handa suggest that this result is widely known, see [24] and [25]. In [26, p. 291] this result is mentioned as "being clear", though the authors deal with diffusions on a compact manifold there, and in this situation the proof is really trivial. We include our own proof of this implication for the particular kind of operators that we are considering.

Proposition 2.1. *Let L be an operator of the form* (2.1), *and let μ be a measure on $\mathbb{R}^n$. If μ is reversible for L, then μ is also invariant.*

PROOF. Let $f \in C_c^\infty(\mathbb{R}^n)$ be arbitrary. Since both f and Lf have compact support, we can find an open ball $B_r(0)$ centered at the origin with radius big enough to contain the supports of both these functions. Take a function $g \in C_c^\infty(\mathbb{R}^n)$ such that $g = 1$ on $\overline{B_r(0)}$. We have

$$\begin{aligned}\int_{\mathbb{R}^n} Lf(x)\mu(dx) &= \int_{B_r(0)} Lf(x)\mu(dx) = \int_{B_r(0)} Lf(x)\, g(x)\mu(dx) \\ &= \int_{\mathbb{R}^n} Lf(x)\, g(x)\mu(dx) = \int_{\mathbb{R}^n} f(x)\, Lg(x)\mu(dx) \\ &= \int_{B_r(0)} f(x)\, Lg(x)\mu(dx) = 0,\end{aligned}$$

where in the last step we use the fact that $Lg = 0$ on $B_r(0)$. ■

In what follows we will need to use the following formula of integration by parts: for any $f, g \in C_c^1(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x)\, g(x)dx = -\int_{\mathbb{R}^n} f(x)\frac{\partial g}{\partial x_i}(x)dx, \quad \text{for } i = 1, \ldots, n. \tag{2.4}$$

This formula of integration by parts is standard in the theory of partial differential equations, and in fact is taken as the basis for defining the notion of weak derivatives, see for instance [27, Definition 7.2.1] or any book on partial differential equations.

For functions f, g, let us denote (f, g) their inner product with respect to Lebesgue measure:

$$(f, g) := \int_{\mathbb{R}^n} f(x)g(x)dx.$$

Note that this expression makes sense at least for all continuous functions with compact support. Also note that, under this newly introduced notation, the equation $(Lf, g) = (f, Lg)$ would mean that the operator L is reversible under Lebesgue measure.

Proposition 2.2. *The Laplace operator Δ satisfies $(\Delta f, g) = (f, \Delta g)$ for all functions $f, g \in C_c^\infty(\mathbb{R}^n)$.*

PROOF. We use the formula of integration by parts (2.4). For arbitrary $f, g \in C_c^\infty(\mathbb{R}^n)$ we have

$$\begin{aligned}(\Delta f, g) &= \int_{\mathbb{R}^n} \Delta f(x) g(x) dx = \int_{\mathbb{R}^n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(x)\, g(x) dx \\ &= \sum_{i=1}^{n} \int_{\mathbb{R}^n} \frac{\partial^2 f}{\partial x_i^2}(x)\, g(x) dx = -\sum_{i=1}^{n} \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x)\, \frac{\partial g}{\partial x_i}(x) dx \\ &= \sum_{i=1}^{n} \int_{\mathbb{R}^n} f(x) \frac{\partial^2 g}{\partial x_i^2}(x) dx = \int_{\mathbb{R}^n} f(x) \sum_{i=1}^{n} \frac{\partial^2 g}{\partial x_i^2}(x) dx \\ &= \int_{\mathbb{R}^n} f(x) \Delta g(x) dx = (f, \Delta g).\end{aligned}$$

■

In this chapter we are interested in the study of conditions for operators L as in (2.1) to admit a reversible measure. As a consequence of Proposition 2.2 we see that the operator $\frac{1}{2}\Delta$ is reversible under Lebesgue measure. Thus it is natural to expect that whether L as defined above has the property of reversibility depends on the nature of the drift term. In Section 2.2 we will present Kolmogorov's criterion of reversibility, which gives conditions on the drift term so that the operator L is reversible (see Theorem 2.2 ahead).

The following technical result is taken from [26, Proposition V.4.4]. Since the authors present this result without a proof, we decided to include our own proof here.

Proposition 2.3. *Assume that the drift $b(x)$ is smooth. Then for the operator L defined in* (2.1) *we have $(Lf, g) = (f, L^*g)$, where L^* is given by*

$$L^*g(x) = \frac{1}{2}\Delta g(x) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(g(x) b_i(x)).$$

Equivalently, we can write

$$L^*g(x) = \frac{1}{2}\Delta g(x) - \operatorname{div}(g(x)b(x)) = \operatorname{div}\left(\frac{1}{2}\nabla g(x) - g(x)b(x)\right).$$

PROOF. Let us write $L = L_2 + L_1$, where

$$L_2 f(x) := \frac{1}{2}\Delta f(x) \qquad \text{and} \qquad L_1 f(x) := \sum_{i=1}^{n} b_i(x) \frac{\partial f}{\partial x_i}(x).$$

We can read L_i as "the i-th order part of L". We will find L^* in the form $L^* = L_2^* + L_1^*$, where L_i^* corresponds to L_i by $(L_i f, g) = (f, L_i^* g)$. As a consequence of Proposition 2.2 we know $L_2^* = L_2$, so it only remains to calculate L_1^*. We will use the formula of integration by parts (2.4) to show that

$$L_1^* g(x) = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}(g(x) b_i(x)).$$

For any $f, g \in C_c^\infty(\mathbb{R}^n)$ we have

$$\begin{aligned}
(L_1 f, g) &= \int_{\mathbb{R}^n} L_1 f(x)\, g(x) dx = \int_{\mathbb{R}^n} \sum_{i=1}^{n} b_i(x) \frac{\partial f}{\partial x_i}(x)\, g(x) dx \\
&= \sum_{i=1}^{n} \int_{\mathbb{R}^n} b_i(x) \frac{\partial f}{\partial x_i}(x)\, g(x) dx \\
&= -\sum_{i=1}^{n} \int_{\mathbb{R}^n} f(x) \frac{\partial}{\partial x_i}(b_i(x) g(x)) dx \\
&= \int_{\mathbb{R}^n} f(x) \left[-\sum_{i=1}^{n} \frac{\partial}{\partial x_i}(g(x) b_i(x)) \right] dx.
\end{aligned}$$

Recall that the divergence of a vector field ξ on $\mathbb{R}^n$ is defined by the expression $\operatorname{div} \xi(x) := \sum_{i=1}^{n} \frac{\partial \xi_i}{\partial x_i}(x)$. Keeping this in mind we can write

$$(L_1 f, g) = \int_{\mathbb{R}^n} f(x) \left[-\operatorname{div}(g(x) b(x)) \right] dx,$$

so we see that $L_1^* g(x) = -\operatorname{div}(g(x) b(x))$. To establish the last part of the claim it suffices to note that $\Delta f(x) = \operatorname{div}(\nabla f(x))$. ■

We saw in Proposition 2.1 that, for a given operator L, reversible measures are necessarily invariant. The question arises whether invariant measures exist at all, for only when this question has a positive answer does it make sense to invest our efforts in studying the reversible measures. Next we turn our attention to the problem of existence of invariant measures. For this, the best references are the papers [10],[11] of V.Bogachev and M.Röckner. In particular, [11, Theorem 1.2] is a result of existence of invariant measures under conditions of local integrability of the drift $b(x)$. Such condition clearly applies to our case since we are considering smooth drifts, but an important difference to point out is that this result is aimed at finding invariant *probability* measures, a restriction that we do not impose since, for instance, Brownian motion does not have an invariant probability measure, but it does have an invariant measure in the sense of Definition 2.1 (*i.e.* Lebesgue measure). The paper [10] has a whole section devoted to the problem of existence of invariant measures, although once again it is focused on invariant probability measures. An existence result is also presented in [26, Proposition V.4.5] for diffusions on a compact Riemannian manifold, and this result establishes also that the invariant measure is unique (up to a multiplicative constant). This result does not apply to the case which concerns us because we are working with diffusions on $\mathbb{R}^n$, a non-compact space. In fact, we will see ahead in Example 2.1 that for the kind of operators we are considering there may be more than one invariant measure, so uniqueness does not hold. For more discussions on invariant measures the reader is referred to the papers [4],[8],[12],[14].

Next we quote without proof a result on regularity of invariant measures.

Theorem 2.1. *Let L be a well-behaved operator of the form* (2.1). *If μ is an invariant measure for L (in the sense of* (2.2)*), then μ has a smooth density with respect to Lebesgue measure, so we may write $\mu(dx) = \varphi(x)dx$ for some $\varphi \in C^\infty(\mathbb{R}^n)$. Moreover, φ is a strictly positive function.*

Remark. This fact is established in [26, Proposition V.4.5], where the authors claim that the smoothness of invariant measures follows from a lemma due to H.Weyl. As stated in [42, Lemma 2], this latter result says that if $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a function such that

$$\int_{\mathbb{R}^n} u(x)\Delta f(x)dx = 0 \qquad \forall f \in C_c^\infty(\mathbb{R}^n),$$

then $u(x)$ is a harmonic function and, in particular, smooth. This result is also quoted in [27, p. 18] and [2, p. 319]; it seems to be known as Weyl's lemma. The way it is used in [26] suggests an extension of Weyl's lemma for second order differential operators of elliptic type. A more general version of Weyl's lemma, probably the one that [26] intends to quote, is presented in [3, Theorem 6.6] as a regularity result for elliptic problems. In the context of elliptic operators, Weyl's lemma establishes that if μ is a solution of $L^*\mu = 0$, that is, if

$$\int_{\mathbb{R}^n} Lf(x)\mu(dx) = 0 \qquad \forall f \in C_c^\infty(\mathbb{R}^n),$$

then μ has a smooth density, so we may write $\mu(dx) = \varphi(x)dx$ for some $\varphi \in C^\infty(\mathbb{R}^n)$. The fact that this density is a strictly positive function is also commented in [26, Proposition V.4.5].

The next proposition emphasizes the importance of the operator L^* for computing the invariant measures of L.

