Stochastic Geometric Mechanics

Joan Andreu Lázaro Camí

UNIVERSIDAD DE ZARAGOZA
DEPARTAMENTO DE FÍSICA TEÓRICA

Programa de Doctorado en Física

Memoria presentada para aspirar al grado de Doctor
por la Universidad de Zaragoza

EDUARDO MARTÍNEZ FERNÁNDEZ, profesor titular del Departamento de Matemática Aplicada de la Universdiad de Zaragoza, y

JUAN PABLO ORTEGA LAHUERTA, *chargé de recherches (CNRS)* del *Département de Mathématiques de l'Université de Franche-Comté*, Besançon,

CERTIFICAN:

Que la presente memoria ha sido realizada por Joan Andreu Lázaro Camí y dirigida por nosotros.

Zaragoza, 10 de septiembre de 2008

Fdo. Eduardo Martínez Fernández

Fdo. Juan Pablo Ortega Lahuerta

Contents

Agradecimientos

Esta tesis no habría sido posible sin la colaboración de muchas personas. Algunas de ellas merecen ser mencionadas explícitamente.

En primer lugar, me gustaría agradecer a Eduardo Martínez y Pepín Cariñena sus consejos, su comprensión, su confianza en mí y toda la libertad que me han dado para escoger mi camino.

Agradezco al Departamento de Física Teórica de la Universidad de Zaragoza y al Département de Mathématiques de l'Université de Franche-Comté el haber puesto a mi disposición todos los medios técnicos para que esta tesis saliera adelante.

Le doy las gracias a Pablo Díaz por haber hecho que mi adaptación a Zaragoza haya sido tan fácil y enriquecedora, dentro y fuera de nuestro despacho.

Finalmente, debo agradecerle a Juan Pablo Ortega todo el trabajo y empeño que ha puesto en esta tesis. Esta tesis no es menos suya que mía. En lo personal, le doy las gracias por haberme abierto las puertas de su casa y dejarme ser su amigo, por mostrarme lo que es el amor a su familia, a María Pilar y Paula. Yo también los quiero infinitamente a ellos.

Preface

This doctoral thesis lies within the framework of stochastic differential geometry and is structured in four chapters. Its goal is conveying to the geometric mechanics community the wealth of global tools available to handle mechanical problems that contain a stochastic component and that do not seem to have been exploited to the full extent of their potential. After an introductory chapter aimed at recalling the main basics of stochastic calculus both in Euclidean spaces and manifolds, the new contributions of the thesis are in the subsequent chapters.

In Chapter 2, we use the global stochastic analysis tools introduced by P. A. Meyer and L. Schwartz to write down a stochastic generalization of the Hamilton equations on a Poisson manifold that, for exact symplectic manifolds, are characterized by a natural critical action principle similar to the one encountered in classical mechanics. Several features and examples in relation with the solution semimartingales of these equations are presented. We extend then some aspects of the Hamilton-Jacobi theory to the category of stochastic Hamiltonian dynamical systems. More specifically, we show that the stochastic action satisfies the Hamilton-Jacobi equation when, as in the classical situation, it is written as a function of the configuration space using a regular Lagrangian submanifold. Additionally, we will use a variation of the Hamilton-Jacobi equation to characterize the generating functions of one-parameter groups of symplectomorphisms that allow to rewrite a given stochastic Hamiltonian system in a form whose solutions are very easy to find; this result recovers in the stochastic context the classical solution method by *reduction to the equilibrium* of a Hamiltonian system.

In Chapter 3, we present reduction and reconstruction procedures for the solutions of symmetric stochastic differential equations, similar to those available for ordinary differential equations. Additionally, we use the local tangent-normal decomposition, available when the symmetry group is proper, to construct local skew-product splittings in a neighborhood of any point in the open and dense principal orbit type. The general methods introduced are then adapted to the Hamiltonian case, which is studied with special care and illustrated with several examples. The Hamiltonian category deserves a separate study since in that situation the presence of symmetries implies in most cases the existence of conservation laws, mathematically described via momentum maps, that should be taken into account in the analysis.

Finally, Chapter 4 proves a version for stochastic differential equations of the Lie-Scheffers Theorem. This result characterizes the existence of nonlinear superposition rules for the general solution of those equations in terms of the involution properties of the distribution generated by the vector fields that define it. When stated in the particular case of standard deterministic systems, our main theorem improves various aspects of the classical Lie-Scheffers result. We show that the stochastic analog of the classical Lie-Scheffers systems can be reduced to the study of Lie group valued stochastic Lie-Scheffers systems; those systems, as well as those taking values in homogeneous spaces are studied in detail. The developments are illustrated with several examples.

1
Preliminaries

The first chapter of this thesis aims at recalling the essential tools on stochastic calculus that we are going to use later on. Although the reader is supposed to be familiar with the basic concepts of probability theory, we have tried to gather most of the basic definitions and results about stochastic processes found in standard textbooks. The idea is, on the one hand, to make this thesis as self-contained as possible and, on the other, to introduce the references on stochastic processes and stochastic differential equations that are appropriate for our purposes. We hope that the readers interested in mechanics who already know very well its geometrical framework but who have not necessarily worked with stochastic calculus will find it useful.

The chapter is structured as follows: in Section 1.1 we recall the definition of a random variable, the different kinds of convergence of sequences of random variables, and the relations among them. We also introduce in this section one of the major concepts in probability theory, that of conditional expectation. In Section 1.2 we switch from random variables to stochastic processes. We present the most important and extensively studied process, the Brownian motion, which is a particular example of some processes playing a prominent role in the theory of stochastic integration, (local) martingales. In addition, other classical properties of processes, such as being Markov, are also presented. Section 1.3 is devoted to the cornerstone of the theory of stochastic processes, the (Itô) stochastic integral. The stochastic integral is a generalization of the Riemann-Stieltjes integral for processes whose paths are not of finite variation. We carefully introduce the Itô stochastic integral, the quadratic variation of a processes, and the Stratonovich integral, linked with the Itô integral by means of the quadratic variation of the process with respect to the semimartingale we integrate. Once stochastic integrals have been introduced, stochastic differential equations can be defined in terms of them and, therefore, we may talk about stochastic systems as those which evolve in time according to the solutions of a given stochastic differential equation. As a consequence of the stochastic integration, the Itô formula gives us the stochastic differential equation fulfilled by a smooth function when composed with a semimartingale. Finally, we show in Section 1.4 how to generalize the tools and concepts of Section 1.3 to the case of manifold valued semimartingales. As we will see, the geometric structures adapted to stochastic integration are the second order tangent and cotangent bundles of a manifold. More concretely, we will be able to integrate processes taking values on the second order cotangent bundle in the Itô sense and those taking values on the cotangent bundle in the Stratonovich's. This extended stochastic integral and the notion of Stratonovich and Schwartz operators will allow us to intrinsically define stochastic differential equations on manifolds.

A few words about notation before we start. Throughout this chapter the triple $(\Omega, \mathcal{F}, P)$ will denote a probability space, where $\mathcal{F}$ is a σ-algebra made out of subsets of Ω and $P : \mathcal{F} \to \mathbb{R}$ denotes the probability measure. Whenever that $\Omega = \mathbb{R}^d$ we will assume that $\mathcal{F} = \mathcal{B}(\mathbb{R}^d)$ is the Borel σ-algebra, that is, the smallest σ-algebra containing the open sets of the topology defined from the standard Euclidean distance. The Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$ will be denoted by λ. We will say that some property in Ω holds P-a.s. (or λ-a.s. in case $\Omega = \mathbb{R}^d$) if it holds except from those $\omega \in \Omega$ contained in a set of probability (measure) zero.

1.1 Random variables

Definition 1.1 *Let $(E, \mathcal{E})$ be a measurable space. A **random variable** is a measurable map $X : \Omega \to E$. The **law** of X is the probability measure P_X induced by X in $(E, \mathcal{E})$ by*

$$P_X(B) := P(\{\omega \in \Omega \mid X(\omega) \in B\}).$$

*In general one writes $P_X(B) = P(\{X \in B\})$ and one says the"probability that X sits in B ". The **σ-algebra** $\sigma(X) \subset \mathcal{F}$ **generated** by a random variable $X : \Omega \to (E, \mathcal{E})$ is the smallest σ-algebra that makes X measurable:*

$$\sigma(X) = \{A = X^{-1}(B) \in \mathcal{F} \mid B \in \mathcal{E}\}.$$

Classical laws are for example the uniform law, the binomial law, the geometric law, or the Poisson law among others. Very important for us is the following example.

Example 1.2 : Random variables with a density. Let μ and ν be two measures on a measurable space $(E, \mathcal{E})$ and suppose that μ is σ-finite, that is, there exists a sequence $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{E}$ such that $E = \cup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < \infty$ for any $n \in \mathbb{N}$. We say that ν is ***absolutely continuous*** with respect to μ if for any $A \in \mathcal{B}(\mathbb{R}^d)$, $\mu(A) = 0$ implies that $\nu(A) = 0$. In this context, the ***Radon-Nikodym Theorem*** states that ν is absolutely continuous with respect to μ if and only if there exists a μ-unique measurable function $f : E \to [0, \infty]$ such that for any $A \in \mathcal{E}$

$$\nu(A) = \int_E f d\mu$$

([K78, Chapter VII § 5], [R87, Theorem 6.9]). In concrete examples, $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and μ is the Lebesgue measure λ. If P_X is absolutely continuous with respect to the Lebesgue measure λ, then $P_X(A) = \int_A f d\lambda$. The function f is called the ***density of the law*** of X or the ***probability density function (pdf)*** of X.

Example 1.3 Some classical densities:

(i) Gaussian or normal law of mean $m \in \mathbb{R}$ and variance $\sigma^2 \in R_+$: usually denoted by $N(m, \sigma^2)$ is the most important law in probability theory. Its pdf is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}.$$

(ii) Uniform law on $[a, b]$: this is a law whose pdf is constant on the whole interval $[a, b]$. Therefore, $p(x) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(x)$.

Let $X : \Omega \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a real valued random variable. Its ***moment*** $E[X^p]$ of order p is defined by

$$E[X^p] := \int_\Omega X^p dP.$$

The case $p = 1$ corresponds to the ***mathematical expectation*** or ***expected value*** of X. $E[X^p]$ obviously exists provided that $\int_\Omega |X|^p dP < \infty$. If $X \in L^2(\Omega, \mathcal{F}, P)$ we define the ***variance*** of X by

$$\mathrm{Var}(X) := E\left[(X - E[X])^2\right] = E[X^2] - (E[X])^2.$$

The ***standard deviation*** is given by $\sigma_X = \sqrt{\mathrm{Var}(X)}$. More in general, if $X : \Omega \to (E, \mathcal{E})$ is a random variable taking values on an arbitrary measurable space and $f : E \to \mathbb{R}$ is measurable, then $f(X) \in L^1(\Omega, \mathcal{F}, P)$ if and only if $f \in L^1(E, \mathcal{E}, P_X)$ and the ***change of variables formula*** reads

$$E[f(X)] = \int_E f(x) P_X(dx).$$

In particular, if $E = \mathbb{R}^n$ and P_X is absolutely continuous with respect to the Lebesgue measure with density $g : \mathbb{R}^n \to \mathbb{R}_+$, then

$$E[f(X)] = \int_{\mathbb{R}^n} f(x) g(x) d\mu.$$

Example 1.4 Given a Gaussian variable X with law $N(m, \sigma^2)$ then

$$E[X] = \int_\mathbb{R} x \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}} dx = m,$$

$$Var[X] = E[(X - m)^2] = \int_\mathbb{R} (x - m)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}} dx = \sigma^2.$$

Independence

Given two events $A, B \in \mathcal{F}$, we say that they are independent whenever $P(A \cap B) = P(A)P(B)$. When $P(B) > 0$, this definition can be interpreted using the ***conditional probability*** induced by B and defined as

$$P(\cdot|B) = \frac{P(\cdot \cap B)}{P(B)}. \tag{1.1}$$

$(\Omega, \mathcal{F}, P(\cdot|B))$ is a new probability space whose measure yields the probability of the events in $\mathcal{F}$ once one knows that B has happened. The events A and B are independent when $P(A|B) = P(A)$. This motivates the following general definition:

Definition 1.5 *Let $\mathcal{B}_1, \ldots, \mathcal{B}_n \subset \mathcal{F}$ be sub σ-algebras of $\mathcal{F}$. We say that $\mathcal{B}_1, \ldots, \mathcal{B}_n$ are* ***independent*** *when*

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdots P(A_n),$$

for any $A_1 \in \mathcal{B}_1, \ldots, A_n \in \mathcal{B}_n$. Let $(E_i, \mathcal{E}_i)$, $i = 1, ..., n$, be a family of measurable spaces. We say that a family of random variables $X_1, \ldots, X_n$, $X_i : \Omega \longrightarrow (E_i, \mathcal{E}_i)$, are independent if the associated σ-algebras $\sigma(X_1), \ldots, \sigma(X_n)$ are independent.

The independence of a family of random variables $X_1, \ldots, X_n$ is equivalent to the condition

$$P(\{X_1 \in F_1\} \cap \cdots \cap \{X_n \in F_n\}) = P(\{X_1 \in F_1\}) \cdots P(\{X_n \in F_n\}),$$

for any $F_1 \in \mathcal{E}_1, \ldots, F_n \in \mathcal{E}_n$. In other words,

$$P_{(X_1,\ldots,X_n)} = P_{X_1} \otimes \cdots \otimes P_{X_n}.$$

In that case, given a family of measurable functions $f_i : (E_i, \mathcal{E}_i) \longrightarrow \mathbb{R}$:

$$E\left[\prod_{i=1}^{n} f_i(X_i)\right] = \prod_{i=1}^{n} E[f_i(X_i)].$$

Additionally, if each random variable X_i takes values in the real line and it is absolutely continuous with pdf p_i, then the random variable $(X_1, \ldots, X_n)$ is absolutely continuous with a pdf p given by

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p_i(x_i)$$

1.1.1 Convergence of random variables

Given a sequence of random variables $\{X_n\}_{n \in \mathbb{N}}$ in the probability space $(\Omega, \mathcal{F}, P)$ there are several notions of convergence that are relevant in the context of probability theory:

(i) Almost sure (a.s.) convergence: we say that $X_n \xrightarrow{a.s.} X$ whenever

$$P\left(\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\right) = 1.$$

(ii) L^p convergence: $X_n \xrightarrow{L^p} X$ whenever $\lim_{n\to\infty} E[|X_n - X|^p] = 0.$

(iii) Convergence in probability: $X_n \xrightarrow{P} X$ whenever for any $\epsilon > 0$:

$$\lim_{n\to\infty} P(\{|X_n - X| > \epsilon\}) = 0.$$

(iv) Convergence in law: consider $C_b^0(\mathbb{R}^d)$ the continuous bounded functions in $\mathbb{R}^d$ with the norm

$$\|\varphi\| = \sup_{x\in\mathbb{R}^d} |\varphi(x)|.$$

We say that a sequence $\{\mu_n\}_{n\in\mathbb{N}}$ of probability measures on $\mathbb{R}^d$ converges to μ if for any $\varphi \in C_b^0(\mathbb{R}^d)$

$$\int \varphi d\mu_n \stackrel{n\to\infty}{\longrightarrow} \int \varphi d\mu.$$

We say that a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ with values in $\mathbb{R}^d$ converges in law to the random variable X (we write $X_n \xrightarrow{l} X$) whenever $P_{X_n} \stackrel{n\to\infty}{\longrightarrow} P_X$. This amounts to saying that for any $\varphi \in C_b^0(\mathbb{R}^d)$:

$$E[\varphi(X_n)] \stackrel{n\to\infty}{\longrightarrow} E[\varphi(X)].$$

It is worth mentioning that if the random variables X_n have a density p_n, there exists a function p such that $p_n(x) \to p(x)$ λ-a.s. On the other hand, if there exists a function $q \geq 0$ in $L^1(\mathbb{R}^d, \mu)$ such that $|p_n(x)| \leq q(x)$ λ-a.s. and the limit $p(x) := \lim_{n\to\infty} p_n(x)$ exists λ-a.s., then $\{X_n\}_{n\in\mathbb{N}}$ converges in law to the law $p(x)d\lambda$. This is a straightforward consequence of the Dominated Convergence Theorem.

These four notions of convergence are not independent. For example, the L^p-convergence implies the convergence in probability. Indeed, applying Chebyshev's inequality yields

$$P(\{|X_n - X| > \varepsilon\}) \leq \frac{1}{\varepsilon^p} E[|X_n - X|^p].$$

On the other hand, almost sure convergence implies the convergence in probability. However, convergence in probability only implies the existence of a subsequence which converges almost surely. By the dominated convergence theorem, $X_n \xrightarrow{a.s.} X$ implies $X_n \xrightarrow{L^p} X$, $p \geq 1$, if the random variables X_n satisfy $|X_n| \leq Y$ a.s. for a fixed nonnegative random variable Y possessing a finite moment of order p. Finally, convergence in probability implies convergence in law, the reciprocal being also true when the limit is constant. See [GS01, Section 7.2] for a proof of theses implications.

Applications: We briefly recall two well known results about the convergence of the average of a sequence of independent random variables. Their proof can be found in [GS01, Section 5.10].

(i) (Strong) Law of Large Numbers: let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent random variables in $L^1(\Omega, \mathcal{F}, P)$ that share the same law. Then

$$\frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{a.s.} E[X_1].$$

(ii) Central Limit Theorem: let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent random variables in $L^2(\Omega, \mathcal{F}, P)$ that share the same law. Let $\sigma^2 = Var(X_1)$. Then

$$\frac{1}{\sqrt{n}}(X_1 + \cdots + X_n - nE[X_1]) \xrightarrow{l} N(0, \sigma^2).$$

1.1.2 Conditional expectations

We showed in the expression (1.1) how to construct out of a given event $B \in \mathcal{F}$ such that $P(B) > 0$ a new probability space $(\Omega, \mathcal{F}, P(\cdot|B))$ that allows us to measure the probability of the other events subjected to the *condition* that B has happened. Given a random variable $X \in L^1(\Omega, \mathcal{F}, P)$, one can compute the expected value of X subjected to the condition that the event B has taken place by computing its expectation with respect to the measure $P(\cdot|B)$. More specifically, one obtains the ***conditional expectation*** of X with respect to the event B

$$E[X|B] := \frac{E[X\mathbf{1}_B]}{P(B)}.$$

One may want to generalize this notion now only defined for a single event to a family of them. Thus, let $\mathcal{B} \subset \mathcal{F}$ be a sub σ-algebra and $X \in L^1(\Omega, \mathcal{F}, P)$. For any $B \in \mathcal{B}$, the function $\mu(B) := E[X\mathbf{1}_B]$ defines a σ-additive set function on $\mathcal{B}$ which is clearly absolutely continuous with respect to P. The Radon-Nikodym Theorem guarantees in this situation the existence of a P-unique function $E[X|\mathcal{B}] : \Omega \to [0, \infty]$ such that

$$\mu(B) = \int_B E[X|\mathcal{B}](\omega) dP(\omega).$$

The random variable $E[X|\mathcal{B}]$ is called the ***conditional expectation*** with respect to a σ-algebra $\mathcal{B}$.

Theorem 1.6 (Conditional expectation) *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{B} \subset \mathcal{F}$ be a sub σ-algebra and $X \in L^1(\Omega, \mathcal{F}, P)$. There exists a unique random variable $E[X|\mathcal{B}] \in L^1(\Omega, \mathcal{B}, P)$ such that for any $B \in \mathcal{B}$*

$$E[X\mathbf{1}_B] = E[E[X|\mathcal{B}]\mathbf{1}_B]. \tag{1.2}$$

Equivalently $E[X|\mathcal{B}]$ can be characterized by saying that for any bounded and $\mathcal{B}$-measurable random variable Z

$$E[XZ] = E[E[X|\mathcal{B}]Z]. \tag{1.3}$$

Remark 1.7 If $\mathcal{B}$ is the σ-algebra generated by a random variable Y we will write

$$E[X|\mathcal{B}] = E[X|\sigma(Y)] = E[X|Y].$$

Remark 1.8 The L^2 case admits an interesting geometric interpretation. Let $\langle X, Y\rangle = E[XY]$ the Euclidean product in L^2, where $X, Y \in L^2(\Omega, \mathcal{F}, P)$. If $\mathcal{B} \subset \mathcal{F}$ is a sub σ-algebra then $L^2(\Omega, \mathcal{B}, P)$ is a closed subspace of $L^2(\Omega, \mathcal{F}, P)$. In this situation, the conditional expectation $E[X|\mathcal{B}]$ is the orthogonal projection of X on $L^2(\Omega, \mathcal{B}, P)$ and hence $E[X|\mathcal{B}]$ can be interpreted as the best approximation of X (in the L^2 norm) by a $\mathcal{B}$-measurable random variable. In fact, if $Y \in L^2(\Omega, \mathcal{B}, P)$, this result comes from the straightforward computation

$$\begin{aligned}
\|X - Y\|_{L^2} &= E[(X-Y)^2] = E[X^2] + E[Y^2] - 2E[XY] \\
&= E[X^2] + E[Y^2] - 2E[E[X|\mathcal{B}]Y] \\
&= E[X^2] - E[E[X|\mathcal{B}]^2] + E[(Y - E[X|\mathcal{B}])^2] \\
&\geq E[X^2] - E[E[X|\mathcal{B}]^2],
\end{aligned} \tag{1.4}$$

where in the second line (1.3) has been used. Since the orthogonal projection minimizes the norm $\|X - Y\|_{L^2}$, this projection is $Y = E[X|\mathcal{B}]$ because only in this case the inequality in the last line of (1.4) is actually an equality.

Elementary properties of the conditional expectation

Let $X : \Omega \to \mathbb{R}$ denote an arbitrary random variable. Then,

(i) If X is $\mathcal{B}$-measurable, $E[X|\mathcal{B}] = X$.

(ii) The map $X \longmapsto E[X|\mathcal{B}]$ is linear.

(iii) A random variable and its conditional expectation have the same expectation:

$$E[E[X|\mathcal{B}]] = E[X].$$

This follows from (1.2) with $B = \Omega$.

(iv) If X and $\mathcal{B}$ are independent, that is, if $\sigma(X)$ and $\mathcal{B}$ are independent, then $E[X|\mathcal{B}] = E[X]$. Indeed, the constant $E[X]$ is clearly $\mathcal{B}$-measurable and for any $B \in \mathcal{B}$ we have

$$E[X\mathbf{1}_B] = E[X]E[\mathbf{1}_B] = E[E[X]\mathbf{1}_B]$$

therefore $E[X|\mathcal{B}] = E[X]$ by (1.2). In general, two sub σ-algebras $\mathcal{B}_1$ and $\mathcal{B}_2$ are independent if and only if for any $\mathcal{B}_2$-measurable random variable X

$$E[X|\mathcal{B}_1] = E[X]. \tag{1.5}$$

This implies in particular that if two random variables X and Y are independent then

$$E[X|Y] = E[X], \tag{1.6}$$

but the converse is not true. It can be shown that the random variables measurable with respect to $\sigma(X)$ are those of the form $h(X)$, with $h : \mathbb{R} \to \mathbb{R}$ a measurable map. Consequently, by (1.5), X and Y are independent if and only if $E[h(X)|Y] = E[h(X)]$ for any h, which obviously implies (1.6) but not the other way around.

(v) If X, Y are real random variables and $X \leq Y$, then $E[X|\mathcal{B}] \leq E[Y|\mathcal{B}]$ a.s.. Taking $Y = |X|$, this monotone property implies that $|E[X|\mathcal{B}]| \leq E[|X||\mathcal{B}]$ a.s.. More generally, Jensen's inequality also holds. That is, if φ is a convex function such that $E[|\varphi(X)|] < \infty$, then

$$\varphi\left(E[X|\mathcal{B}]\right) \leq E[\varphi(X)|\mathcal{B}].$$

(vi) If X, Y are real valued random variables and Y is $\mathcal{B}$-measurable and bounded, then

$$E[YX|\mathcal{B}] = YE[X|\mathcal{B}]. \tag{1.7}$$

Indeed, (1.2) implies that for any bounded and $\mathcal{B}$-measurable random variable $E\left[E[X|\mathcal{B}]Y\right] = E[XY]$ as can be easily checked approximating Y by a suitable sequence of $\mathcal{B}$-measurable elementary process. (1.7) stems immediately from this fact.

(vii) If $\mathcal{B}_1$ and $\mathcal{B}_2$ are sub σ-algebras such that $\mathcal{B}_1 \subset \mathcal{B}_2$, then

$$E[E[X|\mathcal{B}_2]|\mathcal{B}_1] = E[X|\mathcal{B}_1].$$

1.2 Stochastic processes

A stochastic process is a sequence of random variables that evolve in time. In this section $(\Omega, \mathcal{F}, P)$ will be a probability space, $(E, \mathcal{E})$ a measurable space, and T a set of time indices (it may represent the sets $\mathbb{N}$, $\mathbb{R}$, $\mathbb{Z}$, or $\mathbb{R}^d$). A mapping $X : T \times \Omega \to E$ is called a ***stochastic process*** if $X_t : \Omega \to E$ is a random variable for any $t \in T$.

The term stochastic process is sometimes reserved to the case in which $T = \mathbb{R}$. When $T = \mathbb{N}$ or $\mathbb{Z}$ (respectively, $T = \{1, \ldots, N\}$ one uses the term ***time series*** (respectively, ***random vector***). The case $T = \mathbb{R}^d$ corresponds to the so called ***random fields***. In this thesis, however, we are only going to deal with the case $T = \mathbb{R}_+$ or $\mathbb{R}$.

1.2.1 The law of a stochastic process. Continuous processes

The ideas in the following paragraph allow us to code stochastic processes as random variables and hence to apply all the notions that we have previously learnt for these objects. Let $E^T := \{f : T \to E\}$ the space of all the maps from T to E. A ***measurable finite cylinder set*** is a set $C \subset E^T$ of the form

$$C^A_{t_1,\ldots,t_n} := \{f : T \to E \mid (f(t_1), \ldots, f(t_n)) \in A\}$$

for some fixed sequence $\{t_1, \ldots, t_n\} \subset T$, $n \in \mathbb{N}$, and some fixed $A \in \mathcal{E}$. We will denote the family of all measurable cylinder sets by $\mathcal{C}$. In this context, the ***product σ-algebra*** $\mathcal{E}^{\otimes T}$ is the one generated by the measurable finite cylinders sets. That is, $\mathcal{E}^{\otimes T} := \sigma(\mathcal{C})$. Given a stochastic process $X : T \times \Omega \to E$, we will denote by $\overline{X} : \Omega \to E^T$ be the map that assigns to each $\omega \in \Omega$ the path $\{X_t(\omega) \mid t \in T\} \in E^T$.