Proposition 2.4. *Let the measure μ have a smooth density, say $\mu(dx) = \varphi(x)dx$ where $\varphi \in C^\infty(\mathbb{R}^n)$. Then μ is an invariant measure for L if and only if its density solves the equation $L^*\varphi = 0$.*

PROOF. Equation (2.2) now reads

$$\int_{\mathbb{R}^n} Lf(x)\varphi(x)dx = 0 \qquad \forall f \in C_c^\infty(\mathbb{R}^n),$$

or in the newly introduced notation, $(Lf, \varphi) = 0$ for all $f \in C_c^\infty(\mathbb{R}^n)$. But from Proposition 2.3 we know $(Lf, \varphi) = (f, L^*\varphi)$, so the claim follows. ■

Let L be a well-behaved operator of the form (2.1), and let μ be an invariant measure for L. Since the density φ of μ is positive, writing $U(x) = -\log(\varphi(x))$ we see that $e^{-U(x)}dx = \varphi(x)dx = \mu(dx)$. Thus any invariant measure for L can be written in the form $e^{-U(x)}dx$ for some appropriately chosen function $U(x)$. The next result clarifies the meaning of "appropriately chosen" by establishing how the function $U(x)$ must be linked to the vector field $b(x)$ so that the measure $e^{-U(x)}dx$ is invariant for L. This and all remaining results in this section come from [26, Section V.4]. Recall that a vector field ξ on $\mathbb{R}^n$ is called divergence-free when $\operatorname{div}\xi(x) = 0$ for all $x \in \mathbb{R}^n$.

Proposition 2.5. *Let L be a well-behaved operator as in* (2.1). *A measure of the form $e^{-U(x)}dx$ is invariant for L if and only if there is a divergence-free vector field ξ on $\mathbb{R}^n$ such that*

$$b(x) = -\frac{1}{2}\nabla U(x) + \xi(x)e^{U(x)}. \tag{2.5}$$

PROOF. Let us assume that $e^{-U(x)}dx$ is an invariant measure for L. Using the expression for L^* obtained in Proposition 2.3 we get

$$\begin{aligned} L^*(e^{-U(x)}) &= \operatorname{div}\left(\frac{1}{2}\nabla(e^{-U(x)}) - e^{-U(x)}b(x)\right) \\ &= \operatorname{div}\left(e^{-U(x)}\left(-\frac{1}{2}\nabla U(x) - b(x)\right)\right). \end{aligned}$$

Define the vector field ξ as follows

$$\xi(x) := e^{-U(x)}\left(\frac{1}{2}\nabla U(x) + b(x)\right).$$

Proposition 2.4 says that the invariant measures come from solving the equation $L^*(e^{-U}) = 0$. From this and the previous calculation we see that the vector field ξ must have divergence zero, and it is clear that $b(x)$ and $U(x)$ are linked by equation (2.5).

Conversely, if $U \in C^\infty(\mathbb{R}^n)$ is a function that satisfies equation (2.5) for some divergence-free vector field ξ, then using Proposition 2.3 together with the above calculations we get

$$L^*(e^{-U(x)}) = \operatorname{div}\left(e^{-U(x)}\left(-\frac{1}{2}\nabla U(x) - b(x)\right)\right) = -\operatorname{div}\xi(x) = 0,$$

so $e^{-U(x)}dx$ must be an invariant measure by virtue of Proposition 2.4. ■

We have been using the notation $\langle \cdot, \cdot \rangle$ for the usual inner product in $\mathbb{R}^n$. It is also customary to use the "dot" notation for this inner product; so that for instance, $u \cdot v$ is another way to write $\langle u, v \rangle$. The dot notation is especially suited for the vector calculus operations, and we will start using these two notations interchangeably, hoping that this causes no confusion on the reader.

Proposition 2.6. *Let L be a well-behaved operator as in* (2.1). *The measure $e^{-U(x)}dx$ is invariant for L if and only if the function U satisfies*

$$\frac{1}{2}\nabla U \cdot \nabla U + b \cdot \nabla U - \frac{1}{2}\Delta U - \operatorname{div}(b) = 0. \tag{2.6}$$

PROOF. Using the result of Proposition 2.3, a simple calculation yields the following:

$$L^*(e^{-U(x)}) = e^{-U(x)}\left[\frac{1}{2}\nabla U(x) \cdot \nabla U(x) + b(x) \cdot \nabla U(x) - \frac{1}{2}\Delta U(x) - \operatorname{div}(b(x))\right].$$

Thus we see that $L^*(e^{-U(x)}) = 0$ if and only if the function $U(x)$ satisfies equation (2.6). ■

Given a measure of the form $e^{-U(x)}dx$, we will designate by $(f,g)_U$ the inner product associated to this measure:

$$(f,g)_U = \int_{\mathbb{R}^n} f(x)g(x)e^{-U(x)}dx.$$

Proposition 2.7. *For any operator L of the form* (2.1) *we have $(Lf,g)_U = (f,\hat{L}g)_U$, where $\hat{L}$ is given by*

$$\hat{L}g = \frac{1}{2}\Delta g - [b + \nabla U] \cdot \nabla g + \left[\frac{1}{2}\nabla U \cdot \nabla U + b \cdot \nabla U - \frac{1}{2}\Delta U - \operatorname{div}(b)\right] g.$$

PROOF. As in the proof of Proposition 2.3, let us write $L = L_2 + L_1$; we will compute $\hat{L}$ in the form $\hat{L} = \hat{L}_2 + \hat{L}_1$, where $\hat{L}_i$ corresponds to L_i by the equation $(L_i f, g)_U = (f, \hat{L}_i g)_U$, for $i = 1, 2$. We will use the formula of integration by parts (2.4) to show that

$$\hat{L}_1 g(x) = g(x)(b(x) \cdot \nabla U(x)) - g(x)\operatorname{div} b(x) - b(x) \cdot \nabla g(x).$$

and

$$\hat{L}_2 g(x) = \frac{1}{2}\Delta g(x) - \nabla g(x) \cdot \nabla U(x) - \frac{1}{2}g(x)\Delta U(x) + \frac{1}{2}g(x)\nabla U(x) \cdot \nabla U(x).$$

Let us start with the calculation of $\hat{L}_1$. For any $f, g \in C_c^\infty(\mathbb{R}^n)$ we have

$$
\begin{aligned}
(L_1 f, g) &= \int_{\mathbb{R}^n} L_1 f(x)\, g(x) e^{-U(x)} dx = \int_{\mathbb{R}^n} \sum_{i=1}^n b_i(x) \frac{\partial f}{\partial x_i}(x)\, g(x) e^{-U(x)} dx \\
&= \sum_{i=1}^n \int_{\mathbb{R}^n} b_i(x) \frac{\partial f}{\partial x_i}(x)\, g(x) e^{-U(x)} dx \\
&= -\sum_{i=1}^n \int_{\mathbb{R}^n} f(x) \frac{\partial}{\partial x_i} \left[b_i(x) g(x) e^{-U(x)} \right] dx \\
&= -\sum_{i=1}^n \int_{\mathbb{R}^n} f(x) \left[\frac{\partial}{\partial x_i} (b_i(x) g(x)) - b_i(x) g(x) \frac{\partial U}{\partial x_i}(x) \right] e^{-U(x)} dx \\
&= \int_{\mathbb{R}^n} f(x) \left[-\sum_{i=1}^n \frac{\partial}{\partial x_i} (g(x) b_i(x)) + g(x) \sum_{i=1}^n b_i(x) \frac{\partial U}{\partial x_i}(x) \right] e^{-U(x)} dx \\
&= \int_{\mathbb{R}^n} f(x) \left[-\operatorname{div}(g(x) b(x)) + g(x) b(x) \cdot \nabla U(x) \right] e^{-U(x)} dx,
\end{aligned}
$$

so upon computation of the divergence $\operatorname{div}(gb)$ we get the expression for $\hat{L}_1$. Now let us calculate $\hat{L}_2$. For arbitrary $f, g \in C_c^\infty(\mathbb{R}^n)$ we have

$$
\begin{aligned}
(L_2 f, g) &= \int_{\mathbb{R}^n} L_2 f(x)\, g(x) e^{-U(x)} dx \\
&= \int_{\mathbb{R}^n} \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x)\, g(x) e^{-U(x)} dx \\
&= \frac{1}{2} \sum_{i=1}^n \int_{\mathbb{R}^n} \frac{\partial^2 f}{\partial x_i^2}(x)\, g(x) e^{-U(x)} dx \\
&= -\frac{1}{2} \sum_{i=1}^n \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x)\, \frac{\partial}{\partial x_i} \left[g(x) e^{-U(x)} \right] dx \\
&= \frac{1}{2} \sum_{i=1}^n \int_{\mathbb{R}^n} f(x) \frac{\partial^2}{\partial x_i^2} \left[g(x) e^{-U(x)} \right] dx \\
&= \frac{1}{2} \sum_{i=1}^n \int_{\mathbb{R}^n} f(x) \frac{\partial}{\partial x_i} \left[\left(\frac{\partial g}{\partial x_i}(x) - g(x) \frac{\partial U}{\partial x_i}(x) \right) e^{-U(x)} \right] dx \\
&= \int_{\mathbb{R}^n} f(x) \frac{1}{2} \sum_{i=1}^n \frac{\partial}{\partial x_i} \left[\left(\frac{\partial g}{\partial x_i}(x) - g(x) \frac{\partial U}{\partial x_i}(x) \right) e^{-U(x)} \right] dx,
\end{aligned}
$$

so upon differentiating and factoring out $e^{-U(x)}$ the last expression becomes

$$\int_{\mathbb{R}^n} f(x) \sum_{i=1}^{n} \frac{1}{2} \left[\frac{\partial}{\partial x_i} \left(\frac{\partial g}{\partial x_i} - g \frac{\partial U}{\partial x_i} \right)(x) - \left(\frac{\partial g}{\partial x_i} - g \frac{\partial U}{\partial x_i} \right)(x) \frac{\partial U}{\partial x_i}(x) \right] e^{-U(x)} dx,$$

and after performing the last partial differentiation we obtain the expression for $\hat{L}_2$. To conclude the proof we just need to put the two pieces $\hat{L}_1$ and $\hat{L}_2$ together. ■

Corollary 2.1. *If $e^{-U(x)}dx$ is an invariant measure for L, then $\hat{L}$ is given by*

$$\hat{L}g = \frac{1}{2}\Delta g - [b + \nabla U] \cdot \nabla g.$$

PROOF. Follows from Proposition 2.7 and equation (2.6). ■

2.2 Kolmogorov's criterion

In this section we present and prove a very important characterization of reversibility, the criterion of Kolmogorov. This criterion establishes a condition for an operator L to admit a reversible measure. As pointed out earlier, the condition is on the first-order part of L, and it says in brief that reversibility is equivalent to the vector field $b(x)$ being conservative. This criterion can be found in [26, Theorem V.4.6] for the case of diffusions on a compact manifold, the proof that we offer below is an adaptation to the case of diffusions on $\mathbb{R}^n$ that we are considering.

Theorem 2.2. *The operator L defined in* (2.1) *is symmetrizable (* i.e. *has a reversible measure) if and only if $b(x) = \nabla F(x)$ for some $F \in C^\infty(\mathbb{R}^n)$. In this case, the reversible measures are of the form $\quad Constant \cdot e^{2F(x)}dx$.*

PROOF. Suppose first that L is symmetrizable, and let $m(dx)$ be a reversible measure. Then m is also an invariant measure (Proposition 2.1), so we can write $m(dx) = e^{-U(x)}dx$ for some function U. Reversibility under this measure means that $(Lf, g)_U = (f, Lg)_U$ holds for all $f, g \in C_c^\infty(\mathbb{R}^n)$. But Proposition 2.7 says that we also have $(Lf, g)_U = (f, \hat{L}g)_U$ for all $f, g \in C_c^\infty(\mathbb{R}^n)$, so we get $Lg = \hat{L}g$ for all $g \in C_c^\infty(\mathbb{R}^n)$. Taking into account Corollary 2.1 we see that $b(x) = -[b(x) + \nabla U(x)]$, and in consequence $b(x) = \nabla(-\frac{1}{2}U(x))$.

Conversely, if $b(x) = \nabla F(x)$ for some function $F \in C^\infty(\mathbb{R}^n)$, then it is easy to see that $U(x) := -2F(x)$ satisfies equation (2.6). Hence $e^{-U(x)}dx$ is an invariant measure, and we have

$$b(x) = \nabla F(x) = -\frac{1}{2}\nabla U(x).$$

It follows that $\hat{L}g = Lg$ for all $g \in C_c^\infty(\mathbb{R}^n)$, which implies that $(Lf, g)_U = (f, Lg)_U$ holds for all $f, g \in C_c^\infty(\mathbb{R}^n)$. But this says that L is symmetrizable with reversible measure $e^{-U(x)}dx$. ∎

An immediate consequence of this theorem is the following

Corollary 2.2. *If the operator L is symmetrizable, then its reversible measure is unique (up to a positive multiplicative constant).*

As another interesting consequence of Theorem 2.2, we will prove in the next corollary that in the one-dimensional case reversibility is always possible.

Corollary 2.3. *In the one-dimensional case every operator L of the kind* (2.1) *is reversible.*

PROOF. First let us point out that in the one dimensional case the operator L defined in (2.1) acts on functions f according to the expression

$$Lf(x) = \frac{1}{2}f''(x) + b(x)f'(x).$$

Assume that we are given an operator L as above, with a smooth drift coefficient $b(x)$. Define the function $F(x)$ as follows

$$F(x) := \int_0^x b(t)dt.$$

Then it is clear that $F'(x) = b(x)$ for all $x \in \mathbb{R}$. According to Theorem 2.2, we see that any operator L with a smooth drift coefficient is reversible, and the reversible measure is of the form $e^{2F(x)}dx$ with $F(x)$ as defined above. ∎

We know that reversible measures are invariant, and in fact the invariance of the measure $e^{2F(x)}dx$ from the previous proof can be checked directly using equation (2.6) with $U(x) = -2F(x)$. The question arises whether these are all invariant measures for the given operator. The answer to this question involves solving equation (2.6) for $U(x)$, a problem that may be quite hard to tackle depending on the given drift $b(x)$. We will give an explicit answer to this question in the next example, were a specific kind of drift is considered.

Example 2.1 (One-dimensional Gaussian case). Assume we fix $U(x) = \alpha x^2$ with $\alpha > 0$, so we are looking at the Gaussian measure $e^{-\alpha x^2}dx$. First we are going to find for which functions $b(x)$ the operator L has $e^{-\alpha x^2}dx$ as an invariant measure. We know in this case the drift coefficient $b(x)$ of L must satisfy relation (2.6), which in the present case yields the equation

$$b'(x) - 2\alpha x b(x) = 2\alpha^2 x^2 - \alpha.$$

This is a linear differential equation of first order, and we can solve it using the integrating factor $e^{\int(-2\alpha x)dx} = e^{-\alpha x^2}$. After a few computations, we arrive at the general solution

$$b(x) = -\alpha x + Ce^{\alpha x^2}, \tag{2.7}$$

where C is an arbitrary constant. Note that this is nothing but equation (2.5), which we have obtained by a direct computation, the constant C playing the role of the divergence-free vector field. Thus $e^{-\alpha x^2}dx$ is invariant for L if and only if $b(x)$ is of the above form. We see that there are many operators L that are invariant under $e^{-\alpha x^2}dx$, and the reversibility of these operators is guaranteed by Corollary 2.3. Let us pick a particular one of these operators, say by fixing $C = 1$, so the drift coefficient is

$$b(x) = -\alpha x + e^{\alpha x^2}.$$

We can find easily an antiderivative for this function, just define

$$F(x) := -\frac{\alpha x^2}{2} + \int_0^x e^{\alpha t^2}dt.$$

Note that the measure $e^{2F(x)}dx$ symmetrizes the operator L, where

$$2F(x) := -\alpha x^2 + 2\int_0^x e^{\alpha t^2}dt.$$

Being a reversible measure, this is also an invariant measure for the given operator. But we know already that this operator is invariant under $e^{-\alpha x^2}dx$, a measure that is clearly different form $e^{2F(x)}dx$ with $F(x)$ as above. Thus, we learn from this example that for a given operator L there may be more than one invariant measure. ◊

Let us continue the line of reasoning of the previous example and try to apply it to the n-dimensional case, for $n \geq 2$. Once again, let us assume we are given a function $U(x)$, and we want to find operators L for which $e^{-U(x)}dx$ is an invariant measure. We have seen that there must be some divergence-free vector field $\xi(x) = (\xi_1(x), \ldots, \xi_n(x))$ such that the drift coefficient $b(x)$ and the function $U(x)$ are linked by equation (2.5), which for convenience we rewrite here:

$$b(x) = -\frac{1}{2}\nabla U(x) + \xi(x)e^{U(x)}. \tag{2.5}$$

If $\xi \equiv 0$, it is clear that L is symmetrizable with reversible measure $e^{-U(x)}dx$. On the other hand, if ξ is not the zero vector field, it is not immediate how we can construct a potential function $F(x)$ for $b(x)$. Some additional conditions may be required for this to work. This is illustrated in the next example.

Example 2.2 (Two-dimensional Gaussian case). Assume we are given the function $U(x,y) = \alpha x^2 + 2\beta xy + \delta y^2$, where α, δ are positive constants; and we want to see what kind of operators L are reversible under the Gaussian measure $e^{-U(x,y)}dxdy$. The discussion above tells us that in this case we must have $b(x,y) = -\frac{1}{2}\nabla U(x,y)$, so it follows that $b(x,y) = -(\alpha x + \beta y, \beta x + \delta y)$. Note that in particular the operators L for which the Gaussian measure $e^{-U(x,y)}dxdy$ is reversible must have linear drift. Linearity is obvious since we can write $b(x,y) = -M \cdot (x,y)'$ where M is the matrix $M = \begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix}$.

On the other hand, assume we are given an operator L with linear drift coefficient $b(x,y) = -(\alpha x + \beta y, \gamma x + \delta y)$. We want to find conditions for L to be reversible, and in the positive case to compute the reversible measure for such operator. This amounts to finding the corresponding function $U(x,y)$. Taking $\xi \equiv 0$ in equation (2.5), we have

$$\nabla U(x,y) = -2b(x,y) = (2\alpha x + 2\beta y, 2\gamma x + 2\delta y).$$

Let us write $f_1(x,y) = 2\alpha x + 2\beta y$ and $f_2(x,y) = 2\gamma x + 2\delta y$. The equation above says that we must have $\frac{\partial U}{\partial x} = f_1$ and $\frac{\partial U}{\partial y} = f_2$. Since U is smooth we also have $\frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x}$, so necessarily $\frac{\partial f_2}{\partial x} = \frac{\partial f_1}{\partial y}$. But note that $\frac{\partial f_2}{\partial x} = 2\gamma$ and $\frac{\partial f_1}{\partial y} = 2\beta$, so we conclude that $\gamma = \beta$. Note in particular that reversibility forces the matrix $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ to be symmetric.

Next, from $\frac{\partial U}{\partial x} = 2\alpha x + 2\beta y$ we find that the function U must be of the form $U(x,y) = \alpha x^2 + 2\beta xy + h(y)$ for some function $h(y)$ which depends on y only. Differentiating this with respect to y we get $\frac{\partial U}{\partial y} = 2\beta x + h'(y)$. But we know $\frac{\partial U}{\partial y} = 2\gamma x + 2\delta y$, so taking into account the fact that $\gamma = \beta$ we find $h(y) = \delta y^2 + k$ for some arbitrary constant k. Thus the general from of U is

$$U(x,y) = \alpha x^2 + 2\beta xy + \delta y^2 + k.$$

Note that in this case the measure $e^{-U(x,y)}dxdy$ will be Gaussian precisely when $\alpha, \delta > 0$. $\diamondsuit$

Remark. The previous example established incidentally that an operator L with linear drift $b(x,y) = -(\alpha x + \beta y, \gamma x + \delta y)$ is reversible if and only if the matrix $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is symmetric. This allows us to think of examples where reversibility is not possible, just take operators L with linear drift corresponding to a non-symmetric matrix M.

Chapter 3

Reversibility for Brownian motion with drift: second criterion

3.1 Quasi-invariant measures

This chapter explores the relation between reversibility and quasi-invariance for a measure on the state space of the process associated to a given operator L of the form (2.1). Let us recall that our focus so far has been on n-dimensional diffusions, *i.e.* the case when the state space is $\mathbb{R}^n$. Since the discussion in this section is valid for more general state spaces, we will start working on an abstract state space E, and we will specialize to the case $E = \mathbb{R}^n$ in the next section, where we will present and prove a second characterization for reversibility.

Let E be a topological space. For measurability purposes we equip E with the Borel σ-algebra $\mathcal{B}(E)$. By a measurable transformation on E we mean a measurable bijection from E to E with a measurable inverse. For the sake of simplicity, we will use the word 'transformation' to refer to a measurable transformation.

Definition 3.1. Let E be a topological space. A group of transformations on E is a family Γ of transformations on E which is closed under the operation of composition, contains the identity and also contains the inverse of each of its elements.

Suppose that the space E is equipped with a transformation group Γ. Given a measure μ on E, for each $\gamma \in \Gamma$ we can define a new measure μ_γ by means of $\mu_\gamma(A) = \mu(\gamma(A)) = \mu((\gamma^{-1})^{-1}(A))$, for $A \in \mathcal{B}(E)$. It is also customary to write $\mu \circ \gamma$ for the measure μ_γ just defined.