Proposition 1.9 *$X : T \times \Omega \to E$ is a stochastic process if and only if $\overline{X} : \Omega \to (E^T, \mathcal{E}^{\otimes T})$ is a random variable.*

Using this proposition, we rephrase for stochastic processes what we introduced before for random variables. More specifically, the ***law of a stochastic process*** $X : T \times \Omega \to E$ is the law $P_{\overline{X}}$ of the random variable $\overline{X} : \Omega \to (E^T, \mathcal{E}^{\otimes T})$. For example, if $C^A_{t_1,\ldots,t_n}$ is a measurable finite cylinder set, then

$$P_{\overline{X}}(C) = P((X_{t_1}, \ldots, X_{t_n}) \in A_{t_1} \times \cdots \times A_{t_n}).$$

It is worth noticing that given a probability measure Q on $(E^T, \mathcal{E}^{\otimes T})$ there exists at least one stochastic process that has it as a law: indeed, the ***canonical process*** $X : \Omega \times \mathbb{R} \longrightarrow E$ defined on $(\Omega = E^T, \mathcal{E}^{\otimes T})$ by $X_t(\omega) = \omega_t$ is such that $P_{\overline{X}} = Q$.

In order to introduce continuous processes, we are going to suppose in the following paragraphs that $(E, \mathcal{E})$ is a metric space with distance function d, $\mathcal{E} = \mathcal{B}(E)$ is the Borel σ-algebra, and $T = \mathbb{R}$, $\mathbb{R}_+$, or the interval $[a, b]$. Let $C^0(T, E)$ denote the space of continuous maps from T to E. We say that the stochastic process $X : T \times \Omega \to E$ is ***continuous*** if its paths $\{X_t(\omega) \mid t \in T\}$ are continuous for any $\omega \in \Omega$ a.s. and, therefore, $\overline{X} : \Omega \to C^0(T, E)$.

The set $C^0(T, E)$ has a natural topology given by the ***uniform convergence over compact sets*** (also called ***compact convergence***) and hence a natural Borel σ-algebra $\mathcal{B}(C^0(T, E))$

which, in principle, differs from the product σ-algebra $\mathcal{B}(E)^{\otimes T}$. Recall that the topology of the compact convergence is a metric topology with distance function given by

$$d_C(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n}\left[\left(\sup_{0\le t\le n} d(f(t),g(t))\right)\wedge 1\right], \quad f,g \in C^0(T,E).$$

It is not difficult to check that, if T is compact, the compact topology coincides with the topology given by the norm of the uniform convergence:

$$\|f-g\|_\infty = \max_{t\in T}\{d(f(t),g(t))\}.$$

Nevertheless, $\mathcal{B}(C^0(T,E)) \subset \mathcal{B}(E)^{\otimes T}$. Indeed, we can consider the ***Borel finite cylinder sets***, which are sets in E^T of the form

$$\left(C^0\right)^A_{t_1,\dots,t_n} := \{f \in C^0(T,E) \mid (f(t_1),\dots,f(t_n)) \in A\}$$

for some sequence $t_1 \le t_2 \le \dots \le t_n$ and some $A \in \mathcal{B}(E)$. If $\mathcal{C}^0$ denotes the set of all the Borel finite cylinder sets, then clearly $\sigma\left(\mathcal{C}^0\right) \subseteq \mathcal{B}(E)^{\otimes T}$ from the definition of the product σ-algebra $\mathcal{B}(E)^{\otimes T}$. What is less obvious is that $\sigma\left(\mathcal{C}^0\right) = \mathcal{B}(C^0(T,E))$ (see [IW89, Chapter I, Proposition 4.1]), which proves the assertion $\mathcal{B}(C^0(T,E)) \subset \mathcal{B}(E)^{\otimes T}$.

UCP convergence. The *ucp* convergence is the analog of convergence in probability when stochastic processes are coded as random variables and will play a prominent role in the theory of stochastic integration. More specifically, let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of processes. We say that $\{X_n\}_{n\in\mathbb{N}}$ ***converges uniformly on compacts in probability (ucp)*** to a process X whenever for any $\varepsilon > 0$ and any $t \in \mathbb{R}_+$,

$$P\left(\left\{\sup_{0\le s\le t} |X_n - X|_s\right\} > \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 0.$$

1.2.2 Brownian motion

Brownian motion is probably the most important continuous process in stochastic calculus and the most used to model the stochastic behavior of real systems arising from very different disciplines such as statistical physics or finance. Some of the seminal works on Brownian motions were those by Brown (1827), Bachelier (1900), Einstein (1905), and Wiener (1925). We say that a $\mathbb{R}^d$-valued process $B : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ is a ***Brownian motion*** with initial law μ if for any partition $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$ of the interval $[0,t]$, the random variables $B_{t_i} - B_{t_{i-1}}$ are mutually independent, $i = 1,\dots,n$, $P_{B_{t_0}} = \mu$, and $P_{B_{t_i}-B_{t_{i-1}}}$ is absolutely continuous with respect to the Lebesgue measure with Gaussian pdf

$$\frac{1}{(2\pi(t_i - t_{i-1}))^{d/2}}\, \mathrm{e}^{-\frac{\|x\|^2}{2(t_i-t_{i-1})}}.$$

If $d = 1$, this means that $B_{t_i} - B_{t_{i-1}} \sim N(0, t_i - t_{i-1})$. Whenever the initial law μ is not specified it will implicitly assumed that $B_0 = 0$ a.s.. If $W\left(\mathbb{R}^d\right) := \{\omega : \mathbb{R}_+ \to \mathbb{R}^d \mid \omega \text{ continuous}\}$, then the probability law $P_{\bar{B}}$ on $\left(W\left(\mathbb{R}^d\right), \mathcal{B}\left(W\left(\mathbb{R}^d\right)\right)\right)$ induced by a Brownian motion $\bar{B} : \Omega \to W\left(\mathbb{R}^d\right)$ with initial law μ is called the ***Wiener measure*** with initial distribution μ. By [IW89, Chapter I, Theorem 7.1], the Wiener measure P^W_μ exists on $\left(W\left(\mathbb{R}^d\right), \mathcal{B}\left(W\left(\mathbb{R}^d\right)\right)\right)$ for any initial distribution μ which implies, in turn, that the canonical process on $\left(W\left(\mathbb{R}^d\right), \mathcal{B}\left(W\left(\mathbb{R}^d\right)\right), P^W_\mu\right)$ is a Brownian motion. In other words, Brownian motions do exist. We will come back later on the problem of defining a Brownian motion once Markov processes have been introduced. Some of the most relevant properties of Brownian motions are the following:

(i) The autocovariance of a real Brownian motion ($B_0 = 0$) is given by $E[B_t B_s] = \min(s, t)$ ([O03, Lemma 6.2.6]). Indeed, $E[B_t^2] = t$ from the very definition of Brownian motion. If $s < t$,

$$E[B_t B_s] = E[(B_t - B_s)B_s] + E[B_s^2].$$

But $E[(B_t - B_s)B_s] = E[B_t - B_s]E[B_s]$ because $B_t - B_s$ is independent of B_s and $E[B_s] = 0$. Hence

$$E[B_t B_s] = E[B_s^2] = s.$$

(ii) If B is a Brownian motion, then so are ([K97, Theorem 2.1], [KS91, Chapter 2 Lemma 9.4]):

(a) $X_t = a^{-1}B_{a^2 t}$, with $a \neq 0$ (Brownian rescaling property).

(b) $X_t = tB_{1/t}$, with $t > 0$ and $X(0) = 0$ (time inversion).

(c) $X_t = B_{t+t_0} - B_{t_0}$, $t \geq 0$. It is immediate to see that $B_{t+t_0} - B_{t_0}$ fulfills the definition of a Brownian motion.

(iii) The paths of the Brownian motion are nowhere differentiable ([PWZ33], [KS91, Section 2.9 D]). This is a remarkable property that prevents us from defining any natural notion of velocity or derivative associated to a given Brownian motion.

***p*-variation.** Let $X_t : \Omega \to \mathbb{R}$ be a continuous stochastic process. The p-th variation $[X, X]^{(p)}_t(\omega)$ of the path $X(\omega) : \mathbb{R}_+ \to \mathbb{R}$ is defined by

$$[X, X]^{(p)}_t(\omega) = \lim_{\Delta t_k \to 0} \sum_{t_k \leq t} |X_{t_{k+1}}(\omega) - X_{t_k}(\omega)|^p, \tag{1.8}$$

provided this limit exits. The process associated with $p = 1$ is referred to as the ***total variation***; the case $p = 2$ is called the ***quadratic variation***.

A result of Lévy shows that the total variation of the paths of the Brownian motion are $+\infty$ on every time interval. This feature of the Brownian motion makes non-trivial the integration theory that uses it as integrator (recall that the Riemann-Stieltjes is defined only for integrators with bounded variation). The ultimate reason why the stochastic integral that we will introduce later on works is the fact that the Brownian motion has finite quadratic variation. Integration with respect to processes with finite p-variation, $p > 2$, is the subject of the so-called ***Rough Paths Theory*** ([CLT04]), which lies beyond the scope of this thesis. The next theorem gives an explicit expression for the quadratic variation of a (one-dimensional) Brownian motion.

Theorem 1.10 *Let $B : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ a real Brownian motion and let $\{\pi_n\}_{n\in\mathbb{N}}$ be a sequence of partitions of the interval $[0,t]$, that is, $0 = t_0^n \leq t_1^n \leq ... \leq t_{k_n}^n = t$, such that $\|\pi_n\| = \max_{i\in\{1,\dots,k_n\}} |t_i^n - t_{i-1}^n| \to 0$ as $n \to \infty$. Let $\pi_n B := \sum_{t_i\in\pi_n} \left(B_{t_{i+1}} - B_{t_i}\right)^2$. Then $[B,B]_t^{(2)} = \lim_{n\to\infty} \pi_n B = t$ in $L^2(\Omega, \mathcal{F}, P)$.*

Proof. We have

$$\pi_n B - t = \sum_{t_i\in\pi_n} \left\{ \left(B_{t_{i+1}} - B_{t_i}\right)^2 - (t_{i+1} - t_i) \right\} =: \sum_{i=0}^{k_n-1} Y_i$$

where $Y_i := \left(B_{t_{i+1}} - B_{t_i}\right)^2 - (t_{i+1} - t_i)$ are independent random variables with zero mean. Then,

$$E\left[(\pi_n B - t)^2\right] = E\left[\left(\sum_{i=0}^{k_n-1} Y_i\right)^2\right] = \sum_{i=0}^{k_n-1} E\left[Y_i^2\right].$$

Next observe that $\left(B_{t_{i+1}} - B_{t_i}\right)^2 \Big/ (t_{i+1} - t_i)$ has the same distribution as Z^2, where Z is a Gaussian with 0 mean and variance 1. Therefore,

$$\begin{aligned} E\left[(\pi_n B - t)^2\right] &= \sum_{i=0}^{k_n-1} E\left[\left(\frac{Y_i}{(t_{i+1} - t_i)}(t_{i+1} - t_i)\right)^2\right] \\ &= E\left[\left(Z^2 - 1\right)^2\right] \sum_{i=0}^{k_n-1} (t_{i+1} - t_i)^2 \\ &\leq E\left[\left(Z^2 - 1\right)^2\right] \|\pi_n\| \, t, \end{aligned}$$

which tends to 0 as n tends to ∞. ■

Remark 1.11 One can also prove that $[B,B]_t^{(2)} = t$ a.s. ([P05, Chapter I Theorem 28], [CW90, Theorem 6.1]). Morevoer, the Levy's characterization of Brownian motion (Theorem 1.31) claims that $[B,B]_t^{(2)} = t$ characterizes uniquely Brownian motions among those continuous processes which have the additional property of being local martingales (see Subsection 1.2.3).

1.2.3 Filtrations, martingales, stopping times, and Markov processes

Martingales and Markov processes are some of the most important classes of stochastic processes. In order to define them we need the notion of filtration. In this subsection, the time parameter space T will be either $\mathbb{R}_+$ or $\mathbb{N}$.

Definition 1.12 *A **filtration** of the measurable space $(\Omega, \mathcal{F})$ is an increasing sequence $\mathcal{F} = \{\mathcal{F}_t\}_{t\in T}$ of sub σ-algebras of $\mathcal{F}$, that is, $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$. We will usually assume that $\mathcal{F}_0$ contains all the negligible events (complete filtration) and that the map $t \longmapsto \mathcal{F}_t$ is right-continuous, that is, $\mathcal{F}_t = \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$. A stochastic process $X : T \times \Omega \longrightarrow E$ is said to be **adapted** to the filtration $\mathcal{F}$ when $X_t : \Omega \to E$ is $\mathcal{F}_t$-measurable for any $t \in \mathbb{R}_+$. The filtration $\mathcal{F}^X$ induced by the process X is the minimal filtration with respect to which X is adapted; more specifically $\mathcal{F}_t^X = \sigma(X_s \mid s \leq t)$ for any $t \in \mathbb{R}_+$.*

Filtrations are used to model the information available at a given time. In the case of $\mathcal{F}^X$, the σ-algebra $\mathcal{F}_t^X$ represents the information obtained by observing the values taken by X between the instants 0 and t.

Definition 1.13 *A real-valued* ***martingale*** $X : T \times \Omega \to \mathbb{R}^d$ *is a stochastic process such that for every pair* $t, s \in T$ *such that* $s \leq t$*, we have:*

(i) X is $\mathcal{F}_t$-adapted, that is, X_t is $\mathcal{F}_t$-measurable.

(ii) $X_s = E[X_t \mid \mathcal{F}_s]$.

(iii) X_t is integrable: $E[|X_t|] < +\infty$.

For any $p \in [1, \infty)$, X is called a L^p***-martingale*** whenever X is a martingale and $X_t \in L^p(\Omega, P)$ for each t. If $\sup_{t \in \mathbb{R}_+} \mathrm{E}[|X_t|^p] < \infty$, we say that X is L^p***-bounded***.

When the filtration $\mathcal{F}$ is interpreted as the amount of information available at any given time, martingales encode the notion of *fair game*. One feature that makes this plausible is that martingales have constant expectation; taking expected values on both sides of the equality $X_0 = E[X_t|\mathcal{F}_0]$, we actually obtain that $E[X_0] = E[X_t]$.

Examples 1.14

(i) Random walk. Let $T = \mathbb{N}$ and let $\{\xi_n\}_{n \in \mathbb{N}}$ be a sequence of independent integrable random variables that share the same law and that have zero mean. Then, the associated random walk process $S_n := \xi_1 + \cdots + \xi_n$ is a martingale for its own filtration $\mathcal{F}_n^S = \sigma(\xi_i \mid 1 \leq i \leq n)$. Indeed, each S_n is integrable and

$$E[S_{n+1}|\mathcal{F}_n^S] = E[S_n + \xi_{n+1}|\mathcal{F}_n^S] = S_n + E[\xi_{n+1}] = S_n,$$

where we have used linearity and independence.

(ii) The Brownian motion. The Brownian motion B is a martingale with respect to its own filtration $\mathcal{F}^B$. By definition, $B_t - B_s$ is independent of $\mathcal{F}_s^B$ for any $0 \leq s < t$ and hence

$$E[B_t|\mathcal{F}_s^B] = E[B_s + B_t - B_s|\mathcal{F}_s^B] = B_s + E[B_t - B_s] = B_s.$$

We are going to recall now one of the most important concepts in stochastic calculus intrinsically linked with filtrations, that of stopping time. Stopping times are random variables which give us information about the time at which something happens. For instance, the time at which a given process leaves an open set or exceeds some bound are examples of stopping times. In addition, we want the answer to these sort of questions to be *adapted* or *in accordance with* the amount of information available about the process at a certain instant. Thus, we define:

Definition 1.15 *A random variable* $\tau : \Omega \to [0, +\infty]$ *is called a* ***stopping time*** *with respect to the filtration* $\{\mathcal{F}_t\}_{t \in T}$ *if for every* $t \geq 0$ *the set* $\{\omega \mid \tau(\omega) \leq t\}$ *belongs to* $\mathcal{F}_t$*. Given a stopping time* τ*, we define*

$$\mathcal{F}_\tau = \{\Lambda \in \mathcal{F} \mid \Lambda \cap \{\tau \leq t\} \in \mathcal{F}_t \textit{ for any } t \in T\}.$$

Given an adapted process X, it can be shown that the random variable X_τ is $\mathcal{F}_\tau$-measurable. We define the ***stopped process*** X^τ as

$$X_t^\tau := X_{t\wedge\tau} := X_t \mathbf{1}_{\{t\leq\tau\}} + X_\tau \mathbf{1}_{\{t>\tau\}}.$$

The next proposition provides us with a huge amount of stopping times. In particular, all the stopping times we are going to use in this thesis will be of such form.

Proposition 1.16 ([P05, Chapter I Theorem 3]) *Let X be a stochastic process adapted to $\{\mathcal{F}_t\}_{t\in T}$ and let $\Lambda \subseteq \mathbb{R}^d$ a Borel set. For any $\omega \in \Omega$, let*

$$\tau(\omega) := \inf\{t > 0 \mid X_t \in \Lambda\}$$

*the **hitting time** of Λ for X. If the paths of X are right-continuous and have left-limits a.s., then τ is a stopping time.*

If instead of $\Lambda \subseteq \mathbb{R}^d$ we consider its complementary set $\Lambda^c = \mathbb{R}^d \backslash \Lambda$, then τ in Proposition 1.16 is referred to as the ***exit time*** of Λ.

It is customary in stochastic calculus to say that some properties hold locally. The notion of *local* applied to stochastic processes has a quite different meaning than that one uses in geometry, where properties hold locally if they hold on an open neighborhood. Hence, we say that a stochastic process $X : T \times \Omega \to \mathbb{R}^d$ satisfies some property locally if there exists a non-decreasing sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \tau_n = \infty$ a.s. and X^{τ_n} satisfies that property for any $n \in \mathbb{N}$. For example, we will say that $X : T \times \Omega \to \mathbb{R}^d$ is a ***local martingale*** or ***locally bounded*** if there exist a non-decreasing sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ such that X^{τ_n} is a martingale or bounded for any $n \in \mathbb{N}$ respectively.

Homogeneous Markov processes

Let $\left(\Omega, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, P\right)$ be a standard filtered probability space and let $(E, \mathcal{E})$ be a measurable space. A function $p : \mathbb{R}_+ \times E \times \mathcal{E} \to [0, 1]$ is called a ***transition function*** provided that

(i) $p_t(x, \cdot)$ is a probability measure on $\mathcal{E}$ for any $t \in \mathbb{R}_+$ and any $x \in E$.

(ii) $p_t(\cdot, A)$ is $\mathcal{E}/\mathcal{B}([0, 1])$-measurable for any $t \in \mathbb{R}_+$ and any $A \in \mathcal{E}$.

(iii) For any $t, s \in \mathbb{R}_+$, any $x \in S$, and any $A \in \mathcal{E}$,

$$p_{t+s}(x, A) = \int_S p_t(x, dy) p_s(y, A). \tag{1.9}$$

The relationship (1.9) is known as the Chapman-Kolmogorov equation. It says that the probability of being in A at time $t+s$ starting from x is equal to the sum of all the probabilities of being at an intermediate point y at time t to be, s units of time later, in A. A stochastic process X defined on $\left(\Omega, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, P\right)$ and taking values in $(E, \mathcal{E})$ is a ***temporally homogeneous Markov process*** with transition function $p : \mathbb{R}_+ \times E \times \mathcal{E} \to [0, 1]$ if it is adapted and

$$E[f(X_t)|\mathcal{F}_s] = \int_S p_{t-s}(X_s, dy) f(y) \tag{1.10}$$

for any $0 \leq s < t$ and any $f : E \to [0,\infty]$ such that f is $\mathcal{E}/\mathcal{B}([0,\infty])$-measurable. Two Markov processes with the same state space $(E,\mathcal{E})$ are said to be ***equivalent*** if they have the same transition function. It is a cornerstone result in the theory of Markov processes that, given a transition function p on a measurable space $(E,\mathcal{E})$ and a probability measure μ on $(E,\mathcal{E})$, there always exists a filtered probability space $\left(\Omega, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, P^\mu\right)$ and an adapted stochastic process X such that X has μ as initial law and (1.10) holds providing E is σ-compact and $\mathcal{E}$ is the topological σ-algebra of Borel sets ([BG68, Chapter I]). Recall that a topological space is called σ-compact if it is the union of countable many compact subsets. The existence of Markov processes is based on the famous Kolmogorov's Extension Theorem. When E is a topological space, Ω is usually taken as the set of all continuous paths or all right continuous paths with left limits if, for example, the Markov process X is continuous or right continuous with left limits respectively.

Let $\epsilon_x(\cdot)$ be the measure such that $\epsilon_x(A) = 1$ if $x \in A$, $A \in \mathcal{E}$, and 0 otherwise. Let P^x denote the probability associated to $\mu = \epsilon_x(\cdot)$ and $E^x[\cdot]$ the expectation carried out under this law. Observe that (1.10) can be rewritten as

$$E[f(X_t)|\mathcal{F}_s] = E^{X_s}[f(X_{t-s})] \tag{1.11}$$

for any $0 \leq s < t$. Let now $f : E \to [0,\infty]$ $\mathcal{E}/\mathcal{B}([0,\infty])$-measurable and take $t \in \mathbb{R}_+$. We define

$$(P_t f)(x) := \int_E p_t(x,dy)f(y) = E^x[f(X_t)].$$

$\{P_t\}_{t\in\mathbb{R}_+}$ is called the ***transition semigroup*** of the Markov process X because $P_0 f = f$ and $P_{t+s}f = P_t(P_s f)$ if, in addition, $f \in L^\infty(E)$. In this latter case, $\{P_t\}_{t\in\mathbb{R}_+}$ is a ***contraction semigroup***. That is, $\|P_t f\|_\infty \leq \|f\|_\infty$.

Example 1.17 Let $(E,\mathcal{E}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ and

$$p_t(x,dy) = \frac{1}{(2\pi t)^{n/2}}\,\mathrm{e}^{-\frac{\|x-y\|^2}{2t}}\,dy,$$

where dy stands for the Lebesgue measure of $\mathbb{R}^n$. The corresponding Markov process associated to its transition function and $\mu = \epsilon_x(\cdot)$ is the Brownian motion starting at $x \in \mathbb{R}^n$. As we saw in subsection 1.2.2, in this case Ω is $C^0(\mathbb{R}_+, \mathbb{R}^n)$, the space of all continuous paths from $\mathbb{R}_+$ to $\mathbb{R}^n$.

The property (1.11) conveys the idea that the law of X_t conditioned to knowing all the past up to time s (modeled by $\mathcal{F}_s$) depends only on the value X_s. Equivalently, one can say that *a Markov process depends on the past only through the present.* A deterministic analog of the Markov processes are the solutions $x(t)$ of a first order differential equation which are fully determined by their value at any time t, say $t = 0$.

1.3 Stochastic integration and stochastic differential equations

One of the goals of this section is giving a meaning to the expression

$$\int_0^t \phi(s,\omega)dB_s. \tag{1.12}$$

As we already noted, the Brownian motion has infinite total variation and hence the classical theory of (Riemann-Stieltjes) integration is not valid. The integral (1.12) was first defined by Wiener in 1934 for deterministic integrands, that is, $\phi = \phi(t)$. The general stochastic case $\phi = \phi(t, \omega)$ was introduced by Itô in the 40's.

The price to pay for having an integral at this level of generality is that there are choices involved: *there is not a unique stochastic integral*. In order to illustrate this point we outline the strategy for the definition of the integral so that we can pinpoint what the difficulties are.

Take the function $\phi(u, \omega) := X(\omega)\mathbf{1}_{(s,t]}(u)$, for some random variable X. A natural choice for the value of the integral (1.12) in this case is $X(\omega)(B_t(\omega) - B_s(\omega))$. If we require the integral to satisfy the usual linearity properties, the integral should be worth

$$\int_0^t \phi(s, \omega) dB_s = \sum_{i=0}^{n-1} X_i(\omega)(B_{t_{i+1}}(\omega) - B_{t_i}(\omega)),$$

whenever

$$\phi(u, \omega) := \sum_{i=0}^{n-1} X_i(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(u). \tag{1.13}$$

The strategy to define the integral for a general function ϕ consists roughly of finding an approximating sequence $\{\phi_n\}_{n\in\mathbb{N}}$ made of functions of the kind (1.13) and then setting

$$\int \phi dB = \lim_{n\to\infty} \int \phi_n dB.$$

The main difficulty in properly stating this definition consists in the fact that the value of the integral may depend on the choice of approximating sequence and therefore we have to be very specific in this particular point. For instance, take the function $\phi(t, \omega) = B_t(\omega)\mathbf{1}_{[0,T]}(t)$. This function can be approximated by the following two natural sequences of functions, namely

$$\phi_n(t) = \sum_{i=0}^{n-1} B_{t_{i+1}}\mathbf{1}_{(t_i, t_{i+1}]}, \text{ and } \psi_n(t) = \sum_{i=0}^{n-1} B_{t_i}\mathbf{1}_{(t_i, t_{i+1}]}.$$

The choice of one sequence or the other leads to a different definition of the integral. Indeed,

$$\int \phi_n dB - \int \psi_n dB = \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2.$$

This expression tends to the quadratic variation of the Brownian motion as the diameter of the partition goes to zero which, as we saw in Theorem 1.10, $[B, B]_t^{(2)} = t$. The Itô integral consists in taking $\{\psi_n\}_{n\in\mathbb{N}}$ as approximating sequence. This sequence has the feature of being non-anticipative and adapted.

1.3.1 The Itô stochastic integral with respect to a semimartingale

The construction of the stochastic integral that we will present in this section is more general than (1.12) in the sense that we will have as integrator not just the Brownian motion, but an arbitrary *semimartingale*. Semimartingales are the most general and natural setup for stochastic integration and differentiation, in the sense that stochastic differential equations formulated using semimartingales have semimartingales as solutions. We start with several definitions. All along this subsection we will consider *real valued processes*.