Definition 3.2. Let E be a topological space equipped with a transformation group Γ. A measure μ on $(E, \mathcal{B}(E))$ is said to be Γ-quasi-invariant if for each $\gamma \in \Gamma$ the measures μ_γ and μ are equivalent, *i.e.* they have the same sets of measure zero.

The definition of quasi-invariance we have just presented was borrowed from [29] where the study of quasi-invariant measures is approached from an abstract and rather general standpoint. For our specific purposes, we should keep in mind that the abstract space E plays the role of the state space of a process under consideration (*i.e.* that corresponding to a given operator L), and we want to consider an analog of the group of "shifts" $\{S_f\}$ introduced in [24] to establish the connection between reversibility and quasi-invariance. The groups of transformations on E that we consider are indexed by some vector space V, so we will write $S = \{S_v\}_{v \in V}$ to denote a generic transformation group on E. In other words, for each $v \in V$, the mapping $S_v : E \longrightarrow E$ is a bijection, and the following properties hold:

$$
\begin{aligned}
&1) \qquad S_{u+v}(x) = S_u(S_v(x)),\\
&2) \qquad S_0(x) = x.
\end{aligned}
$$

With this in mind, we will adapt the definition of quasi-invariance to suit our purposes. The definition we present next coincides with that used by K. Handa in his papers [24], [25].

Definition 3.3. Let E be a topological space equipped with a transformation group $S = \{S_v\}_{v \in V}$, and let $\Lambda : V \times E \longrightarrow \mathbb{R}$ be such that for each $v \in V$ the function $x \longmapsto \Lambda(v, x)$ is Borel measurable. We say that a measure m on E is S-quasi-invariant with cocycle Λ if for each $v \in V$ the measures m and $m \circ S_v$ are equivalent and the density is given by

$$
\frac{d(m \circ S_v)}{dm}(x) = e^{\Lambda(v,x)}, \qquad m\text{- a.s.}
$$

The first result that we present concerns a certain identity that is necessarily satisfied by the cocycle Λ corresponding to a quasi-invariant measure. This result is already mentioned in the papers [24], [25]; see also [37]. For the sake of clarity we present it here as a proposition with its own proof.

Proposition 3.1. *Let E be a topological space equipped with a transformation group $S = \{S_v\}_{v \in V}$. Consider a measure m on E, and assume that m is quasi-invariant under S with cocycle Λ. Then the map Λ satisfies the following identity:*

$$
\Lambda(u + v, x) = \Lambda(u, S_v(x)) + \Lambda(v, x), \qquad m\text{-}a.s. \tag{3.1}
$$

PROOF. By definition we know that if m is quasi-invariant under S, then for any $v \in V$ the measures m and $m \circ S_v$ are mutually absolutely continuous. Then on one hand, from the definition of Radon-Nikodym density we have

$$dm \circ S_{u+v}(x) = e^{\Lambda(u+v,x)} dm(x).$$

On the other hand, using the transformation group properties we have

$$\begin{aligned} dm \circ S_{u+v}(x) &= dm \circ S_u(S_v(x)) \\ &= e^{\Lambda(u,S_v(x))} dm(S_v(x)) \\ &= e^{\Lambda(u,S_v(x))} e^{\Lambda(v,x)} dm(x) \\ &= e^{\Lambda(u,S_v(x))+\Lambda(v,x)} dm(x). \end{aligned}$$

Hence we see that for m-almost every x, the following equality is valid

$$e^{\Lambda(u+v,x)} = e^{\Lambda(u,S_v(x))+\Lambda(v,x)};$$

and this in turn implies that $\Lambda(v,x)$ satisfies the identity (3.1). ■

In the literature equation (3.1) is commonly referred to as the cocycle identity. In fact, given a transformation group $S = \{S_v\}_{v\in V}$ on the space E, a map $\Lambda : V \times E \longrightarrow \mathbb{R}$ is called an S-cocycle when it satisfies (3.1). The reader can find more details about cocycles in the monograph [1].

3.2 Reversibility and Quasi-invariance

Now let us get back to the case of finite-dimensional diffusions, so from now on we take $E = \mathbb{R}^n$. For the rest of this section we work with a specific group of transformations in $\mathbb{R}^n$ defined as follows: $S_v(x) := x + v$ for $v \in \mathbb{R}^n$. In other words, $S = \{S_v\}_{v\in\mathbb{R}^n}$ is the group of all translations in $\mathbb{R}^n$. The next proposition provides a host of examples of quasi-invariant measures on $\mathbb{R}^n$.

Proposition 3.2. *Let $S = \{S_v\}_{v\in V}$ be a transformation group on $\mathbb{R}^n$, and let $U : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a continuous function. Define the measure $m(dx) = e^{-U(x)}dx$. Then m is S-quasi-invariant with cocycle*

$$\Lambda(v,x) = U(x) - U(S_v(x)). \tag{3.2}$$

PROOF. First, let us establish that a function Λ as defined in (3.2) satisfies the cocycle identity:

$$\begin{aligned}\Lambda(u+v,x) &= U(x) - U(S_{u+v}(x)) \\ &= U(x) - U(S_v(x)) + U(S_v(x)) - U(S_{u+v}(x)) \\ &= U(x) - U(S_v(x)) + U(S_v(x)) - U(S_u(S_v(x))) \\ &= \Lambda(v,x) + \Lambda(u, S_v(x)).\end{aligned}$$

Next, to prove the quasi-invariance, let $v \in V$ be arbitrary. We have

$$(m \circ S_v)(dx) = e^{-U(S_v(x))}dx = e^{U(x)-U(S_v(x))}e^{-U(x)}dx = e^{\Lambda(v,x)}m(dx),$$

so we see that $m \circ S_v$ and m are equivalent with the required cocycle. ■

It is worth pointing out that the proof of this proposition does not depend on the specific group of transformations being used. This fact makes it possible to extend this result to more general spaces, provided that they can be equipped with a "uniform measure" (that is an analog of Lebesgue measure). For instance, this is suggested in [24, comments right after Theorem 2.1] where the author uses equation (3.2) as starting point to derive the cocycle corresponding to certain measure-valued diffusion. The next corollary imitates this derivation for the specific case that we are studying.

Corollary 3.1. *In the setting of* Proposition 3.2, *assume further that U is continuously differentiable. Then the measure $m(dx) = e^{-U(x)}dx$ is S-quasi-invariant with cocycle*

$$\Lambda(v,x) = -\int_0^1 \nabla U(S_{tv}(x)) \cdot v\, dt. \tag{3.3}$$

PROOF. Note that we can write

$$U(x) - U(S_v(x)) = U(S_0(x)) - U(S_v(x)) = -\int_0^1 \frac{d}{dt} U(S_{tv}(x))dt.$$

Using the chain rule we get

$$\frac{d}{dt}U(S_{tv}(x)) = \nabla U(S_{tv}(x)) \cdot \frac{d}{dt}(S_{tv}(x)) = \nabla U(S_{tv}(x)) \cdot v,$$

so it follows that the cocycle $\Lambda(v,x)$ can be expressed as in (3.3). ■

Let us comment on the usefulness of the previous two results for the reversible case. Consider an operator L as in (2.1), and assume that the measure m is reversible for L. From the discussion in Chapter 2 we know that m should be of the form $m(dx) = e^{-U(x)}dx$ for some "potential function" U. Then by Proposition 3.2 the measure m is quasi-invariant with respect to $\{S_v\}_{v\in\mathbb{R}^n}$, and from Corollary 3.1 the cocycle is given by the expression (3.3). Furthermore, we know from Kolmogorov's criterion that in the reversible case we have $b(x) = -\frac{1}{2}\nabla U(x)$, where $b(x)$ is the drift of the operator L. This allows us to further write the cocycle as follows:

$$\Lambda(v, x) = 2\int_0^1 b(S_{tv}(x)) \cdot v\, dt. \tag{3.4}$$

The main theorem of this chapter, Theorem 3.1 ahead, will establish that *a measure m on $\mathbb{R}^n$ is a reversible measure for the operator L if and only if m is quasi-invariant under the group $\{S_v\}_{v\in\mathbb{R}^n}$ of all translations with cocycle defined as in* (3.4).

This theorem makes evident the importance of the expression for the cocycle that we have derived. The proof of this result will be given at the end of this section, after we have established a couple of technical lemmas. It is also important to mention that the expression (3.4) of the cocycle $\Lambda(v, x)$ also works for diffusions on more general state spaces. For instance, [24] uses formally an expression similar to (3.4) for the cocycle associated to a reversible measure-valued Fleming-Viot process.

The next proposition summarizes some properties of the family of transformations $S = \{S_v\}_{v\in\mathbb{R}^n}$ that we may need to use. Note that property (S.1) below merely confirms the fact that this family is a transformation group. The proof will be omitted since it amounts to simple calculations.

Proposition 3.3. *The family $S = \{S_v\}_{v\in\mathbb{R}^n}$ of all translations in $\mathbb{R}^n$ satisfies the following properties:*

(S.1) $S_0(x) = x$ and $S_u(S_v(x)) = S_{u+v}(x)$. In particular, $(S_v)^{-1} = S_{-v}$.

(S.2) For any $x, y \in \mathbb{R}^n$, there is a unique $v \in \mathbb{R}^n$ such that $y = S_v(x)$.

(S.3) For each $v \in \mathbb{R}^n$, $S_v(x)$ is continuously differentiable in x and

$$D_x(S_v(x)) = I_n, \quad \textit{in other words,} \quad \frac{\partial (S_v(x))_i}{\partial x_j} = \delta_{ij}.$$

The following technical lemma will be useful for establishing the proof of Theorem 3.1. This lemma is motivated by [25, Lemma 2.1].

Lemma 3.1. *Let Λ be given by (3.4). Fix an arbitrary $v \in \mathbb{R}^n$. For any given $g \in C_c^\infty(\mathbb{R}^n)$ and $t \in \mathbb{R}$, define*

$$g_t(x) = g(S_{-tv}(x)) \exp\{-\Lambda(tv, S_{-tv}(x))\}.$$

Then $g_t \in C_c^\infty(\mathbb{R}^n)$ *for all* $t \in \mathbb{R}$, *and*

$$\frac{d}{dt}g_t(x) = -2(b(x)\cdot v)g_t(x) - v\cdot\nabla g_t(x). \tag{3.5}$$

PROOF. Let us designate $F_t(x) = \Lambda(tv, S_{-tv}(x))$, so we can write the function g_t as $g_t(x) = g(S_{-tv}(x))e^{-F_t(x)}$. We can see from equation (3.4) that, as function of its second argument, $\Lambda(v,x)$ has continuous derivatives of all orders; so it is clear that $F_t \in C^\infty(\mathbb{R}^n)$. Then it follows from property (S.3) that $g_t \in C_c^\infty(\mathbb{R}^n)$. Note further that

$$\nabla g_t(x) = e^{-F_t(x)}\nabla(g\circ S_{-tv})(x) - g_t(x)\nabla F_t(x). \tag{3.6}$$

Using (S.3) we have $\nabla(g\circ S_{-tv})(x) = D_x(S_{-tv}(x))\nabla g(S_{-tv}(x)) = \nabla g(S_{-tv}(x))$, so it is clear that

$$v\cdot\nabla(g\circ S_{-tv})(x) = v\cdot\nabla g(S_{-tv}(x)) = -\frac{d}{dt}g(S_{-tv}(x)), \tag{3.7}$$

where we have used the chain rule to establish the last equality. Now observe that

$$F_t(x) = \Lambda(tv, S_{-tv}(x)) = 2\int_0^t b(S_{-rv}(x))\cdot v\,dr. \tag{3.8}$$

Let us define an auxiliary function $h(x) := b(x)\cdot v$, so we have $(h\circ S_{-rv})(x) = h(S_{-rv}(x)) = b(S_{-rv}(x))\cdot v$. From this and equation (3.8) it follows that

$$\nabla F_t(x) = 2\int_0^t \nabla(h\circ S_{-rv})(x)\,dr.$$

Then using the same calculation as in (3.7) with $h(x)$ instead of $g(x)$ we get

$$\begin{aligned}
v\cdot\nabla F_t(x) &= 2\int_0^t v\cdot\nabla(h\circ S_{-rv})(x)dr\\
&= -2\int_0^t \frac{d}{dr}h(S_{-rv}(x))dr\\
&= -2\left[h(S_{-tv}(x)) - h(S_0(x))\right]\\
&= -2\left[b(S_{-tv}(x))\cdot v - b(x)\cdot v\right]\\
&= -\frac{d}{dt}F_t(x) + 2b(x)\cdot v.
\end{aligned}$$

Combining the last calculation with expressions (3.6) and (3.7) we obtain

$$\begin{aligned}
v\cdot\nabla g_t(x) &= -e^{-F_t(x)}\frac{d}{dt}g(S_{-tv}(x)) - g_t(x)\left[-\frac{d}{dt}F_t(x) + 2b(x)\cdot v\right]\\
&= -\frac{d}{dt}g_t(x) - 2(b(x)\cdot v)g_t(x).
\end{aligned}$$

This finishes the proof. ■

The next lemma establishes a useful way to prove reversibility of a measure. This result is already contained in [25, proof of Theorem 2.1]. The version that we present here is adapted to suit our purposes, and we also present an explicit proof.

Lemma 3.2. *Let L be an operator of the form (2.1), and let m be a measure on $\mathbb{R}^n$. Then m is a reversible measure for L if and only if*

$$\int (Lf)(x)g(x)m(dx) = -\frac{1}{2}\int \nabla f(x)\cdot\nabla g(x)m(dx) \quad \forall f,g \in C_c^\infty(\mathbb{R}^n). \tag{3.9}$$

PROOF. We know from equation (2.3) that m is a reversible measure for L if and only if the following symmetry holds:

$$\int (Lf)(x)g(x)m(dx) = \int f(x)(Lg)(x)m(dx) \qquad \forall f,g \in C_c^\infty(\mathbb{R}^n).$$

Now, a direct computation shows that the following identity holds for any pair of smooth functions f, g:

$$(Lf)(x)g(x) + f(x)(Lg)(x) - L(fg)(x) = -\nabla f(x)\cdot\nabla g(x). \tag{3.10}$$

Assume that the measure m is reversible for L. In particular, this implies that m is an invariant measure for L. Now take $f, g \in C_c^\infty(\mathbb{R}^n)$, and note that their product fg is also in $C_c^\infty(\mathbb{R}^n)$, so $\int L(fg)(x)m(dx) = 0$. Then integrating both sides of (3.10) with respect to m we get (3.9).

Conversely, assume that (3.9) holds. Since the right hand side of this equation is symmetric in f and g, this implies that the left hand side is also symmetric; *i.e.* (2.3) holds. ■

We are almost ready for the main result of this chapter, Theorem 3.1 to follow. Recall that, according with Definition 3.3, given a measurable function Λ on $\mathbb{R}^n \times \mathbb{R}^n$ we say that the measure m on $\mathbb{R}^n$ is quasi-invariant under $\{S_v\}_{v\in\mathbb{R}^n}$ with cocycle Λ if for every $v \in \mathbb{R}^n$, the measures m and $m \circ S_v = m \circ (S_{-v})^{-1}$ are mutually absolutely continuous with density given by

$$\frac{dm \circ S_v}{dm}(x) = e^{\Lambda(v,x)}, \qquad m\text{- a.s.} \tag{3.11}$$

Note that in order to prove (3.11) we must show that for each $v \in \mathbb{R}^n$ fixed,

$$\int_{\mathbb{R}^n} \Phi(S_{-v}(x))m(dx) = \int_{\mathbb{R}^n} \Phi(x)e^{\Lambda(v,x)}m(dx)$$

holds for all functions Φ in some measure determining class $\mathcal{D}$. Moreover, replacing $\Phi(x)$ by $\Phi(x)e^{-\Lambda(v,x)}$ we see that it is sufficient to prove

$$\int_{\mathbb{R}^n} \Phi(S_{-v}(x))e^{-\Lambda(v,S_{-v}(x))}m(dx) = \int_{\mathbb{R}^n} \Phi(x)m(dx), \qquad \forall \Phi \in \mathcal{D}. \tag{3.12}$$

For the proof of the next theorem we will need to use the notion of "bump function". We will not construct such functions explicitly here, but merely recall that for any $n \geq 1$ and any $0 < r < R$, there is a smooth function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $0 \leq \varphi(x) \leq 1$ everywhere, $\varphi \equiv 1$ on $\overline{B_r(0)}$ and $\varphi \equiv 0$ on $\overline{B_R(0)}^c$. Such function φ is called a bump function for $\overline{B_r(0)}$. Details on the construction of such functions can be found in [30, Lemma 2.22].

Now we are ready for the main result of this chapter. This result was established in [25, Theorem 2.1] for a certain kind of diffusions on a compact subset of Euclidean space. We present our own proof for the kind of diffusions that we are considering.

Theorem 3.1. *Let L be an operator of the form (2.1), and let m be a measure on $\mathbb{R}^n$. Then m is a reversible measure for L if and only if m is quasi-invariant under the group $\{S_v\}_{v\in\mathbb{R}^n}$ of all translations with cocycle defined as*

$$\Lambda(v,x) = 2\int_0^1 b(S_{tv}(x)) \cdot v\, dt. \tag{3.13}$$

Proof. Let us take $g \in C_c^\infty(\mathbb{R}^n)$ and $v \in \mathbb{R}^n$ arbitrary. Integrating both sides of (3.5) with respect to a measure m we get

$$\int (b(x) \cdot v)g_t(x)m(dx) + \frac{1}{2}\int v \cdot \nabla g_t(x)m(dx) = -\frac{1}{2}\int \frac{d}{dt}g_t(x)m(dx).$$

We are going to use this equation to get another, very useful one. For this, consider a ball $B_r(0)$ big enough to contain the supports of all the functions g_t, for $0 \leq t \leq 1$ (see Lemma 3.1 for the definition of g_t), and take φ to be a bump function for $\overline{B_r(0)}$. Now define the function $f(x) := (x \cdot v)\varphi(x)$. We may regard this kind of function f as a "truncated polynomial" of first degree. Note that for $x \in B_r(0)$ we have $f(x) = x \cdot v$; and by simple calculations we can also see that $\nabla f(x) = v$ and $Lf(x) = b(x) \cdot v$. For functions f of the kind just defined the equation above can be given the following form:

$$\int (Lf)(x)g_t(x)m(dx) + \frac{1}{2}\int \nabla f(x) \cdot \nabla g_t(x)m(dx) = -\frac{1}{2}\frac{d}{dt}\int g_t(x)m(dx). \tag{3.14}$$

We know that $g_t \in C_c^\infty(\mathbb{R}^n)$ by Lemma 3.1.

If m is a reversible measure for L, then the left-hand side of (3.14) vanishes by (3.9). Therefore $\int g_t(x)m(dx)$ is constant in t and in particular $\int g_1(x)m(dx) = \int g_0(x)m(dx)$, or equivalently

$$\int g(S_{-v}(x)) \exp\{-\Lambda(v, S_{-v}(x))\}m(dx) = \int g(x)m(dx).$$

Since $g \in C_c^\infty(\mathbb{R}^n)$ and $v \in \mathbb{R}^n$ were chosen arbitrarily, we see that (3.12) is valid, and therefore m is quasi-invariant under S with the desired cocycle.

Conversely, assume that the measure m is quasi-invariant under S with the given cocycle Λ. Fix an arbitrary $g \in C_c^\infty(\mathbb{R}^n)$, and consider equation (3.14) for this fixed function g and for "truncated polynomials" f of the kind defined a few lines above, *i.e.* $f(x) = (x \cdot v)\varphi(x)$ for some $v \in \mathbb{R}^n$. It is clear that the right-hand side of (3.14) vanishes (see Lemma 3.1 and equation (3.12)). Thus (3.9) holds in this case.

Next we note that, for fixed $g \in C_c^\infty(\mathbb{R}^n)$, the set of functions $f \in C_c^\infty(\mathbb{R}^n)$ that satisfy equation (3.9) is closed under multiplication. To see this, assume that $f_1, f_2 \in C_c^\infty(\mathbb{R}^n)$ satisfy (3.9). Then using (3.10) with $f_1 f_2$ replacing f we have

$$\begin{aligned}
\int (Lf_1f_2)(x)g(x)m(dx) = \\
&= \int (Lf_1)(x)f_2 g(x)m(dx) + \int f_1(x)(Lf_2)(x)g(x)m(dx) \\
&\qquad + \int [\nabla f_1(x) \cdot \nabla f_2(x)]g(x)m(dx) \\
&= -\frac{1}{2}\int [\nabla f_1(x) \cdot \nabla (f_2 g)(x)]m(dx) - \frac{1}{2}\int [\nabla f_2(x) \cdot \nabla (f_1 g)(x)]m(dx) \\
&\qquad + \int [\nabla f_1(x) \cdot \nabla f_2(x)]g(x)m(dx) \\
&= -\frac{1}{2}\int f_2(x)[\nabla f_1(x) \cdot \nabla g(x)]m(dx) - \frac{1}{2}\int f_1(x)[\nabla f_2(x) \cdot \nabla g(x)]m(dx) \\
&= -\frac{1}{2}\int [\nabla (f_1 f_2)(x) \cdot \nabla g(x)]m(dx),
\end{aligned}$$

so it is clear that the product $f_1 f_2$ also satisfies (3.9).