We say that the stochastic process $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ has ***finite variation*** whenever it is adapted and has bounded total variation on compact intervals of $\mathbb{R}_+$. More explicitly, a process X has finite variation if for each fixed $\omega \in \Omega$ the path $t \longmapsto X_t(\omega)$ has bounded total variation on compact intervals of $\mathbb{R}_+$. That is, the supremum $\sup\left\{\sum_{i=1}^{p} |X_{t_i}(\omega) - X_{t_{i-1}}(\omega)|\right\}$ over all the partitions $0 = t_0 < t_1 < \cdots < t_p = t$ of the interval $[0,t]$ is finite for any $t \in \mathbb{R}_+$. This is equivalent to the existence of a signed measure μ_ω on $\mathbb{R}_+$ such that, assuming $X_0 = 0$, is given by

$$X_t(\omega) = \mu_\omega([0,t]) = \sup\left\{\sum_{i=1}^{p} |X_{t_i}(\omega) - X_{t_{i-1}}(\omega)|\right\}. \tag{1.14}$$

We can now introduce the processes that we are going to use as ***integrators*** in the stochastic integral, namely, semimartingales.

Definition 1.18 *A* ***continuous semimartingale*** *is the sum of a continuous local martingale and a process with finite variation.*

It can be proved that a given continuous semimartingale has a unique decomposition of the form

$$X = X_0 + V + M, \tag{1.15}$$

with X_0 the initial value of X, V a finite variation process, and M a local continuous semimartingale provided both V and M are null at time equal to zero ([P05, Chapter III Theorem 2]). It is worth noting that if we remove the hypothesis of X being continuous, then the decomposition (1.15) may not be unique.

We proceed by presenting the processes that will be used as ***integrands*** in the stochastic integral, namely, the càglàd processes. Let $\mathbb{L}$ the space of processes $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ whose paths are left-continuous and have right limits. They are usually called ***càglàd*** processes which is the French acronym for left-continuous with right limits. We denote by $\mathbb{D}$ the space of processes whose paths are right-continuous and have left limits. They are usually called ***càdlàg***. The *ucp* topology on $\mathbb{D}$ yields a complete and metrizable space using

$$d(X,Y) = \sum_{n=1}^{\infty} \frac{1}{2^n} E\left[\min\left(1, \sup_{0\le s\le n} |X_s - Y_s|\right)\right].$$

Definition 1.19 *We say that a stochastic process $H \in \mathbb{L}$ is* ***simple predictable*** *if it can be expressed as*

$$H = H_0 \mathbf{1}_{\{0\}} + \sum_{i=1}^{p-1} H_i \mathbf{1}_{(\tau_i, \tau_{i+1}]}, \tag{1.16}$$

where $0 \le \tau_1 \le \cdots \le \tau_{p-1} \le \tau_p$ are stopping times, H_0 and H_i are $\mathcal{F}_0$-measurable and $\mathcal{F}_{\tau_i}$-measurable random variables respectively such that $|H_0| < \infty$ and $|H_i| < \infty$ a.s. for all i,

$\mathbf{1}_{(\tau_i,\tau_{i+1}]}$ *is the characteristic function of the set* $\{(t,\omega) \in R_+ \times \Omega \mid t \in (\tau_i(\omega), \tau_{i+1}(\omega)]\}$*, and* $\mathbf{1}_{\{0\}}$ *that of* $\{(t,\omega) \in R_+ \times \Omega \mid t = 0\}$*. The collection of simple predictable processes is denoted by* $\mathbf{S}$*.*

The Itô integral is defined first for integrands that are simple predictable processes $\mathbf{S}$. Since $\mathbf{S}$ is dense in $\mathbb{L}$ under the *ucp* topology, the continuity of the definition of the integral allows the extension of Itô integral to any càglàd process. We explicitly state the needed chain of definitions and results.

Definition 1.20 *Let* $H \in \mathbf{S}$ *be a simple predictable process and* X *a (continuous) semimartingale. Define the linear map* $J_X : \mathbf{S} \longrightarrow \mathbb{D}$ *as*

$$H \cdot X := \int H dX := \sum_{i=1}^{p-1} H_i (X^{\tau_{i+1}} - X^{\tau_i}). \tag{1.17}$$

where $H = H_0 \mathbf{1}_{\{0\}} + \sum_{i=1}^{p-1} H_i \mathbf{1}_{(\tau_i,\tau_{i+1}]}$ *is given as in (1.16).* $J_X(H)$ *is called the* ***Itô stochastic integral*** *of* H *with respect to* X*.*

Proposition 1.21 $\mathbf{S}$ *is dense in* $\mathbb{L}$ *under the ucp topology.*

Proof. Let $Y \in \mathbb{L}$ be a process and $\rho_n = \inf\{t : |Y_t| > n\}$ a sequence of stopping times. Observe that $\rho_n \to \infty$ a.s.. Then $Y^n = Y^{\rho_n} \mathbf{1}_{\{\rho_n > 0\}} \in \mathbf{b}\mathbb{L}$ converge to Y in *ucp*, where $\mathbf{b}\mathbb{L}$ denotes the set of bounded càglàd processes. Indeed,

$$P\left(\left\{\sup_{0 \le s \le t} |Y_s^n - Y_s| > \varepsilon\right\}\right) \le P(\{\rho_n = 0\}) + P(\{\rho_n < t\}).$$

Observe that $\left\{\rho_{n+1} = 0\right\} \subseteq \{\rho_n = 0\}$ and

$$\lim_{n\to\infty} \{\rho_n = 0\} = \bigcap_{n=1}^{\infty} \{\rho_n = 0\} = \{\omega : |Y_0(\omega)| \ge n,\ \forall n \in \mathbb{N}\}$$

which has zero probability. Therefore $P\{\{\rho_n = 0\}\} \to 0$ as $n \to \infty$. In addition, $P(\{\rho_n < t\}) \to 0$ as $n \to \infty$ because $\rho_n \to \infty$ a.s.. Hence, $\mathbf{b}\mathbb{L}$ is dense in $\mathbb{L}$, so we may suppose that $Y \in \mathbf{b}\mathbb{L}$. Define $Z_t = \lim_{\substack{u \to t \\ u > t}} Y_u$ so that $Z \in \mathbb{D}$. For $\varepsilon > 0$, we define

$$\begin{aligned} \tau_0^\varepsilon &= 0, \\ \tau_{n+1}^\varepsilon &= \inf\left\{t : t > \tau_n^\varepsilon \text{ and } \left|Z_t - Z_{\tau_n^\varepsilon}\right| > \varepsilon\right\}. \end{aligned}$$

Since Z is càdlàg, then τ_n^ε *are stopping times converging to* ∞ *a.s.* as $n \to \infty$. Let $Z^\varepsilon = \sum_n Z_{\tau_n^\varepsilon} \mathbf{1}_{[\tau_n^\varepsilon, \tau_{n+1}^\varepsilon)}$. It is immediate to see that $|Z - Z^\varepsilon| \le \varepsilon$ for any $t \in \mathbb{R}_+$ and any $\omega \in \Omega$, so Z^ε converge uniformly to Z as $\varepsilon \to 0$. Let

$$U^\varepsilon = Y_0 \mathbf{1}_{\{0\}} + \sum_{j=0}^{\infty} Z_{\tau_j^\varepsilon} \mathbf{1}_{\left(\tau_j^\varepsilon, \tau_{j+1}^\varepsilon\right]} \in \mathbf{b}\mathbb{L}.$$

Then, U^ε converges to $Y_0\mathbf{1}_{\{0\}} + Z_- = Y$ in *ucp* as $\varepsilon \to 0$, where $Z_- := \lim_{\substack{u\to t\\ u<t}} Z_u$. Indeed, for any $(t,\omega) \in \mathbb{R}_+ \times \Omega$, t belongs to a unique interval $\left(\tau^\varepsilon_{j_0}(\omega), \tau^\varepsilon_{j_0+1}(\omega)\right]$ for some j_0. Furthermore, if $t \in \left(\tau^\varepsilon_{j_0}(\omega), \tau^\varepsilon_{j_0+1}(\omega)\right]$ then, from the definition of the stopping times τ^ε_j,

$$\left|Z_s(\omega) - Z_{\tau^\varepsilon_{j_0}(\omega)}(\omega)\right| \leq \varepsilon$$

a.s. for any $s < t$ with $s \in \left(\tau^\varepsilon_{j_0}(\omega), \tau^\varepsilon_{j_0+1}(\omega)\right]$. Hence it also holds that

$$\left|(Z_-)_s(\omega) - Z_{\tau^\varepsilon_{j_0}(\omega)}(\omega)\right| \leq \varepsilon$$

with s satisfying the same hypotheses. This implies that $\left\{\sup_{0\leq s\leq t} |U^\varepsilon_s - (Z_-)_s| > \varepsilon\right\}$ has probability zero. Finally, define $Y^{n,\varepsilon} \in \mathbf{S}$ as

$$Y^{n,\varepsilon} = Y_0\mathbf{1}_{\{0\}} + \sum_{j=0}^{n} Z_{\tau^\varepsilon_j}\mathbf{1}_{\left(\tau^\varepsilon_j\wedge n, \tau^\varepsilon_{j+1}\wedge n\right]},$$

which can be made arbitrary close to $Y \in \mathbf{b}\mathbb{L}$ taking ε small and n large enough. ■

In the sequel we will exchangeably use the symbols $H \cdot X$ and $\int HdX$ to denote the Itô stochastic integral.

Theorem 1.22 *If X is a semimartingale, then $J_X : \mathbf{S}_{ucp} \longrightarrow \mathbb{D}_{ucp}$ is continuous.*

Remark 1.23 P. E. Protter introduces semimartingales as the processes for which J_X : $\mathbf{S}_{ucp} \longrightarrow \mathbb{D}_{ucp}$ is continuous. Had we taken his approach, the previous theorem would be empty of content. He shows later on that semimartingales, that is, process for which $J_X : \mathbf{S}_{ucp} \longrightarrow \mathbb{D}_{ucp}$ is continuous, admit a decomposition as in (1.15) ([P05, Chapter III, Theorem 1]).

Definition 1.24 *Let X be a (continuous) semimartingale. The linear map*

$$J_X : \mathbb{L}_{ucp} \longrightarrow \mathbb{D}_{ucp}$$

*obtained as the extension by density of $J_X : \mathbf{S}_{ucp} \longrightarrow \mathbb{D}_{ucp}$ is called the **Itô integral**.*

Given any stopping time τ we define

$$\int_0^\tau YdX := (Y \cdot X)_\tau.$$

It can be shown that $(\mathbf{1}_{[0,\tau]}Y) \cdot X = (Y \cdot X)^\tau = Y \cdot X^\tau$. If there exists a stopping time ζ_X such that the semimartingale X is defined only on the stochastic intervals $[0, \zeta_X)$, then we may define the Itô integral of Y with respect to X on any interval $[0,\tau]$ such that $\tau < \zeta_X$ by means of $Y \cdot X^\tau$.

Remark 1.25 Given the integrator semimartingale X, there are a few considerations on particular cases that deserve being pointed out:

(i) Suppose that X has paths of finite total variation on compact subsets of $\mathbb{R}_+$ (that is, $M = 0$ in the decomposition (1.15)). In this situation, the stochastic integral $\int HdX$ has paths of finite total variation on compact subsets of $\mathbb{R}_+$ and it is indistinguishable from the Riemann-Stieltjes integral, computed path by path ([P05, Chapter II Theorem 17]).

(ii) If X is a local martingale, then $\int HdX$ is also a local martingale ([P05, Chapter III Theorem 33]).

(iii) A corollary of the previous two points is that the stochastic integral with respect to a semimartingale is a semimartingale.

(iv) Brownian motions and the Martingale Representation Theorem. Suppose that $X = B$, a real valued Brownian motion. Then, the integral $\int HdB$ is a local martingale. Conversely, the Martingale Representation Theorem ([O03, Theorem 4.3.4]) asserts that given an arbitrary L^2-martingale M with respect to the canonical filtration $\mathcal{F}^B$ induced by the Brownian motion B, there exists a unique $\mathcal{F}_t$-adapted process H such that $E[\int_0^t H_s^2 ds] < \infty$ and

$$M_t = E[M_0] + \int_0^t H_s dB_s.$$

We now show how the Itô integral can be approximated by finite sums that replace the original integrand by a simple predictable process that tends to it in a sense that we make precise in the following definition.

Definition 1.26 *A sequence $\{\sigma_n\}_{n\in\mathbb{N}}$ of random partitions converging to the identity is a sequence where each σ_n is a finite family of stopping times $\{\tau_j^n\}_{j=0,\ldots,k_n}$ such that*

1. $0 \leq \tau_0^n \leq \tau_1^n \leq \ldots \leq \tau_{k_n}^n$.
2. $\lim_{n\to\infty} \tau_{k_n}^n = \infty$ *a.s.*
3. $\|\sigma_n\| = \sup_k \left|\tau_{k+1}^n - \tau_k^n\right| \to 0$ *a.s. as* $n \to \infty$.

Proposition 1.27 (Approximation property, [P05, Chapter II Theorem 21]) *Let X be a semimartingale and Y a process in $\mathbb{L}$. Let $\{\sigma_n\}_{n\in\mathbb{N}}$ be a sequence of random partitions converging to the identity. Consider*

$$Y^{\sigma_n} = \sum_{i=0}^{k_n-1} Y_{\tau_i^n} \mathbf{1}_{(\tau_i^n, \tau_{i+1}^n]}.$$

Then, as $n \to \infty$,

$$\int_0^t Y^{\sigma_n} dX = \sum_{i=0}^{k_n-1} Y_{\tau_i^n} (X_t^{\tau_{i+1}^n} - X_t^{\tau_i^n}) \xrightarrow[ucp]{} \int_0^t Y dX. \tag{1.18}$$

Alternative approach: Itô integral with respect to L^2-martingales

We are now going to briefly summarize an alternative approach to the Itô integral of a process with respect to a semimartingale that some authors use (see for example [CW90], [LG97], and [IW89]). In the following paragraphs, however, we are not going to deal with the most general case and we are going to assume that the semimartingale $X = X_0 + V + M$ admits a decomposition (1.15) where the local martingale M is actually a L^2-martingale. Since the integral with respect to the finite variation process V may be defined in terms of the Riemann-Stieltjes integral, as far as the definition of the integral with respect to X is concerned we only need to deal with the L^2-martingale M.

Let $\mathcal{R}$ the family of sets of $\mathbb{R}_+ \times \Omega$ the form $\{0\} \times A_0$ and $(s,t] \times A_s$, where $s < t$, $A_0 \in \mathcal{F}_0$, and $A_s \in \mathcal{F}_s$. $\mathcal{R}$ is called the family of ***predictable rectangles***. The σ-algebra of $[0,T] \times \Omega$ generated by $\mathcal{R}$ is called the ***predictable*** σ-algebra and is denoted by $\mathcal{P}$. Let $\mathcal{M}_2^{rc}$ denote the space of $L^2(\Omega, P)$ right-continuous martingales $M : \mathbb{R}_+ \times \Omega \to \mathbb{R}$. Given $M \in \mathcal{M}_2^{rc}$, we define a set function μ_M on $\mathcal{R}$ by $\mu_M((s,t] \times A_s) = E[\mathbf{1}_{A_s}(M_t - M_s)^2]$ for $A_s \in \mathcal{F}_s$ and $s < t \leq T$, and $\mu_M(\{0\} \times A_0) = 0$ for $A_0 \in \mathcal{F}_0$. If M is right-continuous L^2-martingale then μ_M extends to a unique measure on $\mathcal{P}$, called the Doléans measure ([CW90, Section 2.8]). We will continue denoting this measure by μ_M. It can be proved ([CW90, Theorem 4.2]) that, for any $U \in \mathcal{P}$,

$$\mu_M(U) = E\left[\int_0^T \mathbf{1}_U d[M,M]_s\right], \tag{1.19}$$

where $[M,M] : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is the quadratic variation of M (see the next subsection 1.3.2) and the integral in the right hand side of (1.19) is a pathwise integral which must be understood in the Riemann-Stieltjes sense.

We will say that a process H is a ***$\mathcal{R}$-simple process*** process if it can be written as a finite linear combination of indicator functions of predictable rectangles. That is, if H can be written as

$$H = \sum_{k=1}^m d_k \mathbf{1}_{\{0\}\times A_0^k} + \sum_{i=1}^p c_i \mathbf{1}_{(s_i,t_i]\times A_i} \tag{1.20}$$

with $c_i \in \mathbb{R}$, $A_i \in \mathcal{F}_{s_i}$, $s_i < t_i$ in $\mathbb{R}_+$ for $1 \leq i \leq p$, $p \in \mathbb{N}$, and $d_k \in \mathbb{R}$, $A_0^k \in \mathcal{F}_0$ for $1 \leq k \leq m$, $m \in \mathbb{N}$. The Itô integral $\int HdM$ is then naturally defined as

$$\int HdM := \sum_{i=1}^p c_i \mathbf{1}_{A_i}(M_{t_i} - M_{s_i}).$$

It can be checked that the Itô integral does not depend on the particular representation (1.20) of H. Moreover, if $\mathcal{E}$ denotes the space of $\mathcal{R}$-simple process, the map $\mathcal{E} \ni H \mapsto \int HdM$ satisfies the isometry

$$E\left[\left(\int HdM\right)^2\right] = \int_{\mathbb{R}_+\times\Omega} H^2 d\mu_M.$$

On the other hand, $\mathcal{E}$ is dense in the Hilbert space $L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, \mu_M)$ ([CW90, Lemma 2.4]). If we regard $L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, \mu_M)$ and $L^2(\Omega, \mathcal{F}, P)$ as Hilbert spaces, then the map $H \mapsto \int HdM$ is a linear isometry from the dense subspace $\mathcal{E}$ of $L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, \mu_M)$ into $L^2(\Omega, \mathcal{F}, P)$, and hence can be uniquely extended to a linear isometry from $L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, \mu_M)$ into $L^2(\Omega, \mathcal{F}, P)$. For $H \in L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, \mu_M)$, we define $\int H dM$ as the image of H under this isometry. As usual, the expression $\int_0^t H dM$ denotes $\int \mathbf{1}_{[0,t]} H dM$.

1.3.2 The quadratic variation and the Stratonovich integral

All along this section $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ will be a (continuous) semimartingale such that $X_0 = 0$ a.s. In order to simplify the notation, the symbol $[X, X]$ will denote the quadratic variation $[X, X]^{(2)}$ introduced in (1.8). However, we are no longer going to consider the quadratic variation as a quantity computed path-by-path but as a process properly introduced in terms of stochastic integrals. In other words, the quadratic variation will be a process as a whole. Concretely,

Definition 1.28 *Let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ be a càglàd semimartingale. The **quadratic variation process** of X, denoted by $[X, X]_{t \geq 0}$, is defined as*

$$[X, X] = X^2 - 2 \int X dX.$$

Apparently, this definition seems to have nothing in common with the notion of the quadratic variation of a path we gave in (1.8). However, the next proposition shows that the two definitions are actually closer to each other than we might have thought at first sight. In particular, both coincide if the limit in (1.8) is properly taken.

Proposition 1.29 *Let $\{\sigma_n\}_{n \in \mathbb{N}}$ be a sequence of random partitions tending to the identity as in Definition 1.26. Then*

$$[X, X] = X_0^2 + \lim_{\substack{n \to \infty \\ ucp}} \sum_{i=0}^{k_n - 1} \left(X^{\tau_{i+1}^n} - X^{\tau_i^n} \right)^2. \tag{1.21}$$

Additionally, $[X, X]$ is an increasing process.

Proof. Let $\{\tau^n\}_{n \in \mathbb{N}}$ be a sequence of stopping times converging to ∞ a.s.. First of all, observe that if Z is any real process, then $Z^{\tau_n} \to_{ucp} Z$ as $n \to \infty$. Indeed, if $s \in [0, t]$ and $\varepsilon > 0$, then

$$\{ |Z^{\tau_n} - Z|_s > \varepsilon \} \subseteq \{ \tau_n < s \} \subseteq \{ \tau_n < t \}.$$

Hence

$$P\left(\left\{ \sup_{0 \leq s \leq t} |Z^{\tau_n} - Z|_s > \varepsilon \right\} \right) \leq P(\{\tau_n < t\})$$

for any $t \in \mathbb{R}_+$. In addition, $P(\{\tau_n < t\}) \to 0$ in probability as $n \to \infty$ because $\tau_n \to \infty$ a.s.. Once we have made this observation, it is immediate to see that the telescopic sum

$$\left(X^2\right)^{\tau_{k_n}^n} = \sum_{i=0}^{k_n - 1} \left\{ \left(X^2\right)^{\tau_{i+1}^n} - \left(X^2\right)^{\tau_i^n} \right\},$$

converges in *ucp* to $X^2 - X_0^2$ because $\tau_{k_n}^n \to \infty$ a.s. as $n \to \infty$. Now, the sum

$$\sum_{i=0}^{k_n-1} X_{\tau_i^n}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)$$

converges in *ucp* to $\int X dX$ by (1.18). Since $b^2 - a^2 - 2a(b-a) = (b-a)^2$ and $X_{\tau_i^n}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right) = X^{\tau_i^n}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)$, we have

$$\begin{aligned}
\sum_{i=0}^{k_n-1}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)^2 &= \sum_{i=0}^{k_n-1}\left\{\left(X^2\right)^{\tau_{i+1}^n} - \left(X^2\right)^{\tau_i^n} - 2X^{\tau_i^n}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)\right\} \\
&= \sum_{i=0}^{k_n-1}\left\{\left(X^2\right)^{\tau_{i+1}^n} - \left(X^2\right)^{\tau_i^n}\right\} - 2\sum_{i=0}^{k_n-1} X^{\tau_i^n}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right) \\
&\underset{ucp}{\longrightarrow} X^2 - X_0^2 - 2\int X dX.
\end{aligned}$$

Finally, note that if $s < t$, then the approximating sums (1.21) include more non-negative terms, so $[X, X]$ is non-decreasing (see [P05, Chapter II Theorem 22]). ■

Corollary 1.30 *If X is a continuous semimartingale and has paths of finite variation then $[X, X] = X_0^2$.*

Proof. We have

$$\begin{aligned}
\sum_{i=0}^{k_n-1}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)_t^2 &\leq \sup_{i=0,\ldots,k_n-1}\left|X^{\tau_{i+1}^n} - X^{\tau_i^n}\right|_t \sum_{i=0}^{k_n-1}\left|X^{\tau_{i+1}^n} - X^{\tau_i^n}\right|_t \\
&\leq \sup_{i=0,\ldots,k_n-1}\left|X^{\tau_{i+1}^n} - X^{\tau_i^n}\right|_t \mu([0,t])
\end{aligned}$$

where $\mu([0,t])$ is the total variation of X on the interval $[0,t]$ introduced in (1.14). But

$$\sup_{i=0,\ldots,k_n-1}\left|X^{\tau_{i+1}^n} - X^{\tau_i^n}\right|_t$$

tends to 0 as $\|\sigma_n\| \to 0$ because X is continuous, and therefore uniformly continuous, on $[0,t]$.

■

We now state Levy's characterization of Brownian motion for the sake of completeness:

Theorem 1.31 (Levy's Theorem) *A stochastic process $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is a standard Brownian motion if and only if it is a continuous local martingale and $[X, X]_t = t$.*

The next proposition is a consequence of the definition of the stochastic integral and the representation of the quadratic variation of a semimartingale as the limit of the approximative sums in Proposition 1.29. It will be useful in order to prove Itô's formula. See [P05, Chapter II Theorem 30] for its proof.

Proposition 1.32 *Let $\{\sigma_n\}_{n\in\mathbb{N}}$ be a sequence of random partitions tending to the identity as in Definition 1.26, Y a càglàd adapted process, and X a càglàd semimartingales. Then*

$$\int Y_s d[X,X]_s = \lim_{\substack{n\to\infty \\ ucp}} \sum_{i=0}^{k_n-1} Y_{\tau_i^n}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)^2.$$

Definition 1.33 *Let X, Y be two continuous semimartingales such that $X_0 = Y_0 = 0$. The* ***quadratic covariation*** *of X and Y is defined by*

$$[X,Y] = \frac{1}{2}\left([X+Y, X+Y] - [X,X] - [Y,Y]\right). \tag{1.22}$$

Equivalently,

$$[X,Y] := XY - \int X dY - \int Y dX. \tag{1.23}$$

The paths of the quadratic covariation $[X,Y]$ of two continuous semimartingales have finite total variation. Moreover, for any sequence of random partitions tending to the identity $\{\sigma_n\}_{n\in\mathbb{N}}$, we have

$$[X,Y] = \lim_{\substack{n\to\infty \\ ucp}} \sum_{i=0}^{k_n-1} \left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right)\left(Y^{\tau_{i+1}^n} - Y^{\tau_i^n}\right) \tag{1.24}$$

([P05, Chapter II Theorem 23]). Observe that the quadratic covariation is the process we have to add to the usual integration by parts formula in the context of Itô integrals, which consequently no longer holds. In order to remedy this situation, the Stratonovich integral arises.

Definition 1.34 *Given X and Y two semimartingales we define the* ***Stratonovich integral*** *of Y along X as*

$$\int Y\delta X = \int Y dX + \frac{1}{2}[Y,X].$$

Using the limit expressions for the Itô integral and for the quadratic variation, we have that, for any sequence of random partitions tending to the identity $\{\sigma_n\}_{n\in\mathbb{N}}$, the Stratonovich integral can be written as

$$\int Y\delta X = \lim_{\substack{n\to\infty \\ ucp}} \sum_{i=0}^{k_n-1} \frac{\left(Y^{\tau_{i+1}^n} + Y^{\tau_i^n}\right)}{2}\left(X^{\tau_{i+1}^n} - X^{\tau_i^n}\right) \tag{1.25}$$

([P05, Chapter V Theorem 26]). In other words, the difference between the Itô and the Stratonovich integrals can be expressed by saying that the former is obtained by taking a lower endpoint approximation of the integrand on $(\tau_i^n, \tau_{i+1}^n]$ while the latter uses the middle point.

Remark 1.35 (Integration by parts) From the definition of the Stratonovich integral is immediate to check that

$$\int Y\delta X = YX - \int X\delta Y.$$

That is, the Stratonovich integral satisfies the usual integration by parts formula.

1.3.3 The Itô formula

The Itô formula is the cornerstone of stochastic calculus and represents the analog of the chain rule in the context of Itô integration.

Theorem 1.36 (Itô formula) *Let $X^1, \ldots, X^p$ be p continuous semimartingales and $f \in C^2(\mathbb{R}^p)$. Then,*

$$f(X_t^1, \ldots, X_t^p) = f(X_0^1, \ldots, X_0^p) + \sum_{i=1}^{p} \int_0^t \frac{\partial f}{\partial x^i}(X_s^1, \ldots, X_s^p) dX_s^i + \frac{1}{2} \sum_{i,j=1}^{p} \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s^1, \ldots, X_s^p) d[X^i, X^j]_s. \tag{1.26}$$

Proof. (Sketch) We are going to prove only the case $p = 1$. Consider $\{\pi_n\}_{n \in N}$ a sequence of nested deterministic partitions of $[0, t]$ tending to the identity. That is, π_n is a family of finite times such that $0 = t_0^n \leq t_1^n \leq \ldots \leq t_{k_n}^n = t$ and $\|\pi_n\| \to 0$ as $n \to \infty$. Then,

$$f(X_t) - f(X_0) = \sum_{i=0}^{k_n - 1} \left(f\left(X_{t_{i+1}^n}\right) - f\left(X_{t_i^n}\right) \right).$$

Using Taylor's formula, we have

$$f\left(X_{t_{i+1}^n}\right) - f\left(X_{t_i^n}\right) = f'\left(X_{t_i^n}\right)\left(X_{t_{i+1}^n} - X_{t_i^n}\right) + \frac{1}{2} f''\left(X_{t_i^n}\right)\left(X_{t_{i+1}^n} - X_{t_i^n}\right)^2 + R\left(X_{t_i^n}, X_{t_{i+1}^n}\right) \tag{1.27}$$

where $|R(x, y)| \leq r(|x - y|)(x - y)^2$ and $r : \mathbb{R} \to \mathbb{R}$ is an increasing function such that $\lim_{u \to 0} r(u) = 0$. Equation (1.27) is valid for any $f \in C^2(\mathbb{R})$ defined on a compact set. Since there is no reason for X_t to be contained in a compact set, we define the stopping times

$$\tau_m = \inf\{t \ : \ |X_t| \geq m\}.$$

Then the process X^{τ_m} is bounded by m and, if Itô's formula is valid for X^{τ_m} for each m, then it is also valid for X. Therefore, we may assume without loss of generality that X takes values on a compact set.

On the one hand, we have that

$$\lim_{n \to \infty} \sum_{i=0}^{k_n - 1} f'\left(X_{t_i^n}\right)\left(X_{t_{i+1}^n} - X_{t_i^n}\right) = \int_0^t f'(X)\, dX$$

in probability by Proposition 1.27. On the other hand,

$$\lim_{n \to \infty} \sum_{i=0}^{k_n - 1} f''\left(X_{t_i^n}\right)\left(X_{t_{i+1}^n} - X_{t_i^n}\right)^2 = \int_0^t f''(X)\, d[X, X]$$

in probability by 1.32. It remains to consider the third sum $\sum_{i=0}^{k_n-1} R(X_{t_i^n}, X_{t_{i+1}^n})$. This sum is bounded above by

$$\sup_{i=0,\dots,k_n-1} r\left(\left|X_{t_{i+1}^n} - X_{t_i^n}\right|\right)\left(\sum_{i=0}^{k_n-1}\left(X_{t_{i+1}^n} - X_{t_i^n}\right)^2\right)$$

and since $\sum_{i=0}^{k_n-1}(X_{t_{i+1}^n} - X_{t_i^n})^2$ converges in probability to $[X,X]_t$, the last term will tend to 0 if $\lim_{n\to\infty}\sup_{i=0,\dots,k_n-1} r(\left|X_{t_{i+1}^n} - X_{t_i^n}\right|) = 0$. However, the path X_s is a continuous function on $[0,t]$ and hence uniformly continuous. Since $\lim_{n\to\infty}\sup_{i=0,\dots,k_n-1}\left|X_{t_{i+1}^n} - X_{t_i^n}\right| = 0$ by hypothesis, the result follows from the properties of r. ∎

Sometimes (1.26) is presented using a symbolic differential notation, namely:

$$df(X_t^1,\dots,X_t^p) = \sum_{i=1}^p \frac{\partial f}{\partial x^i}(X_s^1,\dots,X_s^p)dX_s^i + \frac{1}{2}\sum_{i,j=1}^p \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s^1,\dots,X_s^p)d[X^i,X^j]_s. \quad (1.28)$$

In this expression one sees that the difference with the chain rule of standard calculus lies in the second term of (1.28) that involves the quadratic variation and that disappears for processes with finite total variation. This hence allows us to visualize Itô calculus as a generalization of standard calculus. When we will later on work globally on manifolds, formula (1.28) will force us to use the second order tangent bundle instead of the standard tangent bundle.

The analog of equality (1.26) for the Stratonovich integral is

$$f(X_t^1,\dots,X_t^p) = f(X_0^1,\dots,X_0^p) + \sum_{i=1}^p \int_0^t \frac{\partial f}{\partial x^i}(X_s^1,\dots,X_s^p)\delta X_s^i$$

or, in differential notation,

$$df(X_t^1,\dots,X_t^p) = \sum_{i=1}^p \frac{\partial f}{\partial x^i}(X_s^1,\dots,X_s^p)\delta X_s^i$$

which coincides with the standard chain rule.

1.3.4 Stochastic differential equations

Let $X = (X^1,\dots,X^p)$ be p real valued continuous semimartingales and $f : \mathbb{R}_+ \times \mathbb{R}^q \to \mathbb{R}^q$ a smooth function. A ***(strong) solution*** of the ***Itô stochastic differential equation***

$$d\Gamma^i = \sum_{j=1}^p f_j^i(t,\Gamma)dX^j \quad (1.29)$$

with initial condition the random vector $\Gamma_0 = (\Gamma_0^1,\dots,\Gamma_0^q)$ is a stochastic process $\Gamma_t = (\Gamma_t^1,\dots,\Gamma_t^q)$ such that

$$\Gamma_t^i - \Gamma_0^i = \sum_{j=1}^p \int_0^t f_j^i(t,\Gamma)dX^j.$$

There is an existence and uniqueness theorem for the solutions of (1.29) which appears in the literature formulated with a great variety of slightly different statements, depending on the hypotheses one imposes on the defining function $f : \mathbb{R}_+ \times \mathbb{R}^q \to \mathbb{R}^q$. Usually, f is only assumed to be locally Lipschitz. Recall that a function $f : \mathbb{R}_+ \times \mathbb{R}^q \to \mathbb{R}^q$ is ***Lipschitz*** if there exists a finite constant K such that

(i) $\|f(t,x) - f(t,y)\| \leq K \|x - y\|$, for each $t \in \mathbb{R}_+$, and $x, y \in \mathbb{R}^q$

(ii) $t \mapsto f(t,x)$ is right continuous with left limits (càdlàg) for each $x \in \mathbb{R}^q$.

The function f is said to be ***locally Lipschitz*** if there exists and increasing sequence of open sets Λ_k such that $\bigcup_k \Lambda_k = \mathbb{R}^q$ and f is Lipschitz with a constant K_k on each Λ_k. For example, if f has continuous but not necessarily bounded derivatives then f is locally Lipschitz. However, we suppose f is differentiable in (1.29) because we prefer to stay within the category of C^∞ functions when considering stochastic differential equations on manifolds. When $f : \mathbb{R}_+ \times \mathbb{R}^q \to \mathbb{R}^q$ is only locally Lipschitz, the solutions of (1.29) are defined up to an ***explosion time***. More explicitly,

Theorem 1.37 ([P05, Chapter V Theorem 38 and 39]) *For any $x \in \mathbb{R}^q$, there exists a stopping time $\zeta(x, \cdot) : \Omega \to \mathbb{R}_+$ and a unique time-continuous solution $X(t, \omega, x)$ of (1.29) with initial condition x defined on the time interval $[0, \zeta(x,\omega))$. Additionally, $\limsup_{t \to \zeta(x,\omega)} \|X_t(\omega)\| = \infty$ a.s. on $\{\zeta < \infty\}$ and X is smooth in x on the open set $\{x \mid \zeta(x,\omega) > t\}$. Finally, the solution X is a semimartingale.*

Had we taken as initial condition in Theorem 1.37 any $\mathcal{F}_0$-measurable random variable X_0 instead of $X_0 = x \in \mathbb{R}^q$ a.s., Theorem 1.37 would remain true. Obviously, an analogous result can be formulated for Stratonovich stochastic differential equations.

Examples 1.38

(i) **The Langevin equation** provides a model for the motion of a particle subjected to damping caused by microscopic collisions. Let V the velocity of the particle in question; the Langevin equation (Langevin (1908)) is:

$$dV_t = -bV_t dt + \sigma dB_t.$$

The solution of the Langevin equation is given by the ***Ornstein-Uhlenbeck*** process

$$V_t = e^{-bt} V_0 + \int_0^t e^{-b(t-s)} \sigma dB_s.$$

(ii) **The geometric Brownian motion** is used as a model for the behavior of the underlying asset in the Black-Scholes formula for the price of an option:

$$dX_t = \mu X_t dt + \sigma X_t dB_t.$$

Its solution is given by

$$X_t = X_0 \exp\left(\mu t - \frac{\sigma^2}{2} t + \sigma B_t\right)$$

for any $\mathcal{F}_0$-measurable random variable X_0.

SDEs and partial differential equations. The Feynman-Kac formula

Let B be a d-dimensional Brownian motion and $V_0, V_1, \ldots, V_d : \mathbb{R}^n \to \mathbb{R}^n$ a collection of Lipschitz autonomous vector fields in $\mathbb{R}^n$ with Lipschitz partial derivatives. Consider the associated (Stratonovich) stochastic differential equation

$$\delta X_t = V_0(X_t)dt + \sum_{i=1}^{d} V_i(X_t)\delta B_t^i. \tag{1.30}$$

The strong solutions of (1.30) are called ***diffusions***. It can be shown that these solutions have moments of every order, have no explosions (that is, $\zeta(x, \cdot) = \infty$ for any $x \in \mathbb{R}^q$), and are homogeneous Markov processes ([P05, Chapter V Theorem 32], [O03, Theorem 7.1.2]). This allows us to associate a transformation semigroup P_t to the stochastic differential equation (1.30) that acts on the space of measurable functions $f : \mathbb{R}^n \to \mathbb{R}$:

$$(P_t f)(x) := E[f(X_t^x)], \tag{1.31}$$

where X_t^x is the (unique) solution of (1.30) such that $X_0^x = x$, a.s.. If $C_0(\mathbb{R}^n)$ denotes the separable Banach space of real valued functions that tend to zero at infinity endowed with the $\|\cdot\|_\infty$ norm, then P_t is a continuous semigroup made of contractions of $(C_0(\mathbb{R}^n), \|\cdot\|_\infty)$, i.e., $\|P_t f\|_\infty \leq \|f\|_\infty$, and the map $(t, f) \mapsto P_t f$ is continuous.

The ***infinitesimal generator*** L of the transformation semigroup P_t is defined by

$$Lf := \lim_{t\to 0} \frac{1}{t}\left(P_t f - f\right).$$

and can be explicitly written down in terms of the vector fields that determine the stochastic differential equation, namely:

$$Lf = V_0[f] + \frac{1}{2}\sum_{i=1}^{d} V_i[V_i[f]] \tag{1.32}$$

By [Y71, Chapter IX Theorem 3.1], the domain $\mathrm{Dom}(L)$ of definition of L is a dense subspace of $(C_0(\mathbb{R}^n), \|\cdot\|_\infty)$ and

$$\frac{d}{dt}P_t f = P_t L f = L P_t f \tag{1.33}$$

for any $f \in \mathrm{Dom}(L)$. This equation and (1.31) show that $u(t, x) = E[f(X_t^x)] = P_t f(x)$ is a solution of the second order, parabolic, linear partial differential equation

$$\begin{cases} \frac{\partial}{\partial t} u & = & Lu \\ u(0, x) & = & f(x) \end{cases} \tag{1.34}$$

with initial condition $f \in C_0(\mathbb{R}^n)$. This statement provides a probabilistic interpretation of the solutions of the partial differential equation (1.34) in terms of an expectation. This is of much importance at the time of numerically computing those solutions using Montecarlo methods.

Example 1.39 The heat equation. Suppose that $X_t^x : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ is a d-dimensional Brownian motion starting at $x \in \mathbb{R}^d$. That is, in our usual notation, $X_t^x = x + B_t$ with $B_0 = 0$. The infinitesimal generator associated to its transition semigroup is just the Laplacian, $L = \Delta$. Consequently, in view of (1.34), $u(t,x) = E[f(X_t^x)] = E^x[f(B_t)]$ is a solution of the heat equation

$$\frac{\partial}{\partial t} u = \Delta u$$

with initial $f \in C_0(\mathbb{R}^n)$ condition at time $t = 0$.

The previous example is the simplest particular case of the ***Feynman-Kac representation formula*** which, under certain technical hypotheses, provides solutions to the Cauchy problems

$$\pm \tfrac{\partial}{\partial t} u_{\pm}(t,x) \quad = \quad \Delta u_{\pm} - k(x) u_{\pm} + g(t,x), \quad t \in [0,T], x \in \mathbb{R}^n, \tag{1.35}$$

restricted to $u_+(0,x) = f(x)$ at time $t = 0$ for the forward equation (sing $+$ in (1.35)) or $u_-(T,x) = f(x)$ at terminal time $t = T$ for the backward equation (sing $-$ in (1.35)). For example, if B denotes an d-dimensional Brownian motion, the function u_- solution of backward heat equation admits the stochastic representation

$$u(t,x) = E^x \left[f\left(B_{T-t}\right) \mathrm{e}^{-\int_0^{T-t} k(s, X_s^{t,x}) ds} + \int_0^{T-t} g(t+u, B_u)\, \mathrm{e}^{-\int_0^u k(B_s) ds} du \right] \tag{1.36}$$

$t \in [0,T]$, $x \in \mathbb{R}^d$, provided that $f : \mathbb{R}^d \to \mathbb{R}$, $k : \mathbb{R}^d \to \mathbb{R}$, and $g : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ are continuous, $u_- : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is continuous and of class $C^{1,2}$ on $[0,T) \times \mathbb{R}^n$, and

$$\max_{0 \le t \le T} |u_-(t,x)| + \max_{0 \le t \le T} |g(t,x)| \le K\, \mathrm{e}^{a\|x\|^2}$$

for some constants $K > 0$ and $0 < a < 1/(2Td)$ (see [KS91, Chapter 4 Theorem 4.2]). (1.36) is known as the ***Feynman-Kac representation formula***. Other versions of the Feynman-Kac representation formula can be found in the literature (see for instance [O03, Theorem 8.2.1]).

Hypoelliptic diffusions and the Kolmogorov-Fokker-Planck equations

A stochastic process X^x with initial condition x is called ***hypoelliptic*** when, for each time $t > 0$ for which X^x is defined, the law of X_t^x is absolutely continuous with respect to the Lebesgue measure and the corresponding probability density function $p_t(x,y)$ is a smooth function on y and t. A stochastic differential equation is called hypoelliptic when each of its solutions of the form X^x is hypoelliptic and the density $p_t(x,y)$ is smooth on all its entries.

Hypoellipticity is usually very hard to check and it is one of the main subjects of the so called *Malliavin calculus*. There exist sufficient conditions for hypoellipticity, the most famous of them being ***Hörmander's condition***. Hörmander's result says that if the vector fields that generate the stochastic differential equation (1.30) are such that the Lie algebra generated by the family

$$\{\{V_i\}_{i=1,\dots,d}, \{[V_i, V_j]\}_{i,j=0,\dots,d}, \{[[V_j, V_j], V_k]\}_{i,j,k=0,\dots,d}, \dots\}$$

spans the whole space $\mathbb{R}^n$ at each point $x \in \mathbb{R}^n$, then the associated diffusion is hypoelliptic. If the spanning condition holds only at the point x, then only the process X^x is hypoelliptic ([N95, Theorem 2.3.2 and 2.3.3]).

When the diffusion is hypoelliptic, the equation (1.33) can be used to write down the differential equation that determines the time evolution of the pdf $p_t(x,y)$. Indeed, endow the space of smooth compactly supported functions $C_c(\mathbb{R}^n)$ with the inner product defined by

$$\langle f, g\rangle = \int_{\mathbb{R}^n} f(y)g(y)dy.$$

Since we are considering a hypoelliptic diffusion, we can write (1.31) as

$$\left(P_t f\right)(x) := E[f(X_t^x)] = \int_{\mathbb{R}^n} f(y)p_t(x,y)dy.$$

Now, for any $f \in C_c(\mathbb{R}^n)$, we have that

$$\frac{d}{dt}\left(P_t f\right)(x) = \int_{\mathbb{R}^n} f(y)\frac{\partial}{\partial t}p_t(x,y)dy = \langle \frac{\partial}{\partial t}p_t(x,\cdot), f\rangle.$$

On the other hand,

$$\left(P_t(Lf)\right)(x) = \int_{\mathbb{R}^n} Lf(y)p_t(x,y)dy = \langle p_t(x,\cdot), Lf\rangle = \langle L^* p_t(x,\cdot), f\rangle.$$

If we now use (1.33) and the fact that the last two equalities are valid for any $f \in C_c(\mathbb{R}^n)$, we conclude that

$$\frac{\partial}{\partial t}p_t(x,y) = L^* p_t(x,y),$$

which is known as the ***forward Kolmogorov*** or ***Fokker-Planck equation*** for the time evolution of the pdf $p_t(x,y)$. In this expression, L^* denotes the adjoint operator of L.

1.4 Manifold valued semimartingales and SDEs

All along this section M will denote a finite dimensional, second-countable, locally compact Hausdorff (and hence paracompact) manifold. The content of this section is mainly based on the excellent book by M. Émery [E89] which, in contrast with other references, gathers the essential tools of stochastic differential geometry using a more modern geometrical language.

Definition 1.40 *A continuous M-valued stochastic process Γ defined on the filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ is called a* ***semimartingale*** *if, for any $f \in C^\infty(M)$, the process $f \circ \Gamma$ is a real valued semimartingale.*

If M and N are two manifolds and $\phi : M \to N$ a smooth mapping between them, it is clear that if Γ is a M-valued semimartingale, then $\phi \circ \Gamma$ is a N-valued semimartingale. The property of being a semimartingale *can be localized* in the sense that one can find a family of charts that covers the manifold such that, roughly speaking, the M-valued process Γ is a semimartingale whenever its restrictions to the charts are (real valued) semimartingales of the type previously studied. We now make this statement more precise.

Definition 1.41 *A stopping time τ is said to be* ***predictable*** *if there exists an increasing sequence $\{\tau_n\}_{n\in\mathbb{N}}$ of stopping times such that $\tau_n \to \tau$ a.s..*

The main tool in proving our statement is the following proposition, which states that given a M-valued continuous process and an open covering of M, one can construct a family of stopping times that can be used to confine the process to the open sets of the covering.

Proposition 1.42 *Let $\{U_k\}_{k\in\mathbb{N}}$ be a countable open covering of coordinate neighborhoods of M and Γ an adapted and continuous M-valued process defined on $[0,\zeta)$, with ζ a predictable stopping time. Then, there exist predictable stopping times τ_n with $\tau_0 = 0$, $\tau_n \leq \tau_{n+1}$ and $\sup_n \tau_n = \zeta$ a.s. such that, on each of the sets $[\tau_n, \tau_{n+1}] \cap \{\tau_{n+1} > \tau_n\}$, Γ takes its values in one set $U_{k(n)}$.*

Proof. Denote by E the set of all predictable stopping times σ for which there exist a finite sequence of predictable stopping times $0 = \tau_0 \leq \tau_1 \leq ... \leq \tau_p = \sigma \leq \zeta$ with the same property as in the statement: Γ takes its values in $U_{k(n)}$ on the interval $[\tau_n, \tau_{n+1}] \cap \{\tau_{n+1} > \tau_n\}$, $n = 1, ..., p$. The proof will be broken down into four steps.

1. $E \neq \emptyset$. Take $U_1 \in \{U_k\}_{k\in\mathbb{N}}$ and consider V an open set such that $\overline{V} \subset U_1$. Define $\tau_0 = 0$ and
$$\tau_1 = \inf\{t > 0,\ \Gamma_t \notin V\}.$$
Obviously, $\tau_1 = 0$ on $\{\Gamma_0 \notin V\}$ by continuity of the process Γ. It is a well know result that τ_1 is a stopping time if Γ is adapted. Moreover, it is predictable. To see this, we can build an increasing sequence of nested closed sets $\{D_n\}_{n\in\mathbb{N}}$ such that $D_n \subseteq V$ for any $n \in \mathbb{N}$ and $d(\partial D_n, \partial V) \longrightarrow 0$ as $n \to \infty$, where $d(\cdot,\cdot)$ is the distance induced on V from the Euclidean one on $\mathbb{R}^m$ by means of the coordinate homeomorphism. Finally, the process Γ takes its values in $\overline{V} \subset U_1$ on the set $[\tau_0, \tau_1] \cap \{\tau_1 > \tau_0\}$. So $\sigma = \tau_1 \in E$.

2. The essential supremum R of E exists. Recall that if $\{Z_j : j \in I\}$ is a family of random variables defined on $(\Omega, \mathcal{F}, P)$, where the index set I may be arbitrary, then there exists a countable subset J of I such that the random variable $Z^* : \Omega \to [0,\infty]$ defined by $Z^* = \sup_{\alpha\in J} Z_\alpha$ is the ***essential supremum***. That is, satisfies the following two properties: (i) $P(\{Z_j \leq Z^*\}) = 1$ for each $j \in I$; (ii) If $\widetilde{Z} : \Omega \to [0,\infty]$ is another random variable satisfying all the previous identities in place of Z^*, then $P(\{Z^* \leq \widetilde{Z}\}) = 1$. The random variable Z^* is called the essential supremum of $\{Z_j : j \in I\}$ relative to P. It is determined by the properties (i) and (ii) uniquely up to a P-null set (see [BP06] and [F06]). In the present situation $I = E$. So the essential supremum R of E exists. Moreover, R may be written as $R = \sup_{\alpha\in J} \sigma_\alpha$ for a countable family $\{\sigma_\alpha\}_{\alpha\in J}$ of stopping times. Using a bijection map $\phi : \mathbb{N} \to J$ if necessary, we may replace J with $\mathbb{N}$ and write $R = \sup_{n\in\mathbb{N}} \sigma_n$. Hence R is again a stopping time. Explicitly, being R the supremum of $\{\sigma_n\}_{n\in\mathbb{N}}$, it holds that $\{R > t\} = \cup_{n\in\mathbb{N}} \{\sigma_n > t\}$ and since each $\{\sigma_n > t\}$ belongs to $\mathcal{F}_t$, so does $\{R > t\}$. Actually, R is predictable, since it may be expressed as $R = \lim_k \widetilde{\sigma}_k$, where $\widetilde{\sigma}_k = \max\{\sigma_1, ..., \sigma_k\}$.

3. The essential supremum R of E is ζ a.s.. Suppose this is not the case. That is, $P(\{R<\zeta\}) > 0$. Observe that R is finite on the set $A := \{R<\zeta\}$. Consequently, Γ_R is well defined on A. There must exist a U_k such that
$$P(A \cap \{\Gamma_R \in U_k\}) > 0,$$
because if $P(A \cap \{\Gamma_R \in U_k\}) = 0$ for any $k \in \mathbb{N}$, then
$$0 < P(A) = P(A \cap \{\Gamma_R \in M\}) \leq \sum_{k\in\mathbb{N}} P(A \cap \{\Gamma_R \in U_k\}) = 0,$$
which is clearly a contradiction. Indeed, as U_k is an open coordinate neighborhood, there must exist another open set V such that $\overline{V} \subsetneq U_k$ and
$$P(A \cap \{\Gamma_R \in V\}) > 0.$$
Define the stopping time S as follows:
$$S = \begin{cases} R & \text{if } R=\infty \text{ or } \Gamma_R \notin V \\ \inf\{t > R : \Gamma_t \notin V\} & \text{if } \Gamma_R \in V. \end{cases}$$
Then, S has non-zero probability of exceeding R and S is predictable because, on the one hand, R was already predictable and, on the other hand, exit stopping times from open sets are predictable. Since R is the essential supremum of E, there exists a sequence $\{\sigma_j\}_{j\in\mathbb{N}} \subseteq E$ such that $\sigma_j \to R$ a.s. as $j \to \infty$. Define $\{\widetilde{\sigma}_j\}_{j\in\mathbb{N}}$ as $\widetilde{\sigma}_j = \sigma_j \wedge S$. It is clear that $\widetilde{\sigma}_j \in E$ because if $\{\tau_{j_i}\}_{i=0,\ldots,p_j}$ are such that $0 = \tau_{j_0} \leq \tau_{j_1} \leq \ldots \leq \tau_{j_p} = \sigma_j$ and the property of the statement holds, then $0 = \tau_{j_0} \wedge S \leq \tau_{j_1} \wedge S \leq \ldots \leq \tau_{j_p} \wedge S = \widetilde{\sigma}_j$ and the property keeps holding for this new sequence. But now $\widetilde{\sigma}_j \to S$ a.s. as $j \to \infty$, which contradicts the fact that R is the essential supremum. Therefore $R = \zeta$ a.s..

4. To sum up, E contains a sequence $\{\sigma_j\}_{j\in\mathbb{N}}$ such that $\sigma_j \to \zeta$ a.s.. In order to prove the statement of the proposition, it suffices to interpolate this sequence by inserting between σ_j and σ_{j+1} the predictable stopping times
$$\tau_{(j+1)_0} \vee \sigma_j, \ldots, \tau_{(j+1)_p} \vee \sigma_j$$
where, as in the previous point, $\tau_{(j+1)_0}, \ldots, \tau_{(j+1)_p}$ are given such that $0 = \tau_{(j+1)_0} \leq \tau_{(j+1)_1} \leq \ldots \leq \tau_{(j+1)_p} = \sigma_{j+1} \in E$.

■

We finish this brief introduction on manifold valued semimartingales by stating a couple of results that, essentially, show that the property of being a semimartingale is local ([E89, page 23]).

Proposition 1.43 *Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a M-valued process and $\{U_k\}_{k\in\mathbb{N}}$ a countable family of coordinate neighborhoods that cover M, with coordinate maps $\{U_k; x_k^i, i = 1, \ldots, m\}$. Let $\{\tau_n\}_{n\in\mathbb{N}}$ be the sequence of stopping times given by Proposition 1.42. Then,*
$$\Gamma \text{ is an } M\text{-valued semimartingale} \Longleftrightarrow x_{k(n)}^i \circ Y^n \text{ is a real valued semimartingale.}$$

Proposition 1.44 *Let Γ be a M-valued semimartingale and $N \subseteq M$ a regular submanifold such that Γ takes its values on N. Then Γ is a N-valued semimartingale.*

1.4.1 The quadratic variations of a M-valued semimartingale

As we saw when dealing with real valued processes, the quadratic variation is a tool of paramount importance. Since the definition that we introduced involves the sum and the multiplication of real numbers, it cannot be trivially rephrased for manifold valued processes. This difficulty is solved by replacing the multiplication by the use of a bilinear form on the manifold; obviously, this yields a different quadratic variation for each choice of bilinear form.