Since (3.9) holds for functions of the form $f(x) = (x \cdot v)\varphi(x)$ for some $v \in \mathbb{R}^n$, an inductive argument shows that (3.9) can be extended to all functions $f(x) = (x \cdot v^1) \cdots (x \cdot v^k)\varphi(x)$ with $v^1, \ldots, v^k \in \mathbb{R}^n$. Using the linearity in f of (3.9) it follows that this expression must be true for all "truncated polynomials" $f(x)$ of arbitrary degree. A suitable approximation procedure (see *e.g.* [22, Appendix 7]) shows that (3.9) is valid for all $f \in C_c^\infty(\mathbb{R}^n)$. Since $g \in C_c^\infty(\mathbb{R}^n)$ was fixed arbitrarily, this establishes the reversibility of m. ∎

Remark. We have established that for n-dimensional Brownian motion with drift, a measure is reversible if and only if it is quasi-invariant under the group of all translations, and the proof clearly relies on calculus in $\mathbb{R}^n$. For other kinds of diffusions, with more general state spaces, it is possible to use similar ideas and establish equivalence between reversibility and quasi-invariance under certain suitably defined groups of transformations on the state space of the process, see for instance [24], [25]. In those cases, the methods are analogous to those used in this section, but the calculus operations have to be defined according to the nature of the state space being used.

Corollary 3.2. *If the operator L as in* (2.1) *has a reversible measure, then the expression defined in* (3.13) *is a cocycle.*

PROOF. Just read the first few lines of the proof of Theorem 3.1. ∎

The last result of this chapter is a partial converse to Theorem 3.1. This result seems to be new.

Theorem 3.2. *Let $b : \mathbb{R}^n \to \mathbb{R}^n$ be a smooth map, and define Λ as in* (3.13). *Assume that Λ satisfies the cocycle identity:*

$$\Lambda(u+v, x) = \Lambda(u, x+v) + \Lambda(v, x) \qquad \forall u, v \in \mathbb{R}^n, \forall x \in \mathbb{R}^n. \tag{3.15}$$

Then b has a potential, i.e. *there is a function $F : \mathbb{R}^n \to \mathbb{R}$ (necessarily smooth) such that $b = \nabla F$.*

PROOF. Note that it is enough to prove that

$$\langle \partial_v b(x), u \rangle = \langle \partial_u b(x), v \rangle \qquad \forall u, v \in \mathbb{R}^n, \forall x \in \mathbb{R}^n. \tag{3.16}$$

For if (3.16) holds, then taking $u = e_i$, $v = e_j$ we get $\partial_j b_i(x) = \partial_i b_j(x)$, and this condition is sufficient for b to have a potential (because the domain $\mathbb{R}^n$ is simply connected, see [44, Proposition 14.3.4]). Now let us establish (3.16). From the cocycle identity (3.15) we get

$$\Lambda(u, x+v) + \Lambda(v, x) = \Lambda(v, x+u) + \Lambda(u, x),$$

or equivalently

$$\Lambda(u, x+v) - \Lambda(u, x) = \Lambda(v, x+u) - \Lambda(v, x).$$

Then using the definition of Λ we get

$$\int_0^1 \langle b(x+v+tu) - b(x+tu), u \rangle dt = \int_0^1 \langle b(x+u+tv) - b(x+tv), v \rangle dt.$$

Replace u by δu and v by εv, for some $\delta, \varepsilon > 0$. Then using the last equation we get

$$\int_0^1 \langle \frac{b(x+\varepsilon v+\delta tu)-b(x+\delta tu)}{\varepsilon}, u\rangle dt = \int_0^1 \langle \frac{b(x+\delta u+\varepsilon tv)-b(x+\varepsilon tv)}{\delta}, v\rangle dt.$$

Letting first $\delta \to 0$ and then $\varepsilon \to 0$ we get

$$\int_0^1 \langle \partial_v b(x), u\rangle dt = \int_0^1 \langle \partial_u b(x), v\rangle dt,$$

from which (3.16) follows clearly. ■

Chapter 4

Reversibility for the Ornstein-Uhlenbeck process on an infinite-dimensional Hilbert space

4.1 Statement of the problem

Let us bring our consideration to the case where the state space is an infinite dimensional separable Hilbert space H. Consider the process $X = (X_t)_{t \geq 0}$ that solves the following linear stochastic differential equation

$$\begin{cases} dX_t = AX_t dt + dW_t, \\ X_0 = x. \end{cases}$$

We assume that the operator $A : \mathcal{D}(A) \subset H \to H$ generates a strongly continuous semigroup $\{T_t\}$ on H, and the process W is a standard cylindrical Wiener process on H, *i.e.* its covariance operator is given by the inner product on H. The reader is referred to the monographs [18], [19] of G.Da Prato and J.Zabczyk for accounts of the meaning of stochastic equations of the kind we mentioned above, as well as their solutions. The particular setting we have described is a special case of that presented in [15]. We say that X is an infinite dimensional Ornstein-Uhlenbeck process.

Although we do not focus on applications, we would like to mention that the kind of processes described above, when considered in finite dimensions, have become interesting because of their various practical uses, for instance in fluid mechanics and mathematical finance. A specific example that motivates the study of infinite dimensional Ornstein-Uhlenbeck processes is a stochastic model of neural response proposed by J. Walsh, where H is taken to be some L^2 space, see [38] and references therein.

We use the notation $\langle \cdot, \cdot \rangle_H$ for the inner product of H. We need to define a special kind of functions on H, the cylindrical functions. Let $D \subset H^*$ be a dense linear subspace of the dual H^*. By $\mathcal{F}C_c^\infty(D)$ we denote the space of all functions $f : H \to \mathbb{R}$ of the form

$$f(x) = \varphi(\ell_1(x), \ldots, \ell_n(x)) \quad \text{where } n \geq 1,\ \varphi \in C_c^\infty(\mathbb{R}^n),\ \ell_1, \ldots, \ell_n \in D \setminus \{0\}.$$

In case we take $D = H^*$, we denote this space as $\mathcal{F}C_c^\infty$ for simplicity.

The generator L of the process X, or more precisely of its corresponding semigroup, can be expressed formally as

$$Lf(x) = \frac{1}{2}\Delta f(x) + \langle Ax, \nabla f(x) \rangle_H, \qquad \text{for } f \in \mathcal{F}C_c^\infty. \tag{4.1}$$

The definitions of gradient and Laplacian for cylindrical functions are presented in the next section. However, we should point out that expression (4.1) only makes sense for $x \in \mathcal{D}(A)$; and this domain can be a proper subset of H. One way to overcome this inconvenience is to write the generator as follows:

$$Lf(x) = \frac{1}{2}\Delta f(x) + \langle x, A^* \nabla f(x) \rangle_H, \qquad \text{for } f \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*)). \tag{4.1$'$}$$

The advantage is that (4.1$'$) is an expression that makes sense for all $x \in H$. The price we pay for this advantage is that now the operator L is restricted to functions from the space $\mathcal{F}C_c^\infty(\mathcal{D}(A^*))$.

Next we recall the definitions of invariant measure and reversible measure for an operator L of the form (4.1$'$). First we specify the allowable measures we deal with. In the rest of this work, by "measure on H" we will always mean a Borel measure μ on H satisfying the following condition:

$$\int_H e^{\langle x,h \rangle_H} \mu(dx) < \infty, \qquad \forall h \in H. \tag{4.2}$$

We impose this condition in order that the integrals we are going to deal with be finite. Also, to avoid trivialities we exclude the zero measure from our discussion.

We say that μ is an invariant measure for L when

$$\int_H Lf(x) \mu(dx) = 0 \qquad \forall f \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*)). \tag{4.3}$$

The measure μ is reversible (or symmetrizing) for L if

$$\int_H Lf(x)\, g(x)\, \mu(dx) = \int_H f(x)\, Lg(x)\, \mu(dx) \quad \forall f, g \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*)). \tag{4.4}$$

The purpose of this chapter is to establish conditions for operators of the kind (4.1′) to be symmetrizable in the sense of (4.4). Based on the intuition that comes from the finite dimensional case, we would like to establish that reversibility is equivalent to quasi-invariance with respect to certain transformation group. In the finite dimensional setting it is the group of all translations that makes this equivalence work. However, in the infinite dimensional case, it is well known that, except for the zero measure, no measure is quasi-invariant with respect to all translations, see [43, p. 111]. We may, however, look for quasi-invariance under some special set of translations, say the group of shifts by vectors from some proper subspace of H.

Let V be a linear subspace of H. For simplicity, we will say that a measure μ on H is V-quasi-invariant when it is quasi-invariant under the group of all translations by vectors from V, *i.e.* under $\{S_v\}_{v \in V}$.

We want to prove the following

Conjecture. *A measure μ on H is symmetrizing for L if and only if it is $\mathcal{D}(A^*)$-quasi-invariant with cocycle*

$$\Lambda(v,x) = 2\int_0^1 \langle x + tv, A^*v\rangle_H \, dt. \tag{4.5}$$

Remark. We conjectured this form of the cocycle based on the analogous expression for the finite dimensional case. In the present case, however, it is easy to intergate out expression (4.5) to get a more useful equivalent form:

$$\Lambda(v,x) = 2\langle x, A^*v\rangle_H + \langle v, A^*v\rangle_H. \tag{4.6}$$

4.2 Technicalities

First we define some calculus operations with the cylindrical functions on H. Recall that for a given dense linear subspace D of the dual H^*, we say that $f \in \mathcal{F}C_c^\infty(D)$ when the function $f : H \to \mathbb{R}$ can be written in the form

$$f(x) = \varphi(\ell_1(x), \ldots, \ell_n(x)) \quad \text{where } n \geq 1,\ \varphi \in C_c^\infty(\mathbb{R}^n),\ \ell_1, \ldots, \ell_n \in D \setminus \{0\}.$$

For these kinds of functions we define the notions of "gradient" $\nabla f(x)$ and "Laplacian" $\Delta f(x)$ as follows:

$$\nabla f(x) = \sum_{i=1}^n \partial_i \varphi(\ell_1(x), \ldots, \ell_n(x))\ell_i, \tag{4.7}$$

$$\Delta f(x) = \sum_{i,j=1}^n \partial_i \partial_j \varphi(\ell_1(x), \ldots, \ell_n(x))\langle \ell_i, \ell_j\rangle_H. \tag{4.8}$$

These definitions are independent of the representation of the function f in terms of φ and $\ell_1, \ldots, \ell_n$, see [38, Section 5].

Remark. In this chapter we often identify H^* with H via $\ell(x) = \langle \ell, x\rangle_H$.

The next proposition says that the gradient and the Laplacian just defined behave similarly as their finite dimensional analogs with respect to products.

Proposition 4.1. *Let D be a dense linear subspace of H^*. For any pair of functions $f, g \in \mathcal{F}C_c^\infty(D)$ the following identities are satisfied:*

$$\nabla(fg) = f\nabla g + g\nabla f,$$

$$\Delta(fg) = f\Delta g + g\Delta f + \langle \nabla f, \nabla g\rangle_H.$$

PROOF. Let $f, g \in \mathcal{F}C_c^\infty(D)$ be arbitrary. Without loss of generality we can consider these two functions depending on the same "coordinates" $\ell_1, \dots, \ell_n$, so we may write $f(x) = \varphi(\ell_1(x), \dots, \ell_n(x))$ and $g(x) = \psi(\ell_1(x), \dots, \ell_n(x))$, where $\varphi, \psi \in C^\infty(\mathbb{R}^n)$. For the next computation we will use the shorthand $\boldsymbol{x} = (\ell_1(x), \dots, \ell_n(x))$ so that the expressions do not get too long. Then we have

$$\begin{aligned}
\nabla(fg)(x) &= \sum_{i=1}^n \partial_i(\varphi\psi)(\ell_1(x), \dots, \ell_n(x))\ell_i \\
&= \sum_{i=1}^n \left[\varphi(\boldsymbol{x})\partial_i\psi(\boldsymbol{x}) + \psi(\boldsymbol{x})\partial_i\varphi(\boldsymbol{x})\right]\ell_i \\
&= \varphi(\boldsymbol{x})\sum_{i=1}^n \partial_i\psi(\boldsymbol{x})\ell_i + \psi(\boldsymbol{x})\sum_{i=1}^n \partial_i\varphi(\boldsymbol{x})\ell_i \\
&= f(x)\nabla g(x) + g(x)\nabla f(x).
\end{aligned}$$

Similar computations show the second identity. ■

Just as in the finite dimensional case, every reversible measure is also invariant. This is established in the next proposition.

Proposition 4.2. *Let μ be a measure on H. If μ is reversible for L, then it is also invariant.*

PROOF. Let μ be a reversible measure for L, and let $f \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*))$ be arbitrary. Write $f(x) = \varphi(\ell_1(x), \dots, \ell_n(x))$ with $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\ell_1, \dots, \ell_n \in \mathcal{D}(A^*)$. Let $B_r(0) \subset \mathbb{R}^n$ be an open ball big enough to contain the support of φ, and let $\psi \in C_c^\infty(\mathbb{R}^n)$ be a bump function for $\overline{B_r(0)}$. Define the function $g(x) = \psi(\ell_1(x), \dots, \ell_n(x))$. Also, define the set

$$D_r := \{x \in H \,:\, (\ell_1(x), \dots, \ell_n(x)) \in B_r(0)\}.$$

Note that f and Lf are identically zero on D_r^c. We have

$$\begin{aligned}\int_H Lf(x)\mu(dx) &= \int_{D_r} Lf(x)\mu(dx) = \int_{D_r} Lf(x)\, g(x)\mu(dx)\\ &= \int_H Lf(x)\, g(x)\mu(dx) = \int_H f(x)\, Lg(x)\mu(dx)\\ &= \int_{D_r} f(x)\, Lg(x)\mu(dx) = 0,\end{aligned}$$

where in the last step we use the fact that $Lg = 0$ on D_r. ■

The following technical lemma will be useful for the proof that reversibility implies quasi-invariance.

Lemma 4.1. *Let Λ be given by (4.5). Fix an arbitrary $v \in \mathcal{D}(A^*)$. For any given $g \in \mathcal{F}C_c^\infty$ and $t \in \mathbb{R}$, define*

$$g_t(x) = g(S_{-tv}(x)) = g(x - tv).$$

Then $g_t \in \mathcal{F}C_c^\infty$ for all $t \in \mathbb{R}$, and the following holds:

$$\frac{d}{dt} g_t(x) = -\langle v, \nabla g_t(x)\rangle_H. \tag{4.9}$$

Furthermore, writing $\widetilde{g}_t(x) := g_t(x)\exp\{-\Lambda(tv, S_{-tv}(x))\}$ we have

$$\frac{d}{dt}\widetilde{g}_t(x) = -\langle v, \nabla \widetilde{g}_t(x)\rangle_H - 2\widetilde{g}_t(x)\langle x, A^*v\rangle_H. \tag{4.10}$$

PROOF. To establish (4.9) we just need to use the definition of g_t and the chain rule. Now let us prove (4.10). Using the expression (4.6) for the cocycle Λ we get

$$\Lambda(tv, x - tv) = 2t\langle x, A^*v\rangle_H - t^2\langle v, A^*v\rangle_H. \tag{4.11}$$

This is useful for computing derivatives of $\Lambda(tv, x - tv)$, as will be needed ahead. Next we compute the gradient $\nabla \widetilde{g}_t(x)$. We have

$$\begin{aligned}\nabla \widetilde{g}_t(x) &= e^{-\Lambda(tv,S_{-tv}(x))}\nabla g_t(x) + g_t(x)\nabla e^{-\Lambda(tv,S_{-tv}(x))}\\ &= e^{-\Lambda(tv,S_{-tv}(x))}\nabla g_t(x) - g_t(x)e^{-\Lambda(tv,S_{-tv}(x))}(2tA^*v)\\ &= e^{-\Lambda(tv,S_{-tv}(x))}\nabla g_t(x) - 2t\widetilde{g}_t(x)A^*v.\end{aligned} \tag{4.12}$$

Now we use the chain rule and (4.11) to compute the derivative of $\widetilde{g}_t(x)$ with respect to t. We have

$$\begin{aligned}\frac{d}{dt}\widetilde{g}_t(x) &= e^{-\Lambda(tv,S_{-tv}(x))}\frac{d}{dt}g_t(x) - g_t(x)e^{-\Lambda(tv,S_{-tv}(x))}\frac{d}{dt}\Lambda(tv, x - tv)\\ &= -e^{-\Lambda(tv,S_{-tv}(x))}\langle v, \nabla g_t(x)\rangle_H - \widetilde{g}_t(x)\left(2\langle x, A^*v\rangle_H - 2t\langle v, A^*v\rangle_H\right)\\ &= -\langle v, e^{-\Lambda(tv,S_{-tv}(x))} - 2t\widetilde{g}_t(x)A^*v\rangle_H - 2\widetilde{g}_t(x)\langle x, A^*v\rangle_H.\end{aligned}$$

The proof is finished using (4.12) in the first term of the last expression. ■

Next we prove a technical lemma which gives an equivalent way to prove reversibility of L, this is analogous to Lemma 3.2 of the previous chapter.

Lemma 4.2. *Let L be as in* (4.1′), *and let μ be a measure on H. Then μ is reversible for L if and only if*

$$\int (Lf)(x)g(x)\mu(dx) = -\frac{1}{2}\int \langle \nabla f(x), \nabla g(x)\rangle_H \mu(dx) \quad \forall f, g \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*)). \tag{4.13}$$

PROOF. Using the identities of Proposition 4.1 we can see that

$$(Lf)(x)g(x) + f(x)(Lg)(x) - L(fg)(x) = -\langle \nabla f(x), \nabla g(x)\rangle_H. \tag{4.14}$$

Assume that the measure μ is reversible for L. In particular, this implies that μ is an invariant measure for L. Now take $f, g \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*))$, and note that their product fg is also in $\mathcal{F}C_c^\infty(\mathcal{D}(A^*))$, so $\int L(fg)(x)\mu(dx) = 0$. Then integrating both sides of (4.14) with respect to μ we get (4.13).

Conversely, assume that (4.13) holds. Since the right hand side of this equation is symmetric in f and g, this implies that the left hand side is also symmetric; *i.e.* (4.4) holds. ■

We will need to use a specific kind of truncation functions that we describe next. Recall that for any given $r > 0$, there is a smooth function $\varphi : \mathbb{R} \to \mathbb{R}$ which is identically equal to one on $[-r, r]$ and vanishes outside some bigger closed interval. This is what we called a bump function for $[-r, r]$. We use this function to define a new one $\Phi(x) := x\varphi(x)$, $x \in \mathbb{R}$. It is clear that Φ is also smooth and compactly supported, and it has the nice property that it equals the identity function when restricted to $[-r, r]$. We will call the function Φ a truncation function for $[-r, r]$.

PROOF OF THE CONJECTURE: REVERSIBILITY IMPLIES QUASI-INVARIANCE. We assume here that the measure μ is reversible for the operator L. Fix $v \in \mathcal{D}(A^*)$ arbitrary. Recall from the previous chapter, equation (3.12), that in order to establish quasi-invariance with the desired cocycle we must prove that for all functions g in some measure determining class we have

$$\int_H g(S_{-v}(x))e^{-\Lambda(v, S_{-v}(x))}\mu(dx) = \int_H g(x)\mu(dx). \tag{4.15}$$

Our strategy will be to prove that for any given $g \in \mathcal{F}C_c^\infty$,

$$Z(t) := \int g_t(x)e^{-\Lambda(tv, x-tv)}\mu(dx) = \int \widetilde{g}_t(x)\mu(dx)$$

is a constant function of $t \in \mathbb{R}$. Once this is established we get (4.15) setting $Z(1) = Z(0)$.