Theorem 1.45 (The quadratic variation) *Let Γ be a continuous M-valued semimartingale. There exists a unique linear mapping*

$$b \longmapsto \int b\langle d\Gamma, d\Gamma\rangle$$

from the space of bilinear forms $\mathcal{T}^2(M)$ on M to the space of real valued continuous processes with finite total variation that is uniquely determined by the following two properties: for any $f, g \in C^\infty(M)$ and any $b \in \mathcal{T}^2(M)$,

$$\int (fb)\langle d\Gamma, d\Gamma\rangle = \int (f\circ\Gamma)\, d\left(\int b\langle d\Gamma, d\Gamma\rangle\right) \tag{1.37a}$$

$$\int (\mathbf{d}f\otimes\mathbf{d}g)\langle d\Gamma, d\Gamma\rangle = \langle f\circ\Gamma, g\circ\Gamma\rangle. \tag{1.37b}$$

The real valued process $\int b\langle d\Gamma, d\Gamma\rangle$ is called the ***b-quadratic variation*** of Γ or the ***integral*** of b along Γ. We need some auxiliary lemmas before the proof of this theorem.

Lemma 1.46 ([E89, Lemma 2.23]) *There exist a natural number $n \geq \dim(M)$ and a finite family of functions $\left\{h^1, ..., h^n\right\} \subset C^\infty(M)$ such that every bilinear form $b \in \mathcal{T}^2(M)$ can be expressed as a finite sum*

$$b = \sum_{i,j=1}^{n} b_{ij}\mathbf{d}h^i \otimes \mathbf{d}h^j, \tag{1.38}$$

where $b_{ij} \in C^\infty(M)$ are n^2 functions that depend on b.

Proof. By Whitney's theorem, M can be imbedded in $\mathbb{R}^n$ for some $n \in \mathbb{N}$. In this context, there exist n smooth functions $\left\{h^1, ..., h^n\right\} \subset C^\infty(M)$, a partition of the unity $(\phi_k)_{k\in K}$, and a family $(J_k)_{k\in K}$ of subsets of $\{1, ..., n\}$ such that, for each k, the family $(h^i)_{i\in J_k}$ is a system of local coordinates in a neighborhood of the support of ϕ_k. If $b \in \mathcal{T}^2(M)$, then $\phi_k b$ is a bilinear form compactly supported on a neighborhood with local coordinates $(h^i)_{i\in J_k}$. Therefore, it can be written as $\phi_k b = \sum_{i,j\in J_k} b_{kij}\mathbf{d}h^i \otimes \mathbf{d}h^j$, where $\operatorname{supp}(b_{kij}) \subseteq \operatorname{supp}(\phi_k)$. The lemma is proved by setting $b_{ij} = \sum_k b_{kij}$ (locally finite sum). ■

Lemma 1.47 (from [E89, Lemma 3.10]) *Let Γ be a continuous M-valued semimartingale. Let u_j, f^j, and g^j, $j = 1, ..., r$, be a finite family of functions on M such that the bilinear form $\sum_{j=1}^r u_j\mathbf{d}f^j \otimes \mathbf{d}g^j$ is identically zero. Then,*

$$\sum_{j=1}^{n} \int (u_j\circ\Gamma)\, d\left[f^j\circ\Gamma, g^j\circ\Gamma\right] = 0.$$

Proof. Let $\{U_k\}_{k\in\mathbb{N}}$ be a countable coordinate neighborhood covering of M and $\{\tau_n\}_{n\in\mathbb{N}}$ the stopping time sequence associated to Γ and subordinated to $\{U_k\}_{k\in\mathbb{N}}$ given by Proposition 1.42. Observe that Γ^τ is defined on the whole interval $[0,\infty)$, so $\sup_n \tau_n = \infty$. Using the Proposition A.2 in the appendix,

$$\sum_{j=1}^{r}\int (u_j\circ\Gamma)\, d\left[f^j\circ\Gamma, g^j\circ\Gamma\right] = \lim_{\substack{ucp\\ n\to\infty}}\sum_{i=1}^{n-1}\sum_{j=1}^{r}\int \mathbf{1}_{(\tau_i,\tau_{i+1}]}\,(u_j\circ\Gamma)\, d\left[f^j\circ\Gamma, g^j\circ\Gamma\right].$$

Now, on $(\tau_i,\tau_{i+1}]$, Γ takes values on $(U_{k(i)}; x^j_{k(i)},\ j=1,...,m)$, so

$$\begin{aligned}&\int \mathbf{1}_{(\tau_i,\tau_{i+1}]}\,(u_j\circ\Gamma)\, d\left[f^j\circ\Gamma, g^j\circ\Gamma\right]\\ &=\sum_{r,s=1}^{m}\int \mathbf{1}_{(\tau_i,\tau_{i+1}]}\left(u_j\frac{\partial f^j}{\partial x^r_{k(i)}}\frac{\partial g^j}{\partial x^s_{k(i)}}\right)(\Gamma)\, d\left[x^r_{k(i)}(\Gamma), x^s_{k(i)}(\Gamma)\right].\end{aligned}$$

Since $\sum_{j=1}^{r} u_j \frac{\partial f^j}{\partial x^r_{k(i)}}\frac{\partial g^j}{\partial x^s_{k(i)}} = 0$ on $U_{k(i)}$ by hypothesis, we conclude

$$\sum_{j=1}^{r}\int (u_j\circ\Gamma)\, d\left[f^j\circ\Gamma, g^j\circ\Gamma\right] = 0.$$

■

Proof (of Theorem 1.45).

Let $b\in\mathcal{T}^2(M)$. By Lemma 1.46, b can be written as a finite sum $b=\sum_{i,j=1}^{n} b_{ij}\mathbf{d}h^i\otimes\mathbf{d}h^j$ for some functions $h^i, h^j\in C^\infty(M)$, $i,j\in\{1,...,n\}$. If (1.37a) and (1.37b) must hold then, necessarily,

$$\int b\langle d\Gamma, d\Gamma\rangle = \sum_{ij=1}^{n}\int b_{ij}(\Gamma)\, d\langle h^i\circ\Gamma, h^j\circ\Gamma\rangle. \tag{1.39}$$

Equation (1.39) allows to define the integral $\int b(d\Gamma, d\Gamma)$ in a unique way. Nevertheless, we need to check that this definition does not depend on the particular decomposition $\sum_{i,j=1}^{n} b_{ij}\mathbf{d}h^i\otimes\mathbf{d}h^j$ of the bilinear form $b\in\mathcal{T}^2(M)$. But this is exactly the content of Lemma 1.47. ■

Remark 1.48 It can be also checked that the b-quadratic variation of Γ depends only on the symmetric part of b. Therefore, if b is antisymmetric, $\int b\langle d\Gamma, d\Gamma\rangle = 0$.

1.4.2 Second order vectors and forms

As we announced when we presented the Itô formula, the new terms that appear in the Itô stochastic version of the chain rule require, when we move onto stochastic global analysis, the use of geometric structures that generalize those used in the standard calculus on manifolds. The most important of these generalized structures are the second order bundles and their associated sections, the so-called second order forms and vectors fields.

Definition 1.49 *Let $p \in M$. A* ***tangent vector at*** p ***of order two with no constant term*** *is a differential operator $L : \mathcal{C}^{\infty}(M) \longrightarrow \mathbb{R}$ that satisfies $L\left[f^3\right](p) = 3f(p)\,L\left[f^2\right](p) - 3f^2(p)\,L[f](p)$. The vector space of tangent vectors of order two at p will be denoted by $\tau_p M$. The manifold $\tau M := \bigcup_{p\in M} \tau_p M$ is referred to as the* ***second order tangent bundle*** *of M. A* ***vector field of order two*** *is a smooth section of the bundle $\tau M \to M$. We denote the set of vector fields order two by $\mathfrak{X}_2(M)$.*

Remark 1.50 *Note that the (first order) tangent bundle TM of M is contained in τM. That is, a vector field $Y \in \mathfrak{X}(M)$ is a vector field of order two. Indeed, if $f \in C^{\infty}(M)$ and applying the Leibniz rule,*

$$3fY\left[f^2\right] - 3f^2Y[f] = 6f^2Y[f] - 3f^2Y[f] = 3f^2Y[f] = Y\left[f^3\right].$$

The following lemma provides various equivalent characterizations of the notion of second order vector field.

Lemma 1.51 *The following statements are equivalent:*

(i) *$L \in \mathfrak{X}_2(M)$ is a vector fields of order 2.*

(ii) *$L : C^{\infty}(M) \to C^{\infty}(M)$ is a $\mathbb{R}$-lineal differential operator of order (at most) 2 with no constant term. Equivalently, on any local coordinate neighborhood $\left(U; x^i, i = 1, ..., m\right)$, L can be written as*

$$L = \sum_{i=1}^{m} l^i \frac{\partial}{\partial x^i} + \sum_{i,j=1}^{m} l^{ij} \frac{\partial^2}{\partial x^i \partial x^j}$$

where $l^i, l^{ij} \in C^{\infty}(U)$, $i, j = 1, ..., m$ and $l^{ij} = l^{ji}$.

(iii) *For any smooth map $F = \left(f^1, ..., f^n\right) : M \to \mathbb{R}^n$ and $\phi : \mathbb{R}^n \to \mathbb{R}$,*

$$L[\phi \circ F] = \frac{\partial \phi}{\partial x^i}(F)\,L\left[f^i\right] + \frac{\partial^2 \phi}{\partial x^i \partial x^j}(F)\,\Upsilon_L\left(f^i, f^j\right),$$

where $\Upsilon_L\left(f^i, f^j\right) = \frac{1}{2}\left(L\left[f^i f^j\right] - f^i L\left[f^j\right] - f^j L\left[f^i\right]\right)$ is the so called ***"carré du champs"****.*

(iv) *$L[1] = 0$ and if $f \in C^{\infty}(M)$ is such that $f(p) = 0$ at $p \in M$, then $L\left[f^3\right](p) = 0$.*

Proof. The only non-trivial implication is $4 \Rightarrow 1$, which we are going to prove explicitly. The rest are left to the reader. Let $f, g, h \in C^{\infty}(M)$. The polarization formula

$$6fgh = (f + g + h)^3 - (g + h)^3 - (f + h)^3 - (f + g)^3 + f^3 + g^3 + h^3$$

and the hypothesis in **(iv)** give $L[fgh](p) = 0$ if $f(p) = g(p) = h(p) = 0$ at some $p \in M$. This implies that L is local. That is, if $f \in C^{\infty}(M)$ vanishes in a neighborhood U of p, taking $g = h$ such that $g(p) = 0$ and $g|_{M\setminus D} = 1$ outside a closed set $p \in D \subset U$, we see that $f = fgh$ and $L[f](p) = 0$. Let now $p \in M$ be a point and $f \in C^{\infty}(M)$ an arbitrary function. Take local coordinates $\left(U_p; x^i, i = 1, ..., m\right)$ on a neighborhood U_p of p such that $x(p) = 0$. Let

$\widetilde{f} : x(U_p) \subseteq \mathbb{R}^m \to \mathbb{R}$ be the smooth function defined by the relation $f = x \circ \widetilde{f}$. Now, in $\mathbb{R}^m$, we can write

$$\widetilde{f}(x) - \widetilde{f}(0) = \sum_{i=1}^{m} x^i \frac{\partial \widetilde{f}}{\partial x^i}(0) + \sum_{i,j=1}^{m} x^i x^j h_{ij}(x),$$

where $h_{ij}(x) = \int_0^1 (1-t) \frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(tx)\, dt$ is a smooth function on $\mathbb{R}^m$ such that $h_{ij}(0) = \frac{1}{2}\frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(0)$. Furthermore,

$$\begin{aligned}\widetilde{f}(x) - \widetilde{f}(0) = &\sum_{i=1}^{m} x^i \frac{\partial \widetilde{f}}{\partial x^i}(0) + \sum_{i,j=1}^{m} x^i x^j \frac{1}{2}\frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(0) \\ &+ \sum_{i,j=1}^{m} x^i x^j \left(h_{ij}(x) - \frac{1}{2}\frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(0)\right).\end{aligned}$$

Therefore, applying L to this last relationship

$$\begin{aligned}L[f](p) = &\sum_{i=1}^{m} L\left[x^i\right](p) \frac{\partial \widetilde{f}}{\partial x^i}(0) + \sum_{i,j=1}^{m} \frac{1}{2} L\left[x^i x^j\right](p) \frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(0) \\ &+ \sum_{i,j=1}^{m} L\left[x^i x^j \left(h_{ij}(x) - \frac{1}{2}\frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(0)\right)\right](p).\end{aligned}$$

The last term vanishes because each of the three functions that make it up vanish at p. Therefore,

$$L[f](p) = \sum_{i=1}^{m} l^i(p) \frac{\partial \widetilde{f}}{\partial x^i}(0) + \sum_{i,j=1}^{m} l^{ij}(p) \frac{\partial^2 \widetilde{f}}{\partial x^i \partial x^j}(0)$$

is a second order differential operator at p with no constant term. Its coefficients are $l^i(p) = L\left[x^i\right](p)$ and $l^{ij}(p) = \frac{1}{2} L\left[x^i x^j\right](p)$. ■

If $\left(\bar{x}^i; i = 1, ..., m\right)$ and $\left(x^j; j = 1, ..., m\right)$ are local charts on $U \subseteq M$, the change of coordinates formula for $L \in \mathfrak{X}_2(U)$ is

$$\bar{l}^i = \sum_{k=1}^{m} l^k \frac{\partial \bar{x}^i}{\partial x^k} + \sum_{k,j=1}^{m} l^{kj} \frac{\partial \bar{x}^i}{\partial x^k \partial x^j}, \tag{1.40a}$$

$$\bar{l}^{ij} = \sum_{l,r=1}^{m} l^{rs} \frac{\partial \bar{x}^j}{\partial x^r} \frac{\partial \bar{x}^i}{\partial x^s}. \tag{1.40b}$$

Equation (1.40a) shows that $\sum_{i=1}^{m} l^i \frac{\partial}{\partial x^i}$, that might be expected to be the first order part of L, does not transform as a tensor and hence it is not intrinsically defined: tangent vectors of order 2 have no intrinsic first order part. On the other hand, (1.40b) satisfies the change of coordinates formula for two times contravariant tensor fields. This shows that we can intrinsically associate to every $L \in \mathfrak{X}_2(M)$ a symmetric tensor $\hat{L} \in \mathfrak{X}(M) \otimes \mathfrak{X}(M)$. Moreover, for any $f, g \in C^{\infty}(M)$,

$$\hat{L}(\mathbf{d}f, \mathbf{d}g) = \Upsilon_L(f, g). \tag{1.41}$$

Since any element $\mathcal{T}_2(M)$ is uniquely defined by its action on pairs $(\mathbf{d}f, \mathbf{d}g)$ of exact one forms, (1.41) fully characterizes $\hat{L}$.

Concerning the vector fields of order two, it can be proved that any $X \in \mathfrak{X}_2(M)$ may be written as a finite sum of second order vector fields of the type YZ and W, where $Y, Z, W \in \mathfrak{X}(M)$.

Definition 1.52 *Let M, N be two manifolds and $\phi : M \to N$ a smooth map. For any $p \in M$, we define the* ***second order tangent map*** *at p*

$$\tau_p\phi : \tau_p M \longrightarrow \tau_{\phi(p)}N$$

as $\tau_p\phi(L)[f] = L[f \circ \phi]$, where $f \in C^\infty(N)$. The restriction of $\tau_p\phi$ to T_pM clearly coincides with $T_p\phi$.

Second order forms are smooth sections of the cotangent bundle of order two, defined as $\tau^*M := \bigcup_{p\in M} \tau_p^*M$, where τ_p^*M is the dual of τ_pM, $p \in M$. The set of second order forms is denoted by $\Omega_2(M)$.

Definition 1.53 *Let $f, g, h \in C^\infty(M)$ and $L \in \mathfrak{X}_2(M)$. We define $d_2 f \in \Omega_2(M)$ by $d_2 f(L) := L[f]$, and $\mathbf{d}f \cdot \mathbf{d}g \in \Omega_2(M)$ as $\mathbf{d}f \cdot \mathbf{d}g[L] := \Upsilon_L(f, g) = \frac{1}{2}(L[fg] - fL[g] - fL[f])$.*

It is easy to show that

$$\mathbf{d}f \cdot \mathbf{d}g\,[ZY] = \frac{1}{2}\left(Z[f]\,Y[g] + Z[g]\,Y[f]\right), \tag{1.42a}$$

$$\mathbf{d}f \cdot \mathbf{d}g\,[W] = 0 \tag{1.42b}$$

for any $Y, Z, W \in \mathfrak{X}(M)$.

More generally, let $\alpha_p, \beta_p \in T_p^*M$ and choose $f, g \in \mathcal{C}^\infty(M)$ two functions such that $\mathbf{d}f(p) = \alpha_p$ and $\mathbf{d}g(p) = \beta_p$. It is easy to check that $(\mathbf{d}f \cdot \mathbf{d}g)(p)$ does not depend on the particular choice of f and g above and hence we can write $\alpha_p \cdot \beta_p$ to denote $(\mathbf{d}f \cdot \mathbf{d}g)(p)$. We define $\alpha \cdot \beta \in \Omega_2(M)$ for any $\alpha, \beta \in \Omega(M)$ as

$$(\alpha \cdot \beta)(p) := \alpha(p) \cdot \beta(p).$$

This product is Abelian and $\mathcal{C}^\infty(M)$-bilinear. Furthermore, every second order form can be locally written as a finite sum of forms of the type $\mathbf{d}f \cdot \mathbf{d}g$ and $d_2 h$.

Theorem 1.54 *There exists a unique linear mapping $d_2 : \Omega(M) \to \Omega_2(M)$ that verifies*

$$d_2(\mathbf{d}f) = d_2 f$$
$$d_2(f\alpha) = \mathbf{d}f \cdot \alpha + f d_2\alpha,$$

where $f \in C^\infty(M)$ and $\alpha \in \Omega(M)$.

Proof. Whitney's Embedding Theorem guarantees that any one-form $\alpha \in \Omega(M)$ can be written as a finite sum

$$\alpha = \sum_{i=1}^{n} f_i \mathbf{d}h^i \tag{1.43}$$

for a finite family of functions $\{f_i, h_i; i = 1, ..., n\}$, $n \geq m$. If d_2 exists, $d_2\alpha$ is necessarily equal to

$$d_2\alpha = \sum_{i=1}^{n} \left(\mathbf{d}f_i \cdot \mathbf{d}h^i + f_i d_2 h^i\right), \tag{1.44}$$

which establishes the uniqueness. In order to prove the existence, we have to verify that the result obtained in (1.44) does not depend on the particular decomposition (1.43) of α. This is equivalent to showing that, if $\alpha = \sum_{i=1}^{n} f_i \mathbf{d}h^i = 0$, then the second order form $\theta = \sum_{i=1}^{n} \left(\mathbf{d}f_i \cdot \mathbf{d}h^i + f_i d_2 h^i\right)$ vanishes. That is, we have to check that $\theta(X) = 0$, where $X \in \mathfrak{X}_2(M)$ is any second order vector field of the form YZ or W, $Y, Z, W \in \mathfrak{X}(M)$. By (1.42b),

$$\theta(W) = \sum_{i=1}^{n} f_i d_2 h^i(W) = \alpha(W) = 0. \tag{1.45}$$

On the other hand, since $YZ - ZY = [Y, Z]$ is a vector field of order one, (1.45) implies that $\theta(YZ) = \theta(ZY)$. Therefore,

$$\begin{aligned} 2\theta(YZ) = \theta(YZ + ZY) &\overset{(1.42a)}{=} \sum_{i=1}^{n} \left(Y[f_i] Z\left[h^i\right] + Y\left[h^i\right] Z[f_i] + f_i YZ\left[h^i\right] + h^i ZY[f_i]\right) \\ &= \sum_{i=1}^{n} \left(Y\left[f_i Z\left[h^i\right]\right] + Z\left[f_i Y\left[h^i\right]\right]\right) = Y[\alpha(Z)] + Z[\alpha(Y)] = 0. \end{aligned}$$

■

1.4.3 Itô and Stratonovich integrals

Definition 1.55 *Let $\theta : \mathbb{R}_+ \times \Omega \to \tau^* M$ be a process. One says that θ is* ***locally bounded*** *if the set $\{\theta_s : 0 \leq s \leq t\}$ is relatively compact in $\tau^* M$ for any t and any ω a.s.. We will say that θ* ***covers*** *(or* ***is over****) $\Gamma : \mathbb{R}_+ \times \Omega \to M$ if $\pi_{\tau^* M}(\theta) = \Gamma$, where $\pi_{\tau^* M} : \tau^* M \to M$ is the canonical projection.*

Theorem 1.56 ([E89, Theorem 6.24]) *Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a continuous semimartingale. There exists a unique linear map $\theta \mapsto \int \langle \theta, d\Gamma \rangle$ from the space of predictable, $\tau^* M$-valued processes θ over Γ to the space of real valued continuous semimartingales that is uniquely determined by the equalities*

$$\int \langle d_2 f \circ \Gamma, d\Gamma \rangle = f(\Gamma) - f(\Gamma_0), \tag{1.46a}$$

$$\int \langle K\theta, d\Gamma \rangle = \int K d\left(\int \langle \theta, d\Gamma \rangle\right), \tag{1.46b}$$

for any $f \in C^\infty(M)$ and any locally bounded, predictable real process K.

Definition 1.57 *The real valued semimartingale $\int \langle \theta, d\Gamma \rangle$ is called the* ***Itô integral of the process*** *θ* ***along*** *Γ. Usually, $\theta = \alpha \circ \Gamma$ for some $\alpha \in \Omega_2(M)$. In this case, we simply write $\int \langle \alpha, d\Gamma \rangle$ and we will call it the* ***Itô integral of the form*** *α* ***along*** *Γ.*

The proof of Theorem 1.56 follows the same pattern as the definition of the b-quadratic variation in Theorem 1.45. Equations (1.46a) and (1.46b) determine uniquely the value of the integral since by Whitney's Embedding Theorem it can be proved that any continuous locally bounded process $\theta : \mathbb{R}_+ \times \Omega \to \tau^* M$ over Γ can be written as a finite sum

$$\theta_t(\omega) = \sum_{\lambda=1}^{r} (K_\lambda)_t(\omega)\, d_2 g^\lambda(\Gamma_t(\omega)), \tag{1.47}$$

where $\{g^\lambda; \lambda = 1, ..., r\} \subset C^\infty(M)$ is a finite family of smooth functions and $\{K_\lambda; \lambda = 1, ..., r\}$ is a finite family of predictable, locally bounded real valued processes. The proof is completed by showing that the resulting expression does not depend on the particular decomposition (1.47) chosen for θ.

Proposition 1.58 *The Itô integral has the following properties:*

(i) *Let $\phi : M \to N$ be a smooth mapping between two manifolds M and N. Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a semimartingale and $\theta : \mathbb{R}_+ \times \Omega \to \tau^* N$ a continuous $\tau^* N$ valued process over $\phi \circ \Gamma$. Then,*

$$\int \langle \theta, d(\phi \circ \Gamma) \rangle = \int \langle \tau^* \phi(\theta), d\Gamma \rangle .$$

(ii) *For any $f, g \in C^\infty(M)$,*

$$\int \langle \mathbf{d}f \cdot \mathbf{d}g, d\Gamma \rangle = \frac{1}{2}[f \circ \Gamma, g \circ \Gamma]. \tag{1.48}$$

(iii) *If $\Gamma : \mathbb{R}_+ \times \Omega \to M$ is a deterministic smooth curve, $\int \langle \theta, d\Gamma \rangle = \int \left\langle \mathbf{R}(\theta)\left(\dot{\Gamma}\right), dt \right\rangle$, where $\mathbf{R} : \tau^* M \to T^* M$ is the dual of the inclusion map $TM \to \tau M$.*

Proof. **(i)** It suffices to verify that the linear mapping $I_\theta : \theta \mapsto \int \langle \tau^* \phi(\theta), d\Gamma \rangle$ has the properties $I_{d_2 f} = f(\phi \circ \Gamma) - f(\phi \circ \Gamma_0)$ and $I_{K\theta} = \int K dI_\theta$ for any $f \in C^\infty(N)$ and any continuous real valued process K. The second property is obvious and the first one is a consequence of the fact that $d_2(f \circ \phi) = \tau^* \phi(d_2 f)$. **(ii)** Using the equivalent expression (1.23) for the quadratic variation of two real processes and the definition of $\mathbf{d}f \cdot \mathbf{d}g$, we have

$$\begin{aligned}
\int \langle \mathbf{d}f \cdot \mathbf{d}g, d\Gamma \rangle &= \int \left\langle \frac{1}{2}(d_2(fg) - f d_2 g - g d_2 f), d\Gamma \right\rangle \\
&= \frac{1}{2}\left(fg \circ \Gamma - fg \circ \Gamma_0 - \int (f \circ \Gamma)\, d(g \circ \Gamma) - \int (g \circ \Gamma)\, d(f \circ \Gamma) \right) \\
&= \frac{1}{2}[f \circ \Gamma, g \circ \Gamma].
\end{aligned}$$

(iii) The proof consists just of checking that, if Γ is a deterministic smooth curve, the definition

$$\int \langle \theta, d\Gamma \rangle = \int \left\langle \mathbf{R}(\theta)(\dot{\Gamma}), dt \right\rangle$$

verifies (1.46a) and (1.46b). The first equality follows from the fact that $\mathbf{R}\left(d_2 f\right) = \mathbf{d}f$ and the second one is obvious. ■

The Stratonovich integral is defined for processes that take values in the ordinary cotangent bundle T^*M. It is completely characterized by two relationships equivalent to (1.46a) and (1.46b) with the operator d_2 replaced by the ordinary external differential $\mathbf{d}$.