Fix an arbitrary $g \in \mathcal{F}C_c^\infty$. Using the result of Lemma 4.1 we have

$$\begin{aligned} Z'(t) &= \int \frac{d}{dt}\widetilde{g}_t(x)\mu(dx) \\ &= -\int \langle v, \nabla \widetilde{g}_t(x)\rangle_H \mu(dx) - 2\int \widetilde{g}_t(x)\langle x, A^*v\rangle_H \mu(dx). \end{aligned}$$

Let us write $g(x) = \psi(\ell_1(x), \ldots, \ell_n(x))$, where $\psi \in C_c^\infty(\mathbb{R}^n)$ and $\ell_1, \ldots, \ell_n \in H^*$. Defining the function

$$\psi_t(y_1, \ldots, y_n) = \psi(y_1 - t\ell_1(v), \ldots, y_n - t\ell_n(v))$$

we can write $g_t(x) = \psi_t(\ell_1(x), \ldots, \ell_n(x))$, and note that $\psi_t \in C_c^\infty(\mathbb{R}^n)$ for all $t \in \mathbb{R}$. Take an open ball $B_r^n(0) \subset \mathbb{R}^n$ (the superscript is used to emphasize the dimension) big enough to contain the supports of all the functions ψ_t for $0 \le t \le 1$. Let $\Phi \in C_c^\infty(\mathbb{R})$ be a truncation function for $[-r, r]$, and define the function $f(x) = \Phi(\langle x, v\rangle_H)$. It is clear that $f \in \mathcal{F}C_c^\infty(\mathcal{D}(A^*))$. Define the set

$$D_r := \{x \in H \, : \, (\ell_1(x), \ldots, \ell_n(x), \langle x, v\rangle_H) \in B_r^n(0) \times [-r, r]\}.$$

Note that on the complement of D_r either f vanishes or all the functions g_t, $0 \le t \le 1$ vanish. On the other hand, for $x \in D_r$ we have $f(x) = \langle x, v\rangle_H$, $\nabla f(x) = v$ and $Lf(x) = \langle x, A^*v\rangle_H$. Then the last line in the computation of $Z'(t)$ above yields

$$Z'(t) = -\int \langle \nabla f(x), \nabla \widetilde{g}_t(x)\rangle_H \mu(dx) - 2\int \widetilde{g}_t(x) Lf(x)\mu(dx)$$

Therefore the proof of this part is completed if we verify that the functions $\widetilde{g}_t$ belong to a class for which (4.13) still holds (for every $f \in \mathcal{F}C_c^\infty$). ■

Conclusions

Here are the major highlights of this thesis.

In Chapter 2 we proved Kolmogorov's criterion of reversibility establishing that an operator of the form (2.1) is reversible if and only if its drift term is of gradient form. This criterion was established in [26] for diffusions on a compact Riemannian manifold; we relied on this reference to provide our own proof of this criterion for the case of "Brownian motion with drift", where the state space is all of $\mathbb{R}^n$ (a non-compact space). As immediate consequences of this criterion we saw that in the reversible case the reversible measure is unique (up to a multiplicative constant), and also that in dimension one reversibility is always possible.

In Chapter 3 we proved a second characterization of reversibility in finite dimensions. This characterization tests measures on $\mathbb{R}^n$ (the state space of the process) and establishes that reversibility is equivalent to quasi-invariance under the group of all translations with a cocycle given in terms of the drift coefficients. The proof of this theorem is more elaborate than the proof of the criterion of Kolmogorov, but it is worth the effort since this new criterion seems to be suited for diffusions in more general state spaces, as the papers [24], [25] of K.Handa suggested. The paper [25] is of particular importance in this chapter; we borrowed some of its ideas to construct our own proof of the characterization of reversibility in terms of quasi-invariance.

While in finite dimensions the characterizations of reversibility work nicely, in infinite dimensions some difficulties arise. As a consequence, our analysis of the link between reversibility and quasi-invariance for the infinite dimensional Ornstein-Uhlenbeck process is incomplete, and the goal of fully characterizing reversible measures for this case still awaits solution. However, we do make significant progress in this direction, and the approach we use in this thesis seems to show some promise.

Another major contribution of this thesis is the extensive list of meaningful references provided in the bibliography. This should furnish those who follow with a well-paved road for future research in this area.

Bibliography

[1] J. AARONSON, *An Introduction to Infinite Ergodic Theory.* AMS Mathematical Surveys and Monographs, No. 50, Providence, Rhode Island, 1997.

[2] D.R. ADAMS and L.I. HEDBERG, *Function spaces and potential theory.* Springer Verlag, Berlin, 1996.

[3] S. AGMON, *Lectures on Elliptic Boundary Value Problems* . Van Nostrand Mathematical Studies, No. 2. Van Nostrand, Princeton, New Jersey, 1965.

[4] S. ALBEVERIO, V.I. BOGACHEV and M. RÖCKNER, On uniqueness of invariant measures for finite and infinite dimensional diffusions. *Comm. Pure Appl. Math.* **52** (1999), 325-362.

[5] S. ALBEVERIO, Y.G. KONDRATIEV and M. RÖCKNER, Ergodicity for the stochastic dynamics of quasi-invariant measures with applications to Gibbs states. *J. Funct. Anal.* **149** (1997), 415-469.

[6] R.F. BASS, *Diffusions and Elliptic Operators.* Springer Verlag, New York, 1998.

[7] H. BAUER, *Mesure and Integration Theory.* Walter de Gruyter, Berlin - New York, 2001.

[8] V.I. BOGACHEV, N.V. KRYLOV and M. RÖCKNER, Regularity of invariant measures: the case of non-constant diffusion part. *J. Funct. Anal.* **138** (1996), 223-242.

[9] V.I. BOGACHEV, N.V. KRYLOV and M. RÖCKNER, Elliptic regularity and essential self-adjointness of Dirichlet operators on $\mathbb{R}^n$. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **24** (1997), no. 3, 451-461.

[10] V.I. BOGACHEV and M. RÖCKNER, Regularity of invariant measures on finite and infinite dimensional spaces and applications. *J. Funct. Anal.* **133** (1995), 168-223.

[11] V.I. BOGACHEV and M. RÖCKNER, A generalization of Khasminskii's theorem on the existence of invariant measures for locally integrable drifts. *Theory Probab. Appl.* **45** (2000), no. 3.

[12] V.I. BOGACHEV, M. RÖCKNER and F.Y. WANG, Elliptic equations for invariant measures on finite and infinite dimensional manifolds. *J. Math. Pure Appl.* **80** (2001), 177-221.

[13] V.I. BOGACHEV, M. RÖCKNER and F.Y. WANG, Elliptic equations for invariant measures on Riemannian manifolds: existence and regularity of solutions. *C.R. Acad. Sci. Paris* **332** (2001), 333-338.

[14] V.I. BOGACHEV, M. RÖCKNER and T.S. ZHANG, Existence and uniqueness of invariant measures: an approach via sectorial forms. *Appl. Math. Optim.* **41** (2000), 87-109.

[15] A. CHOJNOWSKA-MICHALIK and B. GOLDYS, Symmetric Ornstein-Uhlenbeck semigroups and their generators. *Probab. Theory Relat. Fields* **124** (2002), 459-486.

[16] K.L. CHUNG, *Lectures from Markov processes to Brownian motion.* Springer Verlag, New York, 1982.

[17] J.B. CONWAY, *A course in functional analysis.* 2nd ed. Springer Verlag, New York, 1990.

[18] G. DA PRATO and J. ZABCZYK, *Stochastic equations in infinite dimensions.* Cambridge University Press, Cambridge - New York, 1992.

[19] G. DA PRATO and J. ZABCZYK, *Ergodicity for infinite dimensional systems* . Cambridge University Press, Cambridge - New York, 1996.

[20] R.M. DUDLEY, *Real Analysis and Probability,* Cambridge University Press, Cambridge - New York, 2002.

[21] K-J. ENGEL and R. NAGEL, *One-parameter semigroups for linear evolution equations.* Springer Verlag, New York, 2000.

[22] S. N. ETHIER and T. KURTZ, *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.

[23] K. HANDA, Quasi-invariant measures and their characterization by conditional probabilities. *Bull. Sci. Math.* **125** (2001), 583-604.

[24] K. HANDA, Quasi-invariance and reversibility in the Fleming-Viot process. *Probab. Theory Relat. Fields* **122** (2002), 545-566.

[25] K. HANDA, Reversible distributions of multi-allelic Guillespie-Sato diffusion models. *Ann. I. H. Poincaré - PR*, **40** (2004), 569-597.

[26] N. IKEDA and S. WATANABE, *Stochastic Differential Equations and Diffusion Processes.* 2nd ed. North-Holland, Amsterdam, 1989.

[27] J. JOST, *Partial Differential Equations.* Springer Verlag, New York, 2002.

[28] I. KARATZAS and S.E. SHREVE, *Brownian motion and stochastic calculus.* 2nd ed. Springer Verlag, New York, 1991.

[29] A.B. KARAZISHVILI, *Transformation Groups and Invariant Measures.* World Scientific, Singapore - New Jersey, 1998.

[30] J.M. LEE, *Introduction to Smooth Manifolds.* Springer Verlag, New York, 2003.

[31] T.M. LIGGETT, *Interacting Particle Systems.* Springer Verlag, New York, 1985.

[32] E. NELSON, The adjoint Markoff process. *Duke Math. J.* **25** (1958), 671-690.

[33] A. PAZY, *Semigroups of linear operators and applications to partial differential equations.* Springer Verlag, New York, 1983.

[34] R.G. PINSKY, *Positive harmonic functions and diffusion.* Cambridge University Press, Cambridge, 1995.

[35] D. REVUZ and M. YOR, *Continuous Martingales and Brownian Motion.* 3rd ed. Springer Verlag, Berlin, 1999.

[36] L.C.G. ROGERS and D. WILLIAMS, *Diffusions, Markov processes, and Martingales, Vol. 1: Foundations.* 2nd ed. Cambridge University Press, Cambridge, 2000.

[37] B. SCHMULAND and W. SUN, A cocycle proof that reversible Fleming-Viot processes have uniform mutation. *C.R. Math. Rep. Acad. Sci. Canada* **24** (2002) 124-128.

[38] B. SCHMULAND, A Course on Infinite Dimensional Dirichlet Forms. Lecture notes presented at the Institute for Mathematical Sciences, Yonsei University in Seoul, 1995. Corrected version 1999.

[39] A.N. SHIRYAEV, *Probability.* 2nd ed. Springer Verlag, New York, 1996.

[40] A.V. Skorokhod, *Lectures on the Theory of Stochastic Processes.* VSP, Utrecht, and TBiMC, Kiev, 1996.

[41] D.W. Stroock, *Probability theory: an analytic view.* Rev. ed. Cambridge University Press, Cambridge, 1999.

[42] H. Weyl, The method of orthogonal projection in potential theory. *Duke Math. J.* **7** (1940), 411-444.

[43] Y. Yamasaki, *Measures on Infinite Dimensional Spaces.* World Scientific, Singapore - New Jersey, 1985.

[44] V.A. Zorich, *Mathematical Analysis II.* Springer Verlag, New York, 2004.