Theorem 1.59 *Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a semimartingale. There exists a unique linear map $\theta \mapsto \int \langle \theta, \delta\Gamma \rangle$ from the space of continuous, T^*M-valued semimartingales θ over Γ to the space of real valued continuous semimartingales that is uniquely determined by the equalities*

$$\int \langle \mathbf{d}f \circ \Gamma, \delta\Gamma \rangle = f\left(\Gamma\right) - f\left(\Gamma_0\right), \tag{1.49a}$$

$$\int \langle K\theta, \delta\Gamma \rangle = \int K\delta\left(\int \langle \theta, \delta\Gamma \rangle\right), \tag{1.49b}$$

*for any $f \in C^\infty\left(M\right)$ and any continuous real semimartingale K. The real valued semimartingale $\int \langle \theta, \delta\Gamma \rangle$ is called the **Stratonovich integral of the process θ along** Γ. Usually, $\theta = \alpha \circ \Gamma$ for some $\alpha \in \Omega\left(M\right)$. In this case, we simply write $\int \langle \alpha, \delta\Gamma \rangle$ and we will call it the **Stratonovich integral of the form α along** Γ.*

The proof of Theorem 1.59 consists, once more, of using Whitney's Embedding Theorem to write down the continuous locally bounded process $\theta : \mathbb{R}_+ \times \Omega \to T^*M$ over Γ as the finite sum

$$\theta_t\left(\omega\right) = \sum_{\lambda=1}^{r} \left(K_\lambda\right)_t\left(\omega\right) \mathbf{d}g^\lambda\left(\Gamma_t\left(\omega\right)\right), \tag{1.50}$$

where $\left\{g^\lambda ; \lambda = 1, ..., r\right\} \subset C^\infty\left(M\right)$ is a finite family of smooth functions and $\left\{K_\lambda ; \lambda = 1, ..., r\right\}$ is a finite family of continuous, locally bounded real valued processes. The uniqueness of the Stratonovich integral follows from using (1.49a) and (1.49b) on the decomposition (1.50). The existence follows from showing that the resulting expression is independent of the particular decomposition (1.50) of θ.

When we integrate forms instead of arbitrary T^*M or τ^*M valued processes over Γ, the operator d_2 yields a convenient relation between the Itô and the Stratonovich integrals.

Proposition 1.60 *Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a semimartingale and $\alpha \in \Omega(M)$ a one-form. Then,*

$$\int \langle \alpha, \delta\Gamma \rangle = \int \langle d_2\alpha, d\Gamma \rangle. \tag{1.51}$$

Proof. We start as in (1.43) by writing

$$\alpha = \sum_{i=1}^{n} f_i \mathbf{d}h^i \tag{1.52}$$

for a finite family of functions $\left\{f_i, h_i ; i = 1, ..., n\right\}$, $n \geq m$. Then, on one hand

$$\int \langle \alpha, \delta\Gamma \rangle = \sum_{i=1}^{n} \int \langle f_i \mathbf{d}h^i, \delta\Gamma \rangle = \sum_{i=1}^{n} \int f_i \delta\left(\int \langle \mathbf{d}h^i, \delta\Gamma \rangle\right) = \sum_{i=1}^{n} \int f_i(\Gamma)\delta h^i(\Gamma).$$

On the other hand

$$\int \langle d_2\alpha, d\Gamma\rangle = \sum_{i=1}^{n} \int \langle d_2\left(f_i \mathbf{d} h^i\right), d\Gamma\rangle = \sum_{i=1}^{n} \int \langle \mathbf{d} f_i \cdot \mathbf{d} h^i + f_i d_2 h^i, d\Gamma\rangle$$
$$= \frac{1}{2}\sum_{i=1}^{n}[f_i \circ \Gamma, h^i \circ \Gamma] + \sum_{i=1}^{n} \int f_i(\Gamma) dh^i(\Gamma) = \sum_{i=1}^{n} \int f_i(\Gamma)\delta h^i(\Gamma),$$

as required. ■

Remark 1.61 The Itô and Stratonovich integrals for real valued processes introduced in definitions 1.24 and 1.34 can be recovered from the theorems-definition 1.56 and 1.59 by taking in (1.46b) (respectively (1.49b)) a real valued semimartingale $X : \mathbb{R}_+ \times \Omega \to R$ as Γ, another real valued semimartingale Y as K, and $\theta := d_2 t \circ X$ (respectively $\theta := dt \circ X$). With those choices, by (1.46a) (respectively (1.49a)) we have that

$$\int \langle K\theta, d\Gamma\rangle = \int Y d\left(\int \langle d_2 t \circ X, dX\rangle\right) = \int Y d\left(X - X_0\right) = \int Y dX,$$
$$\int \langle K\theta, \delta\Gamma\rangle = \int Y \delta\left(\int \langle \mathbf{d} t \circ X, \delta X\rangle\right) = \int Y \delta\left(X - X_0\right) = \int Y \delta X.$$

1.4.4 Stochastic differential equations on manifolds

We start by defining the Stratonovich stochastic differential equations. This concept is based on the Stratonovich integral of forms over a semimartingale and on the notion of Stratonovich operator that we now introduce.

Definition 1.62 *Let M and N be two manifolds. A **Stratonovich operator** from M to N is a family $\{S(x,y)\}_{x\in M, y\in N}$ of maps such that $S(x,y) : T_xM \to T_yN$ is a linear mapping that depends smoothly on its two entries. Equivalently, we require that*

$$S : TM \times N \longrightarrow TN$$

is a smooth map.

Let $S^*(x,y) : T_y^*N \to T_x^*M$ be the adjoint of $S(x,y)$. Let X be a M-valued semimartingale.

Definition 1.63 *We say that a N-valued semimartingale Y is a solution of the the **Stratonovich stochastic differential equation***

$$\delta Y = S(X,Y)\delta X \tag{1.53}$$

if, for any $\alpha \in \Omega(N)$, the following equality between Stratonovich integrals holds:

$$\int \langle \alpha, \delta Y\rangle = \int \langle S^*(X,Y)\alpha, \delta X\rangle.$$

It can be shown that given a semimartingale X in M, a $\mathcal{F}_0$-measurable random variable Y_0, and a Stratonovich operator S from M to N, there are a predictable (and hence progressively measurable) stopping time ζ and a solution Y of (1.53) with initial condition Y_0 defined on the set $\{(t,\omega) \in \mathbb{R}_+ \times \Omega \mid t \in [0, \zeta(\omega))\}$ that has the following maximality and uniqueness property: if ζ' is another stopping time such that $\zeta' < \zeta$ and Y' is another solution defined on $\{(t,\omega) \in \mathbb{R}_+ \times \Omega \mid t \in [0, \zeta'(\omega))\}$, then Y' and Y coincide in this set. Moreover, if ζ is finite, then Y explodes at time ζ. This means that the path $(Y_t)_{t\in[0,\zeta)}$ is not contained in any compact subset of N ([E89, Theorem 7.21]).

Stochastic differential equations from the Itô integration point of view require the notion of a ***Schwartz operator*** whose construction we now briefly review.

Definition 1.64 *Let M, N be two manifolds. Given $x \in M$ and $y \in N$, a linear map from $\tau_x M$ into $\tau_y N$ is called a **Schwartz morphism** whenever $f(T_xM) \subset T_yN$ and $\widehat{f(L)} = \left(f|_{T_xM} \otimes f|_{T_xM}\right)(\widehat{L})$, for any $L \in \tau_x M$. The symbol $\widehat{L} \in T_xM \otimes T_xM$ denotes the unique symmetric element that is intrinsically attached to $L \in \tau_x M$ (Eq. (1.41)). A **Schwartz operator** from M to N is a family $\{f(x,y)\}_{x\in M, y\in N}$ such that $f(x,y) : \tau_x M \to \tau_y N$ is a Schwartz operator that depends smoothly on its two entries. That is,*

$$f : \tau M \times N \longrightarrow \tau N$$

is a smooth map.

Let $f^*(x,y) : \tau_y^* N \to \tau_x^* M$ be the adjoint of $f(x,y)$ and X a M-valued semimartingale.

Definition 1.65 *We say that a N-valued semimartingale is a solution of **the Itô stochastic differential equation***

$$dY = f(X,Y)dX \tag{1.54}$$

if, for any $\alpha \in \Omega_2(N)$, the following equality between Itô integrals holds:

$$\int \langle \alpha, dY \rangle = \int \langle f^*(X,Y)\alpha, dX \rangle.$$

Analogously, there is an existence and uniqueness result for the solutions of these stochastic differential equations analogous to that one available for Stratonovich differential equations ([E89, Theorem 6.41]).

Given a Stratonovich operator S from M to N, there exists a unique Schwartz operator $f : \tau M \times N \to \tau N$ associated to S and defined as follows. Let $\gamma(t) = (x(t), y(t)) \in M \times N$ be a smooth curve that verifies $S(x(t), y(t))(\dot{x}(t)) = \dot{y}(t)$ for all t. We define $f(x(t), y(t))\left(L_{\ddot{x}(t)}\right) := \left(L_{\ddot{y}(t)}\right)$, where the second order differential operators $\left(L_{\ddot{x}(t)}\right) \in \tau_{x(t)} M$ and $\left(L_{\ddot{y}(t)}\right) \in \tau_{y(t)} N$ are defined as $\left(L_{\ddot{x}(t)}\right)[h] := \frac{d^2}{dt^2} h\left(x\left(t\right)\right)$ and $\left(L_{\ddot{y}(t)}\right)[g] := \frac{d^2}{dt^2} g\left(y\left(t\right)\right)$ for any $h \in C^\infty(M)$ and $g \in C^\infty(N)$. This relation completely determines f since the vectors of the form $L_{\ddot{x}(t)}$ span $\tau_{x(t)} M$. Moreover, the Itô and Stratonovich equations $\delta Y = S(X,Y)\delta X$ and $dY = f(X,Y)dX$ are equivalent: they have the same solutions for the same initial $\mathcal{F}_0$-measurable random variables Y_0.

2

Stochastic Hamiltonian dynamical systems

The generalization of classical mechanics to the context of stochastic dynamics has been an active research subject ever since K. Itô introduced the theory of stochastic differential equations in the 1950s (see for instance [N67, B81, Y81, YZ82, MZ84, TZ97, TZ97a, A03, CD06, BO07, BO07a], and references therein). The motivations behind some pieces of work related to this field lay in the hope that a suitable stochastic generalization of classical mechanics should provide an explanation of the intrinsically random effects exhibited by quantum mechanics within the context of the theory of diffusions . In other instances the goal is establishing a framework adapted to the handling of mechanical systems subjected to random perturbations or whose parameters are not precisely determined and are hence modeled as realizations of a random variable.

Most of the pieces of work in the first category use a class of processes that have a stochastic derivative introduced in [N67] and that has been subsequently refined over the years. This derivative can be used to formulate a real valued action and various associated variational principles whose extremals are the processes of interest.

The approach followed in this chapter is closer to the one introduced in [B81] in which the action has its image in the space of real valued processes and the variations are taken in the space of processes with values in the phase space of the system that we are modeling. Our work in this chapter can be actually seen as a generalization of some of the results in [B81] in the following directions:

(i) We make extensive use of the global stochastic analysis tools introduced by P. A. Meyer [M81, M82] and L. Schwartz [S82] to handle non-Euclidean phase spaces. This feature not only widens the spectrum of systems that can be handled but it is also of paramount importance at the time of reducing them with respect to the symmetries that they may eventually have (see Chapter 3); indeed, the orbit spaces obtained after reduction are generically non-Euclidean, even if the original phase space is.

(ii) The stochastic dynamical components of the system are modeled by continuous semimartingales and are not limited to Brownian motion.

(iii) We handle stochastic Hamiltonian systems on Poisson manifolds and not only on symplectic manifolds.

(iv) The variational principle that we propose in Theorem 2.34 is not just satisfied by the stochastic Hamiltonian equations (as in [B81]) but fully characterizes them.

There are various reasons that have lead us to consider these generalized Hamiltonian systems. First, even though the laws that govern the dynamics of classical mechanical systems are, in principle, completely known, the finite precision of experimental measurements yields impossible the estimation of the parameters of a particular given one with total accuracy. Second, the modeling of complex physical systems involves most of the time simplifying assumptions or idealizations of parts of the system, some of which could be included in the description as a stochastic component; this modeling philosophy has been extremely successful in the social sciences [BJ76]. Third, even if the model and the parameters of the system are known with complete accuracy, the solutions of the associated differential equations may be of great complexity and exhibit high sensitivity to the initial conditions hence making the probabilistic treatment and description of the solutions appropriate. Finally, we will see (Subsection 2.2.3) how stochastic Hamiltonian modeling of microscopic systems can be used to model dissipation and macroscopic damping.

This chapter is structured as follows: in Section 2.1 we introduce the stochastic Hamilton equations with phase space a given Poisson manifold and we study some of the fundamental properties of the solution semimartingales like, for instance, the preservation of symplectic leaves or the characterization of the conserved quantities. This section contains a discussion on two notions on non-linear stability, almost sure Lyapunov stability and stability in probability, that reduce in the deterministic setup to the standard definition of Lyapunov stability. We formulate criteria that generalize to the Hamiltonian stochastic context the standard energy methods to conclude the stability of a Hamiltonian equilibrium using existing conservation laws. More specifically, there are two different natural notions of conserved quantity in the stochastic context that, via a stochastic Dirichlet criterion (Theorem 2.15) allow one to conclude the different kinds of stability that we have mentioned above. Section 2.2 contains several examples: in the first one we show how the systems studied by Bismut in [B81] fall in the category introduced in Section 2.1. We also see that a damped oscillator can be described as the average motion of the solution semimartingale of a natural stochastic Hamiltonian system, and that Brownian motion in a manifold is the projection onto the base space of very simple Hamiltonian stochastic semimartingale defined on the cotangent bundle of the manifold or of its orthonormal frame bundle, depending on the availability or not of a parallelization for the manifold in question. Section 2.3 is dedicated to showing that the stochastic Hamilton equations are characterized by a critical action principle that generalizes the one found in the treatment of deterministic systems. In order to make this part more readable, the proofs of most of the technical results needed to prove the theorems in this section have been included separately at the end of the chapter. Finally, we show in Section 2.4 that the stochastic action satisfies a generalized version of the Hamilton-Jacobi equation when written as a function of the configuration space using a Lagrangian submanifold (see Theorem 2.39). As an application of the results in this section we show in Example 2.41 how the exponential of the expectation of the so called projected stochastic action can be used to construct solutions of the heat equation corrected with a potential, in a way that strongly resembles the Feynman-Kac formula.

This chapter is a transcription of the two papers [LO07] and [LO08b] written by the author of this thesis in collaboration with Juan Pablo Ortega.

Conventions: All the manifolds will be finite dimensional, second-countable, locally compact, and Hausdorff (and hence paracompact).

2.1 The stochastic Hamilton equations

In this section we present a natural generalization of the standard Hamilton equations in the stochastic context. Even though the arguments gathered in the following paragraphs as motivation for these equations are of formal nature, we will see later on that, as it was already the case for the standard Hamilton equations, they satisfy a natural variational principle.

We recall that a ***symplectic manifold*** is a pair (M, ω), where M is a manifold and $\omega \in \Omega^2(M)$ is a closed non-degenerate two-form on M, that is, $\mathbf{d}\omega = 0$ and, for every $m \in M$, the map $v \in T_mM \mapsto \omega(m)(v, \cdot) \in T_m^*M$ is a linear isomorphism between the tangent space T_mM to M at m and the cotangent space T_m^*M. Using the nondegeneracy of the symplectic form ω, one can associate each function $h \in C^\infty(M)$ a vector field $X_h \in \mathfrak{X}(M)$, defined by the equality

$$\mathbf{i}_{X_h}\omega = \mathbf{d}h. \tag{2.1}$$

We will say that X_h is the ***Hamiltonian vector field*** associated to the ***Hamiltonian function*** h. The expression (2.1) is referred to as the ***Hamilton equations***.

A ***Poisson manifold*** is a pair $(M, \{\cdot,\cdot\})$, where M is a manifold and $\{\cdot,\cdot\}$ is a bilinear operation on $C^\infty(M)$ such that $(C^\infty(M), \{\cdot,\cdot\})$ is a Lie algebra and $\{\cdot,\cdot\}$ is a derivation (that is, the Leibniz identity holds) in each argument. The functions in the center $\mathcal{C}(M)$ of the Lie algebra $(C^\infty(M), \{\cdot,\cdot\})$ are called ***Casimir functions***. From the natural isomorphism between derivations on $C^\infty(M)$ and vector fields on M it follows that each $h \in C^\infty(M)$ induces a vector field on M via the expression $X_h = \{\cdot, h\}$, called the ***Hamiltonian vector field*** associated to the ***Hamiltonian function*** h. Hamilton's equations $\dot{z} = X_h(z)$ can be equivalently written in Poisson bracket form as $\dot{f} = \{f, h\}$, for any $f \in C^\infty(M)$. The derivation property of the Poisson bracket implies that for any two functions $f, g \in C^\infty(M)$, the value of the bracket $\{f, g\}(z)$ at an arbitrary point $z \in M$ (and therefore $X_f(z)$ as well), depends on f only through $\mathbf{d}f(z)$ which allows us to define a contravariant antisymmetric two–tensor $B \in \Lambda^2(M)$ by $B(z)(\alpha_z, \beta_z) = \{f, g\}(z)$, where $\mathbf{d}f(z) = \alpha_z \in T_z^*M$ and $\mathbf{d}g(z) = \beta_z \in T_z^*M$. This tensor is called the ***Poisson tensor*** of M. The vector bundle map $B^\sharp : T^*M \to TM$ naturally associated to B is defined by $B(z)(\alpha_z, \beta_z) = \langle \alpha_z, B^\sharp(\beta_z)\rangle$.

We start by rewriting the solutions of the standard Hamilton equations in a form that we will be able to mimic in the stochastic differential equations context. All the necessary prerequisites on stochastic calculus on manifolds have already been introduced in Chapter 1.

Proposition 2.1 *Let (M, ω) be a symplectic manifold and $h \in C^\infty(M)$. The smooth curve $\gamma : [0, T] \to M$ is an integral curve of the Hamiltonian vector field X_h if and only if for any $\alpha \in \Omega(M)$ and for any $t \in [0, T]$*

$$\int_{\gamma|_{[0,t]}} \alpha = -\int_0^t \mathbf{d}h(\omega^\sharp(\alpha)) \circ \gamma(s) ds, \tag{2.2}$$

where $\omega^\sharp : T^*M \to TM$ is the vector bundle isomorphism induced by ω. More generally, if M is a Poisson manifold with bracket $\{\cdot, \cdot\}$ then the same result holds with (2.2) replaced by

$$\int_{\gamma|_{[0,t]}} \alpha = -\int_0^t \mathbf{d}h(B^\sharp(\alpha)) \circ \gamma(s) ds, \tag{2.3}$$

Proof. Since in the symplectic case $\omega^\sharp = B^\sharp$, it suffices to prove (2.3). As (2.3) holds for any $t \in [0, T]$, we can take derivatives with respect to t on both sides and we obtain the equivalent form

$$\langle \alpha(\gamma(t)), \dot{\gamma}(t) \rangle = -\langle \mathbf{d}h(\gamma(t)), B^\sharp(\gamma(t))(\alpha(\gamma(t))) \rangle. \tag{2.4}$$

Let $f \in C^\infty(M)$ be such that $\mathbf{d}f(\gamma(t)) = \alpha(\gamma(t))$. Then (2.4) can be rewritten as

$$\langle \mathbf{d}f(\gamma(t)), \dot{\gamma}(t) \rangle = -\langle \mathbf{d}h(\gamma(t)), B^\sharp(\gamma(t))(\mathbf{d}f(\gamma(t))) \rangle = \{f, h\}(\gamma(t)),$$

which is equivalent to $\dot{\gamma}(t) = X_h(\gamma(t))$, as required. ■

We will now introduce the stochastic Hamilton equations by mimicking in the context of Stratonovich integration the integral expressions (2.2) and (2.3). In the next definition we will use the following notation: let $f : M \to W$ be a differentiable function that takes values on the vector space W. We define the ***differential*** $\mathbf{d}f : TM \to W$ as the map given by $\mathbf{d}f = p_2 \circ Tf$, where $Tf : TM \to TW = W \times W$ is the tangent map of f and $p_2 : W \times W \to W$ is the projection onto the second factor. If $W = \mathbb{R}$ this definition coincides with the usual differential. If $\{e_1, \ldots, e_n\}$ is a basis of W and $f = \sum_{i=1}^n f^i e_i$ then $\mathbf{d}f = \sum_{i=1}^n \mathbf{d}f^i \otimes e_i$.

Definition 2.2 *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a semimartingale that takes values on the vector space V with $X_0 = 0$, and $h : M \to V^*$ a smooth function. Let $\{\epsilon^1, \ldots, \epsilon^r\}$ be a basis of V^* and $h = \sum_{i=1}^r h_i \epsilon^i$. The **Hamilton equations** with **stochastic component** X, and **Hamiltonian function** h are the Stratonovich stochastic differential equation*

$$\delta \Gamma^h = H(X, \Gamma) \delta X, \tag{2.5}$$

defined by the Stratonovich operator $H(v, z) : T_v V \to T_z M$ given by

$$H(v, z)(u) := \sum_{j=1}^r \langle \epsilon^j, u \rangle X_{h_j}(z). \tag{2.6}$$

The dual Stratonovich operator $H^*(v,z) : T_z^*M \to T_v^*V$ of $H(v,z)$ is given by $H^*(v,z)(\alpha_z) = -\mathbf{d}h(z) \cdot B^\sharp(z)(\alpha_z)$. Hence, the results quoted in Subsection 1.4.4 show that for any $\mathcal{F}_0$ measurable random variable Γ_0, there exists a unique semimartingale Γ^h such that $\Gamma_0^h = \Gamma_0$ and a maximal stopping time ζ^h that solve (2.5), that is, for any $\alpha \in \Omega(M)$,

$$\int \langle \alpha, \delta\Gamma^h \rangle = -\int \langle \mathbf{d}h(B^\sharp(\alpha))(\Gamma^h), \delta X \rangle. \tag{2.7}$$

We will refer to Γ^h as the ***Hamiltonian semimartingale*** associated to h with initial condition Γ_0.

Remark 2.3 The stochastic component X encodes the random behavior exhibited by the stochastic Hamiltonian system that we are modeling and the Hamiltonian function h specifies how it embeds in its phase space. Unlike the situation encountered in the deterministic setup we allow the Hamiltonian function to be vector valued in order to accommodate higher dimensional stochastic dynamics.

Remark 2.4 The generalization of Hamilton's equations proposed in Definition 2.2 by using a Stratonovich operator is inspired by one of the transfer principles presented in [E90] to provide stochastic versions of ordinary differential equations. This procedure can be also used to carry out a similar generalization of the equations induced by a Leibniz bracket (see [OP04]).

Remark 2.5 Stratonovich versus Itô integration: at the time of proposing the equations in Definition 2.2 a choice has been made, namely, we have chosen Stratonovich integration instead of Itô or other kinds of stochastic integration. The option that we took is motivated by the fact that by using Stratonovich integration, most of the geometric features underlying classical deterministic Hamiltonian mechanics are preserved in the stochastic context (see the next section). Additionally, from the mathematical point of view, this choice is the most economical one in the sense that the classical geometric ingredients of Hamiltonian mechanics plus a noise semimartingale suffice to construct the equations; had we used Itô integration we would have had to provide a Schwartz operator (see Subsection 1.4.4) and the construction of such an object via a transfer principle like in [E90] involves the choice of a connection. The use of Itô integration in the modeling of physical phenomena is sometimes preferred because the definition of this integral is not *anticipative*, that is, it does not assume any knowledge about the behavior of the system in future times. Even though we have used Stratonovich integration to write down our equations, we also share this feature because the equations in Definition 2.2 can be naturally translated to the Itô framework (see Proposition 2.8). This is a particular case of a more general fact since given any Stratonovich stochastic differential equation there always exists an equivalent Itô stochastic differential equation, in the sense that both equations have the same solutions. Note that the converse is in general not true.

2.1.1 Elementary properties of the stochastic Hamilton's equations

Proposition 2.6 *Let $(M, \{\cdot,\cdot\})$ be a Poisson manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a semimartingale that takes values on the vector space V with $X_0 = 0$ and $h : M \to V^*$ a smooth function. Let*

Γ_0 be a $\mathcal{F}_0$ measurable random variable and Γ^h the Hamiltonian semimartingale associated to h with initial condition Γ_0. Let ζ^h be the corresponding maximal stopping time. Then, for any stopping time $\tau < \zeta^h$, the Hamiltonian semimartingale Γ^h satisfies

$$f(\Gamma^h_\tau) - f(\Gamma^h_0) = \sum_{j=1}^{r} \int_0^\tau \{f, h_j\}(\Gamma^h)\delta X^j, \tag{2.8}$$

where $\{h_j\}_{j\in\{1,\ldots,r\}}$ and $\{X^j\}_{j\in\{1,\ldots,r\}}$ are the components of h and X with respect to two given dual bases $\{e_1, \ldots, e_r\}$ and $\{\epsilon^1, \ldots, \epsilon^r\}$ of V and V^*, respectively. Expression (2.8) can be rewritten in differential notation as

$$\delta f(\Gamma^h) = \sum_{j=1}^{r} \{f, h_j\}(\Gamma^h)\delta X^j.$$

Proof. It suffices to take $\alpha = \mathbf{d}f$ in (2.7). Indeed, by (1.49a)

$$\int_0^\tau \langle \mathbf{d}f, \delta\Gamma^h \rangle = f(\Gamma^h_\tau) - f(\Gamma^h_0).$$

At the same time

$$\begin{aligned} -\int_0^\tau \langle \mathbf{d}h(B^\sharp(\mathbf{d}f))(\Gamma^h), \delta X\rangle &= -\sum_{j=1}^{r}\int_0^\tau \langle (\mathbf{d}h_j \otimes \epsilon^j(B^\sharp(\mathbf{d}f)))(\Gamma^h), \delta X\rangle \\ &= \sum_{j=1}^{r}\int_0^\tau \langle \{f, h_j\}(\Gamma^h)\epsilon^j, \delta X\rangle. \end{aligned}$$

By (1.49b) this equals $\sum_{j=1}^r \int_0^\tau \{f, h_j\}(\Gamma^h)\delta\left(\int\langle \epsilon^j, \delta X\rangle\right)$. Given that $\int\langle\epsilon^j, \delta X\rangle = X^j - X^j_0$, the equality follows. ■

Remark 2.7 Notice that if in Definition 2.2 we take $V^* = \mathbb{R}$, $h \in C^\infty(M)$, and $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ the deterministic process given by $(t, \omega) \longmapsto t$, then the stochastic Hamilton equations (2.7) reduce to

$$\int \langle \alpha, \delta\Gamma^h\rangle = \int \langle \alpha, X_h\rangle(\Gamma^h_t)dt. \tag{2.9}$$

A straightforward application of (2.8) shows that $\Gamma^h_t(\omega)$ is necessarily a differentiable curve, for any $\omega \in \Omega$, and hence the Riemann-Stieltjes integral in the left hand side of (2.9) reduces, when evaluated at a given $\omega \in \Omega$, to a Riemann integral identical to the one in the left hand side of (2.3), hence proving that (2.9) reduces to the standard Hamilton equations. Indeed, let $\Gamma^h_{t_0}(\omega) \in M$ be an arbitrary point in the curve $\Gamma^h_t(\omega)$, let U be a coordinate patch around $\Gamma^h_{t_0}(\omega)$ with coordinates $\{x^1, \ldots, x^n\}$, and let $x(t) = (x^1(t), \ldots, x^n(t))$ be the expression of $\Gamma^h_t(\omega)$ in these coordinates. Then by (2.8), for $h \in \mathbb{R}$ sufficiently small, and $i \in \{1, \ldots, n\}$,

$$x^i(t_0 + h) - x^i(t_0) = \int_{t_0}^{t_0+h} \{x^i, h\}(x(t))dt.$$

Hence, by the Fundamental Theorem of Calculus, $x^i(t)$ is differentiable at t_0, with derivative

$$\dot{x}^i(t_0) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(x^i(t_0+\epsilon) - x^i(t_0)\right) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\int_{t_0}^{t_0+\epsilon}\{x^i,h\}(x(t))dt\right) = \{x^i,h\}(x(t_0)),$$

as required.

The following proposition provides an equivalent expression of the Stochastic Hamilton equations in the Itô form (see Subsection 1.4.4).

Proposition 2.8 *The stochastic Hamilton's equations in Definition 2.2 admit an equivalent description using Itô integration by using the Schwartz operator $\mathcal{H}(v,m) : \tau_v V \to \tau_m M$ naturally associated to the Hamiltonian Stratonovich operator H and that can be described as follows. Let $L \in \tau_v M$ be a second order vector and $f \in C^\infty(M)$ arbitrary, then*

$$\mathcal{H}(v,m)(L)[f] = \left\langle \sum_{i,j=1}^{r}\{f,h_j\}(m)\epsilon^j + \{\{f,h_j\},h_i\}(m)\epsilon^i\cdot\epsilon^j, L\right\rangle.$$

Moreover, expression (2.8) in the Itô representation is given by

$$f(\Gamma^h_\tau) - f(\Gamma^h_0) = \sum_{j=1}^{r}\int_0^\tau \{f,h_j\}(\Gamma^h)dX^j + \frac{1}{2}\sum_{j,i=1}^{r}\int_0^\tau \{\{f,h_j\},h_i\}(\Gamma^h)d\left[X^j,X^i\right]. \quad (2.10)$$

*We will refer to $\mathcal{H}$ as the **Hamiltonian Schwartz operator** associated to h.*

Proof. According to the remarks made in Subsection 1.4.4, the Schwartz operator $\mathcal{H}$ naturally associated to H is constructed as follows. For any second order vector $L_{\ddot{v}} \in \tau_v M$ associated to the acceleration of a curve $v(t)$ in V such that $v(0) = v$ we define $\mathcal{H}(v,m)(L_{\ddot{v}}) := L_{\ddot{m}(0)} \in \tau_m M$, where $m(t)$ is a curve in M such that $m(0) = m$ and $\dot{m}(t) = H(v(t), m(t))\dot{v}(t)$, for t in a neighborhood of 0. Consequently,

$$\begin{aligned}
\mathcal{H}(v,m)(L_{\ddot{v}})[f] &= \left.\frac{d^2}{dt^2}\right|_{t=0} f(m(t)) = \left.\frac{d}{dt}\right|_{t=0}\langle \mathbf{d}f(m(t)), \dot{m}(t)\rangle \\
&= \left.\frac{d}{dt}\right|_{t=0}\langle \mathbf{d}f(m(t)), H(v(t),m(t))\dot{v}(t)\rangle \\
&= \left.\frac{d}{dt}\right|_{t=0}\sum_{j=1}^{r}\langle\epsilon^j,\dot{v}(t)\rangle\langle \mathbf{d}f(m(t)), X_{h_j}(m(t))\rangle = \left.\frac{d}{dt}\right|_{t=0}\sum_{j=1}^{r}\langle\epsilon^j,\dot{v}(t)\rangle\{f,h_j\}(m(t)) \\
&= \sum_{j=1}^{r}\langle\epsilon^j,\ddot{v}(0)\rangle\{f,h_j\}(m) + \langle\epsilon^j,\dot{v}(0)\rangle\langle \mathbf{d}\{f,h_j\}(m),\dot{m}(0)\rangle \\
&= \sum_{j=1}^{r}\langle\epsilon^j,\ddot{v}(0)\rangle\{f,h_j\}(m) + \langle\epsilon^j,\dot{v}(0)\rangle\sum_{i=1}^{r}\langle\epsilon^i,\dot{v}(0)\rangle\{\{f,h_j\},h_i\}(m) \\
&= \left\langle \sum_{i,j=1}^{r}\{f,h_j\}(m)\epsilon^j + \{\{f,h_j\},h_i\}(m)\epsilon^i\cdot\epsilon^j, L_{\ddot{v}}\right\rangle.
\end{aligned}$$

In order to establish (2.10) we need to calculate $\mathcal{H}^*(v,m)(d_2 f(m))$ for a second order form $d_2 f(m) \in \tau_m^* M$ at $m \in M$, $f \in C^\infty(M)$. Since $\mathcal{H}^*(v,m)(d_2 f(m))$ is fully characterized by its action on elements of the form $L_{\ddot{v}} \in \tau_v V$ for some curve $v(t)$ in V such that $v(0) = v$, we have

$$\begin{aligned}\langle \mathcal{H}^*(v,m)(d_2 f(m)), L_{\ddot{v}} \rangle &= \langle d_2 f(m), \mathcal{H}(v,m)(L_{\ddot{v}}) \rangle = \mathcal{H}(v,m)(L_{\ddot{v}})[f] \\ &= \left\langle \sum_{i,j=1}^{r} \{f, h_j\}(m)\epsilon^j + \{\{f,h_j\},h_i\}(m)\epsilon^i \cdot \epsilon^j, L_{\ddot{v}} \right\rangle.\end{aligned}$$

Consequently, $\mathcal{H}^*(v,m)(d_2 f(m)) = \sum_{i,j=1}^{r} \{f,h_j\}(m)\epsilon^j + \{\{f,h_j\},h_i\}(m)\epsilon^i \cdot \epsilon^j$. Hence, if Γ_h is the Hamiltonian semimartingale associated to h with initial condition Γ_0, $\tau < \zeta^h$ is any stopping time, and $f \in C^\infty(M)$, we have by (1.46a), (1.46b), and (1.48)

$$\begin{aligned} f(\Gamma_\tau^h) - f(\Gamma_0^h) &= \int_0^\tau \left\langle d_2 f, d\Gamma^h \right\rangle = \int_0^\tau \left\langle \mathcal{H}^*(X, \Gamma^h)(d_2 f), dX \right\rangle \\ &= \sum_{j=1}^{r} \int_0^\tau \left\langle \{f,h_j\}(\Gamma^h)\epsilon^j, dX \right\rangle + \sum_{j,i=1}^{r} \int_0^\tau \left\langle \{\{f,h_j\},h_i\}(\Gamma^h)\epsilon^i \cdot \epsilon^j, dX \right\rangle \\ &= \sum_{j=1}^{r} \int_0^\tau \{f,h_j\}(\Gamma^h) dX^j + \frac{1}{2} \sum_{j,i=1}^{r} \int_0^\tau \{\{f,h_j\},h_i\}(\Gamma^h) d\left[X^i, X^j\right]. \end{aligned}$$

■

Proposition 2.9 (Preservation of the symplectic leaves by Hamiltonian semimartingales) *In the setup of Definition 2.2, let $\mathcal{L}$ be a symplectic leaf of (M,ω) and Γ^h a Hamiltonian semimartingale with initial condition $\Gamma_0(\omega) = Z_0$, where Z_0 is a random variable such that $Z_0(\omega) \in \mathcal{L}$ for all $\omega \in \Omega$. Then, there exists a stopping time $\zeta_{\mathcal{L}}^h \leq \zeta^h$ such that for any stopping time $\tau < \zeta_{\mathcal{L}}^h$ we have that $\Gamma_\tau^h \in \mathcal{L}$. If the symplectic leaf $\mathcal{L}$ is a closed subset of M then $\zeta_{\mathcal{L}}^h = \zeta^h$.*

Proof. Expression (2.6) shows that for any $z \in \mathcal{L}$, the Stratonovich operator $H(v,z)$ takes values in the characteristic distribution associated to the Poisson structure $(M, \{\cdot,\cdot\})$, that is, in the tangent space $T\mathcal{L}$ of $\mathcal{L}$. Consequently, H induces another Stratonovich operator $H_{\mathcal{L}}(v,z) : T_v V \to T_z \mathcal{L}$, $v \in V$, $z \in \mathcal{L}$, obtained from H by restriction of its range. It is clear that if $i : \mathcal{L} \hookrightarrow M$ is the inclusion then

$$H_{\mathcal{L}}^*(v,z) \circ T_z^* i = H^*(v,z). \tag{2.11}$$

Let $\Gamma_{\mathcal{L}}^h$ be the semimartingale in $\mathcal{L}$ that is a solution of the Stratonovich stochastic differential equation

$$\delta\Gamma_{\mathcal{L}}^h = H_{\mathcal{L}}(X, \Gamma_{\mathcal{L}}^h)\delta X \tag{2.12}$$

with initial condition Γ_0. We now show that $\overline{\Gamma} := i \circ \Gamma_{\mathcal{L}}^h$ is a solution of

$$\delta\overline{\Gamma} = H(X, \overline{\Gamma})\delta X.$$

The uniqueness of the solution of a stochastic differential equation will guarantee in that situation that Γ^h necessarily coincides with $\overline{\Gamma}$, in the times in which both are defined. More specifically, $\Gamma^h = \overline{\Gamma}$ up to $\zeta^h_{\mathcal{L}}$, with $\zeta^h_{\mathcal{L}}$ the maximal stopping time associated to the solution of (2.12), with initial condition $\overline{\Gamma}_0(\omega) = Z_0 \in \mathcal{L}$; this will prove the statement. Indeed, for any $\alpha \in \Omega(M)$,

$$\int \langle \alpha, \delta\overline{\Gamma}\rangle = \int \langle \alpha, \delta(i \circ \Gamma^h_{\mathcal{L}})\rangle = \int \langle T^* i \cdot \alpha, \delta \Gamma^h_{\mathcal{L}}\rangle.$$

Since $\Gamma^h_{\mathcal{L}}$ satisfies (2.12) and $T^* i \cdot \alpha \in \Omega(\mathcal{L})$, by (2.11) this equals

$$\int \langle H^*_{\mathcal{L}}(X, \Gamma^h_{\mathcal{L}})(T^* i \cdot \alpha), \delta X\rangle = \int \langle H^*(X, i \circ \Gamma^h_{\mathcal{L}})(\alpha), \delta X\rangle = \int \langle H^*(X, \overline{\Gamma})(\alpha), \delta X\rangle,$$

that is, $\delta\overline{\Gamma} = H(X, \overline{\Gamma})\delta X$, as required. The statement on the equality $\zeta^h_{\mathcal{L}} = \zeta^h$ under the hypothesis that the symplectic leaf $\mathcal{L}$ is closed is a consequence of [E82, Theorem 3 page 123]. ■

Proposition 2.10 (The stochastic Hamilton equations in Darboux-Weinstein coordinates) *Let $(M, \{\cdot,\cdot\})$ be a Poisson manifold and Γ^h be a solution of the Hamilton equations (2.5) with initial condition $x_0 \in M$. There exists an open neighborhood U of x_0 in M and a stopping time τ_U such that $\Gamma^h_t(\omega) \in U$, for any $\omega \in \Omega$ and any $t \leq \tau_U(\omega)$. Moreover, U admits local Darboux coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n, z_1, \ldots, z_l)$ in which (2.8) takes the form*

$$q^i(\Gamma^h_\tau) - q^i(\Gamma^h_0) = \sum_{j=1}^{r} \int_0^\tau \frac{\partial h_j}{\partial p_i} \delta X^j,$$

$$p_i(\Gamma^h_\tau) - p_i(\Gamma^h_0) = -\sum_{j=1}^{r} \int_0^\tau \frac{\partial h_j}{\partial q^i} \delta X^j,$$

$$z_i(\Gamma^h_\tau) - z_i(\Gamma^h_0) = \sum_{j=1}^{r} \int_0^\tau \{z_i, h_j\}_T \delta X^j,$$

where $\{\cdot,\cdot\}_T$ is the transverse Poisson structure of $(M, \{\cdot,\cdot\})$ at x_0.

Proof. Let U be an open neighborhood of x_0 in M for which Darboux coordinates can be chosen. Define $\tau_U = \inf_{t\geq 0}\{\Gamma^h_t \in U^c\}$ (τ_U is the exit time of U). It is a standard fact in the theory of stochastic processes that τ_U is a stopping time. The proposition follows by writing (2.8) for the Darboux-Weinstein coordinate functions $(q^1, \ldots, q^n, p_1, \ldots, p_n, z_1, \ldots, z_l)$.

■

Let $\zeta : M \times \Omega \to [0, \infty]$ be the map such that, for any $z \in M$, $\zeta(z)$ is the maximal stopping time associated to the solution of the stochastic Hamilton equations (2.5) with initial condition $\Gamma_0 = z$ a.s.. Let φ be the ***flow*** of (2.5), that is, for any $z \in M$, $\varphi(z) : [0, \zeta(z)) \to M$ is the solution semimartingale of (2.5) with initial condition z. The map $z \in M \longmapsto \varphi_t(z, \omega) \in M$ is a local diffeomorphism of M, for each $t \geq 0$ and almost all $\omega \in \Omega$ in which this map is defined (see [IW89]). In the following result, we show that, in the symplectic context, Hamiltonian flows preserve the symplectic form and hence the associated volume form $\theta = \omega \wedge \overset{n}{\ldots} \wedge \omega$. This has already been shown for Hamiltonian diffusions (see Example 2.2.1) by Bismut [B81].

Theorem 2.11 (Stochastic Liouville's Theorem) *Let (M,ω) be a symplectic manifold, $X:\mathbb{R}_+\times\Omega\to V^*$ a semimartingale, and $h:M\to V^*$ a Hamiltonian function. Let φ be the associated Hamiltonian flow. Then, for any $z\in M$ and any $(t,\eta)\in[0,\zeta(z))$,*

$$\varphi_t^*(z,\eta)\,\omega=\omega.$$

Proof. By [K81, Theorem 3.3] (see also [W80]), given an arbitrary form $\alpha\in\Omega^k(M)$ and $z\in M$, the process $\varphi(z)^*\alpha$ satisfies the following stochastic differential equation:

$$\varphi(z)^*\alpha=\alpha(z)+\sum_{j=1}^{r}\int\varphi(z)^*\left(\pounds_{X_{h_j}}\alpha\right)\delta X^j.$$

In particular, if $\alpha=\omega$ then $\pounds_{X_{h_j}}\omega=0$ for any $j\in\{1,...,r\}$, and hence the result follows. ■

2.1.2 Conserved quantities and stability

Conservation laws in Hamiltonian mechanics are extremely important since they make easier the integration of the systems that have them and, in some instances, provide qualitative information about the dynamics. A particular case of this is their use in concluding the nonlinear stability of certain equilibrium solutions using Dirichlet type criteria that we will generalize to the stochastic setup using the following definitions.

Definition 2.12 *A function $f\in C^\infty(M)$ is said to be a* ***strongly*** *(respectively,* ***weakly****)* ***conserved quantity*** *of the stochastic Hamiltonian system associated to $h:M\to V^*$ if for any solution Γ^h of the stochastic Hamilton equations (2.5) we have that $f(\Gamma^h)=f(\Gamma^h_0)$ (respectively, $E[f(\Gamma^h_\tau)]=E[f(\Gamma^h_0)]$, for any stopping time τ).*

Notice that strongly conserved quantities are obviously weakly conserved and that the two definitions coincide for deterministic systems with the standard definition of conserved quantity. The following result provides in the stochastic setup an analogue of the classical characterization of the conserved quantities in terms of Poisson involution properties.

Proposition 2.13 *Let $(M,\{\cdot,\cdot\})$ be a Poisson manifold, $X:\mathbb{R}_+\times\Omega\to V$ a semimartingale that takes values on the vector space V such that $X_0=0$, and $h:M\to V^*$ and $f\in C^\infty(M)$ two smooth functions. If $\{f,h_j\}=0$ for every component h_j of h then f is a strongly conserved quantity of the stochastic Hamilton equations (2.5). Conversely, suppose that the semimartingale $X=\sum_{j=1}^r X^j\epsilon_j$ is such that $\left[X^i,X^j\right]=0$ if $i\neq j$. If f is a strongly conserved quantity then $\{f,h_j\}=0$, for any $j\in\{1,...,r\}$ such that $\left[X^j,X^j\right]$ is an strictly increasing process at 0. The last condition means that there exists $A\in\mathcal{F}$ and $\delta>0$ with $P(A)>0$ such that for any $t<\delta$ and $\omega\in A$ we have $[X^j,X^j]_t(\omega)>[X^j,X^j]_0(\omega)$, for all $j\in\{1,...,r\}$.*

Proof. Let Γ^h be the Hamiltonian semimartingale associated to h with initial condition Γ^h_0. As we saw in (2.10),

$$f(\Gamma^h)=f(\Gamma^h_0)+\sum_{j=1}^{r}\int\{f,h_j\}(\Gamma^h)dX^j+\frac{1}{2}\sum_{j,i=1}^{r}\int\{\{f,h_j\},h_i\}(\Gamma^h)d\left[X^i,X^j\right]. \tag{2.13}$$

If $\{f, h_j\} = 0$ for every component h_j of h then all the integrals in the previous expression vanish and therefore $f(\Gamma^h) = f(\Gamma_0^h)$ which implies that f is a strongly conserved quantity of the Hamiltonian stochastic equations associated to h. Conversely, suppose now that f is a strongly conserved quantity. This implies that for any initial condition Γ_0^h, the semimartingale $f(\Gamma^h)$ is actually time independent and hence of finite variation. Equivalently, the (unique) decomposition of $f(\Gamma^h)$ into two processes, one of finite variation plus a local martingale, only has the first term. In order to isolate the local martingale term of $f\left(\Gamma^h\right)$ recall first that the quadratic variations $\left[X^i, X^j\right]$ have finite variation and that the integral with respect to a finite variation process has finite variation (see [LG97, Proposition 4.3]). Consequently, the last summand in (2.13) has finite variation. As to the second summand, let M^j and A^j, $j = 1, \ldots, r$, local martingales and finite variation processes, respectively, such that $X^j = A^j + M^j$. Then,

$$\int \{f, h_j\} (\Gamma^h) dX^j = \int \{f, h_j\} (\Gamma^h) dM^j + \int \{f, h_j\} (\Gamma^h) dA^j.$$

Given that for each j, $\int \{f, h_j\} (\Gamma^h) dA^j$ is a finite variation process and $\int \{f, h_j\} (\Gamma^h) dM^j$ is a local martingale (see [P05, Theorem 29, page 128]) we conclude that $Z := \sum_{j=1}^r \int \{f, h_j\} (\Gamma^h) dM^j$ is the local martingale term of $f(\Gamma^h)$ and hence equal to zero. We notice now that any continuous local martingale $Z : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is also a local $L^2(\Omega)$-martingale. Indeed, consider the sequence of stopping times $\tau^n = \{\inf t \geq 0 \mid |Z_t| = n\}$, $n \in \mathbb{N}$. Then $E[(Z^{\tau^n})_t^2] \leq E\left[n^2\right] = n^2$, for all $t \in \mathbb{R}_+$. Hence, $Z^{\tau^n} \in L^2(\Omega)$ for any n. In addition, $E[(Z^{\tau^n})_t^2] = E\left[\left[Z^{\tau^n}, Z^{\tau^n}\right]_t\right]$ (see [P05, Corollary 3, page 73]). On the other hand by Proposition A.1,

$$Z^{\tau^n} = \left(\sum_{j=1}^r \int \{f, h_j\} (\Gamma^h) dM^j\right)^{\tau^n} = \sum_{j=1}^r \int \mathbf{1}_{[0,\tau^n]} \{f, h_j\} (\Gamma^h) dM^j.$$

Thus, by [P05, Theorem 29, page 75] and the hypothesis $\left[X^i, X^j\right] = 0$ if $i \neq j$,

$$\begin{aligned}
E\left[\left(Z^{\tau^n}\right)_t^2\right] &= E\left[\left[Z^{\tau^n}, Z^{\tau^n}\right]_t\right] \\
&= \sum_{j,i=1}^r E\left[\left[\int \mathbf{1}_{[0,\tau^n]} \{f, h_j\} (\Gamma^h) dM^j, \int \mathbf{1}_{[0,\tau^n]} \{f, h_i\} (\Gamma^h) dM^i\right]_t\right] \\
&= \sum_{j,i=1}^r E\left[\left(\int \mathbf{1}_{[0,\tau^n]} (\{f, h_j\} \{f, h_i\}) (\Gamma^h) d\left[M^j, M^i\right]\right)_t\right] \\
&= \sum_{j,i=1}^r E\left[\left(\int \mathbf{1}_{[0,\tau^n]} (\{f, h_j\} \{f, h_i\}) (\Gamma^h) d\left[X^j, X^i\right]\right)_t\right] \\
&= \sum_{j=1}^r E\left[\left(\int \mathbf{1}_{[0,\tau^n]} \{f, h_j\}^2 (\Gamma^h) d\left[X^j, X^j\right]\right)_t\right].
\end{aligned}$$

Since $\left[X^j, X^j\right]$ is an increasing process of finite variation then $\int \mathbf{1}_{[0,\tau^n]} \{f, h_j\}^2 (\Gamma^h) d\left[X^j, X^j\right]$ is a Riemann-Stieltjes integral and hence for any $\omega \in \Omega$

$$\left(\int \mathbf{1}_{[0,\tau^n]} \{f, h_j\}^2 (\Gamma^h) d\left[X^j, X^j\right]\right)(\omega) = \int \mathbf{1}_{[0,\tau^n(\omega)]} \{f, h_j\}^2 (\Gamma^h(\omega)) d\left(\left[X^j, X^j\right](\omega)\right).$$

As $\left[X^j, X^j\right](\omega)$ is an increasing function of $t \in \mathbb{R}_+$, then for any $j \in \{1, \ldots, r\}$

$$E\left[\int \mathbf{1}_{[0,\tau^n]} \{f, h_j\}^2 (\Gamma^h) d\left[X^j, X^j\right]\right] \geq 0. \tag{2.14}$$

Additionally, since $E[(Z^{\tau^n})_t^2] = 0$, we necessarily have that the inequality in (2.14) is actually an equality. Hence,

$$\int_0^t \mathbf{1}_{[0,\tau^n]} \{f, h_j\}^2 (\Gamma^h) d\left[X^j, X^j\right] = 0. \tag{2.15}$$

Suppose now that $[X^j, X^j]$ is strictly increasing at 0 for a particular j. Hence, there exists $A \in \mathcal{F}$ with $P(A) > 0$, and $\delta > 0$ such that $[X^j, X^j]_t(\omega) > [X^j, X^j]_0(\omega)$ for any $t < \delta$. Take now a fixed $\omega \in A$. Since $\tau^n \to \infty$ a.s., we can take n large enough to ensure that $\tau^n(\omega) > t$, where $t \in [0, \delta)$. Thus, we may suppose that $\mathbf{1}_{[0,\tau^n]}(t, \omega) = 1$. As $\left[X^j, X^j\right](\omega)$ is an strictly increasing process at zero $\int_0^t \{f, h_j\}^2 \left(\Gamma^h(\omega)\right) d\left[X^j, X^j\right](\omega) > 0$ unless $\{f, h_j\}^2 \left(\Gamma^h(\omega)\right) = 0$ in a neighborhood $[0, \widetilde{\delta}_\omega)$ of 0 contained in $[0, \delta)$. In principle $\widetilde{\delta}_\omega > 0$ might depend on $\omega \in A$, so the values of $t \in [0, \delta)$ for which $\{f, h_j\}^2 \left(\Gamma_t^h(\omega)\right) = 0$ for any $\omega \in A$ are those verifying $0 \leq t \leq \inf_{\omega \in A} \widetilde{\delta}_\omega$. In any case (2.15) allows us to conclude that $\{f, h_j\}^2 \left(\Gamma_0^h(\omega)\right) = 0$ for any $\omega \in A$. Finally, consider any Γ^h solution to the Stochastic Hamilton equations with constant initial condition $\Gamma_0^h = m \in M$ an arbitrary point. Then, for any $\omega \in A$,

$$0 = \{f, h_j\}^2 \left(\Gamma_0^h(\omega)\right) = \{f, h_j\}^2 (m).$$

Since $m \in M$ is arbitrary we can conclude that $\{f, h_j\} = 0$. ∎

We now use the conserved quantities of a system in order to formulate sufficient Dirichlet type stability criteria. Even though the statements that follow are enunciated for processes that are not necessarily Hamiltonian, it is for these systems that the criteria are potentially most useful. We start by spelling out the kind of nonlinear stability that we are after.

Definition 2.14 *Let M be a manifold and let*

$$\delta\Gamma = e(X, \Gamma)\delta X \tag{2.16}$$

be a Stratonovich stochastic differential equation whose solutions $\Gamma : \mathbb{R} \times \Omega \to M$ take values on M. Given $x \in M$ and $s \in \mathbb{R}$, denote by $\Gamma^{s,x}$ the unique solution of (2.16) such that $\Gamma_s^{s,x}(\omega) = x$, for all $\omega \in \Omega$. Suppose that the point $z_0 \in M$ is an ***equilibrium*** *of (2.16), that is, the constant process $\Gamma_t(\omega) := z_0$, for all $t \in \mathbb{R}$ and $\omega \in \Omega$, is a solution of (2.16). Then we say that the equilibrium z_0 is*

(i) ***Almost surely (Lyapunov) stable*** *when for any open neighborhood U of z_0 there exists another neighborhood $V \subset U$ of z_0 such that for any $z \in V$ we have $\Gamma^{0,z} \subset U$, a.s.*

(ii) ***Stable in probability.*** *For any $s \geq 0$ and $\epsilon > 0$*

$$\lim_{x \to z_0} P\left\{\sup_{t > s} d\left(\Gamma_t^{s,x}, z_0\right) > \epsilon\right\} = 0,$$

where $d : M \times M \to \mathbb{R}$ is any distance function that generates the manifold topology of M.

Theorem 2.15 (Stochastic Dirichlet's Criterion) *Suppose that we are in the setup of the previous definition and assume that there exists a function $f \in C^\infty(M)$ such that $\mathbf{d}f(z_0) = 0$ and that the quadratic form $\mathbf{d}^2 f(z_0)$ is (positive or negative) definite. If f is a strongly (respectively, weakly) conserved quantity for the solutions of (2.16) then the equilibrium z_0 is almost surely stable (respectively, stable in probability).*

Proof. Since the stability of the equilibrium z_0 is a local statement, we can work in a chart of M around z_0 with coordinates $(x_1, \ldots, x_n)$ in which z_0 is modeled by the origin. Moreover, using the Morse lemma and the hypotheses on the function f, and assuming without loss of generality that $f(z_0) = 0$, we choose the coordinates $(x_1, \ldots, x_n)$ so that $f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$. Hence, in the definition of stability in probability, we can use the distance function $d(x, z_0) = f(x)$. Suppose now that f is a strongly conserved quantity and let U be an open neighborhood of z_0. Let $r > 0$ be such that $V := f^{-1}([0, r]) \subset U$. Let $z \in V$ with $f(z) = r'$. As f is a strongly conserved quantity $f(\Gamma^{0,z}) = r' \leq r$ and hence $\Gamma^{0,z} \subset U$, as required. In order to study the case in which f is a weakly conserved quantity, let $\epsilon > 0$ and let U_ϵ be the ball of radius ϵ around z_0. Then, for any $x \in U_\epsilon$ and $s \in \mathbb{R}_+$, let τ_{U_ϵ} be the first exit time of $\Gamma^{s,x}$ with respect to U_ϵ. Notice first that if $\omega \in \Omega$ belongs to the set $\{\omega \in \Omega \mid \sup_{0 \leq s < t} d(\Gamma_t^{s,x}, z_0) > \epsilon\} = \{\omega \in \Omega \mid \sup_{0 \leq s < t} f(\Gamma_t^{s,x}) > \epsilon^2\}$, then $\tau_{U_\epsilon}(\omega) \leq t$ and hence the stopped process $(\Gamma^{s,x})^{\tau_{U_\epsilon}}$ satisfies that

$$f\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}(\omega)\right) = f\left(\Gamma^{s,x}_{\tau_{U_\epsilon}(\omega)}(\omega)\right) = \epsilon^2,$$

for those values of ω. This ensures that

$$\epsilon^2 \mathbf{1}_{\{\omega \in \Omega \mid \sup_{0 \leq s < t} d\left(\Gamma_t^{s,x}, z_0\right) > \epsilon\}} \leq f\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}\right).$$

Taking expectations in both sides of this inequality we obtain

$$P\left(\sup_{0 \leq s < t} d\left(\Gamma_t^{s,x}, z_0\right) > \epsilon\right) \leq \frac{E[f\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}\right)]}{\epsilon^2}.$$

Since by hypothesis f is a weakly conserved quantity, we can rewrite the right hand side of this inequality as

$$\frac{E[f\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}\right)]}{\epsilon^2} = \frac{E\left[f\left(\Gamma^{s,x}_{\tau_{U_\epsilon} \wedge t}\right)\right]}{\epsilon^2} = \frac{E[f(\Gamma_s^{s,x})]}{\epsilon^2} = \frac{f(x)}{\epsilon^2},$$

and we can therefore conclude that

$$P\left(\sup_{0 \leq s < t} d\left(\Gamma_t^{s,x}, z_0\right) > \epsilon\right) \leq \frac{f(x)}{\epsilon^2}. \tag{2.17}$$

Taking the limit $x \to z_0$ in this expression and recalling that $f(z_0) = 0$, the result follows. ■

A careful inspection of the proof that we just carried out reveals that in order for (2.17) to hold, it would suffice to have $E[f(\Gamma_\tau)] \leq E[f(\Gamma_0)]$, for any stopping time τ and any solution Γ, instead of the equality guaranteed by the weak conservation condition. This motivates the next definition.

Definition 2.16 *Suppose that we are in the setup of Definition 2.14. Let U be an open neighborhood of the equilibrium z_0 and let $V : U \to \mathbb{R}$ be a continuous function. We say that V is a **Lyapunov function** for the equilibrium z_0 if $V(z_0) = 0$, $V(z) > 0$ for any $z \in U \setminus \{z_0\}$, and*

$$E[V(\Gamma_\tau)] \leq E[V(\Gamma_0)], \tag{2.18}$$

for any stopping time $\tau \leq \tau_U$ smaller than the first exit time from U and any solution Γ of (2.16).

This definition generalizes to the stochastic context the standard notion of Lyapunov function that one encounters in dynamical systems theory. If (2.16) is the stochastic differential equation associated to an Itô diffusion and the Lyapunov function is twice differentiable, the inequality (2.18) can be ensured by requiring that $A[V](z) \leq 0$, for any $z \in U \setminus \{z_0\}$, where A is the infinitesimal generator of the diffusion, and by using Dynkin's formula.

Theorem 2.17 (Stochastic Lyapunov's Theorem) *Let $z_0 \in M$ be an equilibrium solution of the stochastic differential equation (2.16) and let $V : U \to \mathbb{R}$ be a continuous Lyapunov function for z_0. Then z_0 is stable in probability.*

Proof. Let U_ϵ be the ball of radius ϵ around z_0 and let $V_\epsilon := \inf_{x \in U \setminus U_\epsilon} V(x)$. Using the same notation as in the previous theorem we denote, for any $x \in U_\epsilon$ and $s \in \mathbb{R}_+$, τ_{U_ϵ} as the first exit time of $\Gamma^{s,x}$ with respect to U_ϵ. Using the same approach as above we notice that if $\omega \in \Omega$ belongs to the set $\{\omega \in \Omega \mid \sup_{0 \leq s < t} d(\Gamma_t^{s,x}, z_0) > \epsilon\}$, then $\tau_{U_\epsilon}(\omega) \leq t$ and hence the stopped process $(\Gamma^{s,x})^{\tau_{U_\epsilon}}$ satisfies that

$$V\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}(\omega)\right) = V\left(\Gamma^{s,x}_{\tau_{U_\epsilon}(\omega)}(\omega)\right) \geq V_\epsilon,$$

for those values of ω, since $\Gamma^{s,x}_{\tau_{U_\epsilon}(\omega)}(\omega)$ belongs to the boundary of U_ϵ. This ensures that

$$V_\epsilon \mathbf{1}_{\{\omega \in \Omega \mid \sup_{0 \leq s < t} d(\Gamma_t^{s,x}, z_0) > \epsilon\}} \leq V\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}\right).$$

Taking expectations in both sides of this inequality we obtain

$$P\left(\sup_{0 \leq s < t} d(\Gamma_t^{s,x}, z_0) > \epsilon\right) \leq \frac{E[V\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}\right)]}{V_\epsilon}.$$

We now use that V being a Lyapunov function satisfies (2.18) and hence

$$\frac{E[V\left((\Gamma^{s,x})_t^{\tau_{U_\epsilon}}\right)]}{V_\epsilon} = \frac{E\left[V\left(\Gamma^{s,x}_{\tau_{U_\epsilon} \wedge t}\right)\right]}{V_\epsilon} \leq \frac{E[V(\Gamma_s^{s,x})]}{V_\epsilon} = \frac{V(x)}{V_\epsilon}.$$

We can therefore conclude that

$$P\left(\sup_{0 \leq s < t} d(\Gamma_t^{s,x}, z_0) > \epsilon\right) \leq \frac{V(x)}{V_\epsilon}.$$

Taking the limit $x \to z_0$ in this expression and recalling that $V(z_0) = 0$, the result follows. ∎

Remark 2.18 This theorem has been proved by Gihman [G96] and Hasminskii [H80] for Itô diffusions.

2.2 Examples

2.2.1 Stochastic perturbation of a Hamiltonian mechanical system and Bismut's Hamiltonian diffusions

Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold and $h_j \in C^\infty(M)$, $j = 0, ..., r$, smooth functions. Let $h : M \longrightarrow \mathbb{R}^{r+1}$ be the Hamiltonian function $m \longmapsto (h_0(m), \dots, h_r(m))$, and consider the semimartingale $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{r+1}$ given by $(t, \omega) \longmapsto \left(t, B_t^1(\omega), \dots, B_t^r(\omega)\right)$, where B^j, $j = 1, ..., r$, are r-independent Brownian motions. Lévy's characterization of Brownian motion shows (see for instance [P05, Theorem 40, page 87]) that $[B^j, B^i]_t = t\delta^{ji}$. In this setup, the equation (2.8) reads

$$f(\Gamma_\tau^h) - f(\Gamma_0^h) = \int_0^\tau \{f, h_0\}(\Gamma^h)dt + \sum_{j=1}^r \int_0^\tau \{f, h_j\}(\Gamma^h)\delta B^j \tag{2.19}$$

for any $f \in C^\infty(M)$. According to (2.10), the equivalent Itô version of this equation is

$$f(\Gamma_\tau^h) - f(\Gamma_0^h) = \int_0^\tau \{f, h_0\}(\Gamma^h)dt + \sum_{j=1}^r \int_0^\tau \{f, h_j\}(\Gamma^h)dB^j + \int_0^\tau \{\{f, h_j\}, h_j\}(\Gamma^h)dt.$$

Equation (2.19) may be interpreted as a stochastic perturbation of the classical Hamilton equations associated to h_0, that is,

$$\frac{d(f \circ \gamma)}{dt}(t) = \{f, h_0\}(\gamma(t)).$$

by the r Brownian motions B^j. These equations have been studied by Bismut in [B81] in the particular case in which the Poisson manifold $(M, \{\cdot, \cdot\})$ is just the symplectic Euclidean space $\mathbb{R}^{2n}$ with the canonical symplectic form. He refers to these particular processes as ***Hamiltonian diffusions***.

If we apply Proposition 2.13 to the stochastic Hamiltonian system (2.2.1), we obtain a generalization to Poisson manifolds of a result originally formulated by Bismut (see [B81, Théorèmes 4.1 and 4.2, page 231]) for Hamiltonian diffusions. See also [M99].

Proposition 2.19 *Consider the stochastic Hamiltonian system introduced in (2.2.1). Then $f \in C^\infty(M)$ is a conserved quantity if and only if*

$$\{f, h_0\} = \{f, h_1\} = \ldots = \{f, h_r\} = 0. \tag{2.20}$$

Proof. If (2.20) holds then f is clearly a conserved quantity by Proposition 2.13. Conversely, notice that as $[B^i, B^j] = t\delta^{ij}$, $i, j \in \{1, \dots, r\}$, and $X^0(t, \omega) = t$ is a finite variation process then $[X^i, X^j] = 0$ for any $i, j \in \{0, 1, \dots, r\}$ such that $i \neq j$. Consequently, by Proposition 2.13, if f is a conserved quantity then

$$\{f, h_1\} = \ldots = \{f, h_r\} = 0. \tag{2.21}$$

Moreover, (2.19) reduces to

$$\int_0^{\tau} \{f, h_0\} (\Gamma^h) dt = 0,$$

for any Hamiltonian semimartingale Γ^h and any stopping time $\tau \leq \zeta^h$. Suppose that $\{f, h_0\} (m_0) > 0$ for some $m_0 \in M$. By continuity there exists a compact neighborhood U of m_0 such that $\{f, h_0\}|_U > 0$. Take Γ^h the Hamiltonian semimartingale with initial condition $\Gamma_0^h = m_0$, and let ξ be the first exit time of U for Γ^h. Then, defining $\tau := \xi \wedge \zeta$,

$$\int_0^{\tau} \{f, h_0\} (\Gamma^h) dt \geq \int_0^{\tau} \min \{\{f, h_0\} (m) \mid m \in U\} dt > 0,$$

which contradicts (2.21). Therefore, $\{f, h_0\} = 0$ also, as required. ■

Remark 2.20 Notice that, unlike what happens for standard deterministic Hamiltonian systems, the energy h_0 of a Hamiltonian diffusion does not need to be conserved if the other components of the Hamiltonian are not involution with h_0. This is a general fact about stochastic Hamiltonian systems that makes them useful in the modeling of dissipative phenomena. We see more of this in the next example.

2.2.2 Integrable stochastic Hamiltonian dynamical systems

Let (M, ω) be a $2n$-dimensional manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a semimartingale, and $h : M \to V^*$ such that $h = \sum_{i=1}^{r} h_i \epsilon^i$, with $\left\{\epsilon^1, ..., \epsilon^r\right\}$ a basis of V^*. Let H be the associated Stratonovich operator in (2.6).

Suppose that there exists a family of functions $\{f_{r+1}, ..., f_n\} \subset C^{\infty}(M)$ such that the n-functions $\{f_1 := h_1, ..., f_r := h_r, f_{r+1}, ..., f_n\} \subset C^{\infty}(M)$ are in Poisson involution, that is, $\{f_i, f_j\} = 0$, for any $i, j \in \{1, ..., n\}$. Moreover, assume that $F := (f_1, ..., f_n)$ satisfies the hypotheses of the Liouville-Arnold Theorem [A89]: F has compact and connected fibers and its components are independent. In this setup, we will say that the stochastic Hamiltonian dynamical system associated to H is ***integrable***.

As it was already the case for standard (Liouville-Arnold) integrable systems, there is a symplectomorphism that takes (M, ω) to $\left(\mathbb{T}^n \times \mathbb{R}^n, \sum_{i=1}^{n} \mathbf{d}\theta^i \wedge \mathbf{d}I_i\right)$ and for which $F \equiv F(I_1, ..., I_n)$. In particular, in the action-angle coordinates $\left(I_1, ..., I_n, \theta^1, ..., \theta^n\right)$, $h_j \equiv h_j(I_1, ..., I_n)$ with $j \in \{1, ..., r\}$. In other words, the components of the Hamiltonian function depend only on the actions $\mathbf{I} := (I_1, ..., I_n)$. Therefore, for any random variable Γ_0 and any $i \in \{1, ..., n\}$

$$I_i(\Gamma) - I_i(\Gamma_0) = \sum_{j=1}^{r} \int \{I_i, h_j(\mathbf{I})\} (\Gamma) \, \delta X^j = 0 \tag{2.22a}$$

$$\theta^i(\Gamma) - \theta^i(\Gamma_0) = \sum_{j=1}^{r} \int \left\{\theta^i, h_j(\mathbf{I})\right\} (\Gamma) \, \delta X^j = \sum_{j=1}^{r} \int \frac{\partial h_j}{\partial I_i} (\Gamma) \, \delta X^j. \tag{2.22b}$$

Consequently, the tori determined by fixing $\mathbf{I} =$ constant are left invariant by the stochastic flow associated to (2.22). In particular, as the paths of the solutions are contained in compact sets, the stochastic flow is defined for any time and the flow is complete. Moreover, the restriction of this stochastic differential equation to the torus given by say, $\mathbf{I}_0$, yields the solution

$$\theta^i(\Gamma) - \theta^i(\Gamma_0) = \sum_{j=1}^{r} \omega_j(\mathbf{I}_0) X^j, \tag{2.23}$$

where $\omega_j(\mathbf{I}_0) := \frac{\partial h_j}{\partial I_i}(\mathbf{I}_0)$ and where we have assumed that $X_0 = 0$. Expression (2.23) clearly resembles the integration that can be carried out for deterministic integrable systems.

Additionally, the Haar measure $\mathbf{d}\theta^1 \wedge ... \wedge \mathbf{d}\theta^n$ on each invariant torus is left invariant by the stochastic flow (see Theorem 2.11 and [Li08]). Therefore, if we can ensure that there exists a unique invariant measure μ (for instance, if (2.23) defines a non-degenerate diffusion on the torus $\mathbb{T}^n$, the invariant measure is unique up to a multiplicative constant by the compactness of $\mathbb{T}^n$ (see [IW89, Proposition 4.5])) then μ coincides necessarily with the Haar measure.

2.2.3 The Langevin equation and viscous damping

Hamiltonian stochastic differential equations can be used to model dissipation phenomena. The simplest example in this context is the damping force experienced by a particle in motion in a viscous fluid. This dissipative phenomenon is usually modeled using a force in Newton's second law that depends linearly on the velocity of the particle (see for instance [LL76, §25]). The standard microscopic description of this motion is carried out using the ***Langevin*** stochastic differential equation (also called the ***Orstein-Uhlenbeck*** equation) that says that the velocity $\dot{q}(t)$ of the particle with mass m is a stochastic process that solves the stochastic differential equation

$$m\, d\dot{q}(t) = -\lambda \dot{q}(t) dt + b dB_t, \tag{2.24}$$

where $\lambda > 0$ is the damping coefficient, b is a constant, and B_t is a Brownian motion. A common physical interpretation for this equation (see [CH06]) is that the Brownian motion models random instantaneous bursts of momentum that are added to the particle by collision with lighter particles, while the mean effect of the collisions is the slowing down of the particle. This fact is mathematically described by saying that the expected value $q_e := \mathrm{E}[q]$ of the process q determined by (2.24) satisfies the ordinary differential equation $\ddot{q}_e = -\lambda \dot{q}_e$. Even though this description is accurate it is not fully satisfactory given that it *does not provide any information about the mechanism that links the presence of the Brownian perturbation to the emergence of damping* in the equation. In order for the physical explanation to be complete, a relation between the coefficients b and λ should be provided in such a way that *the damping vanishes when the Brownian collisions disappear*, that is, $\lambda = 0$ when $b = 0$.

We now show that the motion of a particle of mass m in one dimension subjected to viscous damping with coefficient λ and to a harmonic potential with Hooke constant k is a Hamiltonian stochastic differential equation. More explicitly, we will give a stochastic Hamiltonian system such that the expected value q_e of its solution semimartingales satisfies the ordinary differential equation of the damped harmonic oscillator, that is,

$$m\, \ddot{q}_e(t) = -\lambda \dot{q}_e(t) - k q_e(t). \tag{2.25}$$

This description provides a mathematical mechanism by which the stochastic perturbations in the system generate an average damping.

Consider $\mathbb{R}^2$ with its canonical symplectic form and let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ be the real semimartingale given by $X_t(\omega) = (t + \nu B_t(\omega))$ with $\nu \in \mathbb{R}$ and B_t a Brownian motion. Let now $h : \mathbb{R}^2 \to \mathbb{R}$ be the energy of a harmonic oscillator, that is, $h(q,p) := \frac{1}{2m}p^2 + \frac{1}{2}\rho q^2$. By (2.10), the solution semimartingales Γ^h of the Hamiltonian stochastic equations associated to h and X satisfy

$$q(\Gamma^h) - q(\Gamma_0^h) = \frac{1}{2m}\int \left(2p(\Gamma_t^h) - \nu^2 \rho q(\Gamma_t^h)\right) dt + \frac{\nu}{m}\int p(\Gamma_t^h) dB_t, \tag{2.26}$$

$$p(\Gamma^h) - p(\Gamma_0^h) = -\frac{\rho}{2m}\int \left(\nu^2 p(\Gamma_t^h) + 2mq(\Gamma_t^h)\right) dt - \nu\rho \int q(\Gamma_t^h) dB_t. \tag{2.27}$$

Given that $E\left[\int p(\Gamma_t^h) dB_t\right] = E\left[\int q(\Gamma_t^h) dB_t\right] = 0$, if we denote

$$q_e(t) := E\left[q(\Gamma_t^h)\right], \qquad p_e(t) := E\left[p(\Gamma_t^h)\right],$$

Fubini's Theorem guarantees that

$$\dot{q}_e(t) = \frac{1}{m}p_e(t) - \frac{\nu^2\rho}{2m}q_e(t) \quad \text{and} \quad \dot{p}_e(t) = -\frac{\nu^2\rho}{2m}p_e(t) - \rho q_e(t). \tag{2.28}$$

From the first of these equations we obtain that

$$p_e(t) = m\dot{q}_e + \frac{\nu^2\rho}{2}q_e$$

whose time derivative is

$$\dot{p}_e(t) = m\ddot{q}_e + \frac{\nu^2\rho}{2}\dot{q}_e.$$

These two equations substituted in the second equation of (2.28) yield

$$m\ddot{q}_e(t) = -\nu^2\rho\dot{q}_e(t) - \rho\left(\frac{\nu^4\rho}{4m} + 1\right)q_e(t), \tag{2.29}$$

that is, the expected value of the position of the Hamiltonian semimartingale Γ^h associated to h and X satisfies the differential equation of a damped harmonic oscillator (2.25) with constants

$$\lambda = \nu^2\rho \quad \text{and} \quad k = \rho\left(\frac{\nu^4\rho}{4m} + 1\right).$$

Notice that the dependence of the damping and elastic constants on the coefficients of the system is physically reasonable. For instance, we see that the more intense the stochastic perturbation is, that is, the higher ν is, the stronger the damping becomes ($\lambda = \nu^2\rho$ increases). In particular, if there is no stochastic perturbation, that is, if $\nu = 0$, then the damping vanishes, $k = \rho$ and (2.29) becomes the differential equation of a free harmonic oscillator of mass m and elastic constant ρ.

The stability of the resting solution. It is easy to see that the constant process $\Gamma_t(\omega) = (0,0)$, for all $t \in \mathbb{R}$ and $\omega \in \Omega$ is an equilibrium solution of (2.26) and (2.27). One can show using the stochastic Dirichlet's criterion (Theorem 2.15) that this equilibrium is almost surely Lyapunov stable since the Hamiltonian function h is a strongly conserved quantity (by (2.8)) that exhibits a critical point at the origin with definite Hessian.

The Langevin equation. In the previous paragraphs we succeeded in providing a microscopic Hamiltonian description of the harmonic oscillator subjected to Brownian perturbations whose macroscopic counterpart via expectations yields the equations of the damped harmonic oscillator. In view of this, is such a stochastic Hamiltonian description available for the pure Langevin equation (2.24)? The answer is no. More specifically, it can be easily shown (proceed by contradiction) that (2.24) cannot be written as a stochastic Hamiltonian differential equation on $\mathbb{R}^2$ with its canonical symplectic form with a noise semimartingale of the form $X_t(\omega) = (f_0(t, B_t), f_1(t, B_t))$ and a Hamiltonian function $h(q,p) = (h_0(q,p), h_1(q,p))$, $f_0, f_1, h_0, h_1 \in C^\infty(\mathbb{R})$. Nevertheless, if we put aside for a moment the stochastic Hamiltonian category and we use Itô integration, the Langevin equation can still be written in phase space, that is,

$$dq_t = v_t dt, \quad dv_t = -\lambda v_t dt + b dB_t, \tag{2.30}$$

as a stochastic perturbation of a deterministic system, namely, a free particle whose evolution is given by the differential equations

$$dq_t = v_t dt \quad \text{and} \quad dv_t = 0. \tag{2.31}$$

Let $\{u^1, u^2\}$ be global coordinates on $\mathbb{R}^2$ associated to the canonical basis $\{e_1, e_2\}$ and consider the global basis $\{d_2 u^i, d_2 u^i \cdot d_2 u^j\}_{i,j=1,2}$ of $\tau^*\mathbb{R}^2$. Define a dual Schwartz operator $\mathcal{S}^*(x, (q,v)) : \tau^*_{(q,v)}\mathbb{R}^2 \longrightarrow \tau^*_x\mathbb{R}^2$ characterized by the relations

$$d_2 q \longmapsto v d_2 u^1, \quad d_2 v \longmapsto b d_2 u^2 - \lambda v \left(d_2 u^2 \cdot d_2 u^2\right),$$

where $(q,v) \in \mathbb{R}^2$ is an arbitrary point in phase space and $x \in \mathbb{R}^2$. If $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^2$ is such that $X(t,\omega) = (t, bB_t(\omega))$, for any $(t,\omega) \in \mathbb{R}_+ \times \Omega$, it is immediate to see that the Itô equations associated to $\mathcal{S}^*$ and X are (2.30). Moreover, if we set $b = 0$, that is, we switch off the Brownian perturbation then we recover (2.31), as required.

2.2.4 Brownian motions on manifolds

The mathematical formulation of Brownian motions (or Wiener processes) on manifolds has been the subject of much research and it is a central topic in the study of stochastic processes on manifolds (see [IW89, Chapter 5], [E89, Chapter V], and references therein for a good general review of this subject).

In the following paragraphs we show that Brownian motions can be defined in a particularly simple way using the stochastic Hamilton equations introduced in Definition 2.2. More specifically we will show that Brownian motions on manifolds can be obtained as the projections onto the base space of very simple Hamiltonian stochastic semimartingales defined on the cotangent bundle of the manifold or of its orthonormal frame bundle, depending on the availability or not of a parallelization for the manifold in question.

We will first present the case in which the manifold in question is parallelizable or, equivalently, when the coframe bundle on the manifold admits a global section, for the construction is particularly simple in this situation. The parallelizability hypothesis is verified by many important examples. For instance, any Lie group is parallelizable; the spheres S^1, S^3, and S^7 are parallelizable too. At the end of the section we describe the general case.

The notion of manifold valued Brownian motion that we will use is the following. A M-valued process Γ is called a Brownian motion on (M, g), with g a Riemannian metric on M, whenever Γ is continuous and adapted and for every $f \in C^\infty(M)$

$$f(\Gamma) - f(\Gamma_0) - \frac{1}{2}\int \Delta_M f(\Gamma) dt$$

is a local martingale. We recall that the Laplacian $\Delta_M(f)$ is defined as $\Delta_M(f) = \mathrm{Tr}\,(\mathrm{Hess}\, f)$, for any $f \in C^\infty(M)$, where $\mathrm{Hess}\, f := \nabla(\nabla f)$, with $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, the Levi-Civita connection of g. $\mathrm{Hess}\, f$ is a symmetric $(0,2)$-tensor such that for any $X, Y \in \mathfrak{X}(M)$,

$$\mathrm{Hess}\, f(X, Y) = X\left[g(\mathrm{grad}\, f, Y)\right] - g(\mathrm{grad}\, f, \nabla_X Y). \tag{2.32}$$

Brownian motions on parallelizable manifolds. Suppose that the n-dimensional manifold (M, g) is parallelizable and let $\{Y_1, ..., Y_n\}$ be a family of vector fields such that for each $m \in M$, $\{Y_1(m), ..., Y_n(m)\}$ forms a basis of $T_m M$ (a parallelization). Applying the Gram-Schmidt orthonormalization procedure if necessary, we may suppose that this parallelization is orthonormal, that is, $g(Y_i, Y_j) = \delta_{ij}$, for any $i, j = 1, ..., n$.

Using this structure we are going to construct a stochastic Hamiltonian system on the cotangent bundle T^*M of M, endowed with its canonical symplectic structure, and we will show that the projection of the solution semimartingales of this system onto M are M-valued Brownian motions in the sense specified above. Let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{n+1}$ be the semimartingale given by $X(t, \omega) := (t, B_t^1(\omega), \ldots, B_t^n(\omega))$, where B^j, $j = 1, ..., n$, are n-independent Brownian motions and let $h = (h_0, h_1, \ldots, h_n) : T^*M \to \mathbb{R}^{n+1}$ be the function whose components are given by

$$\begin{array}{rccl} h_0 : & T^*M & \longrightarrow & \mathbb{R} \\ & \alpha_m & \longmapsto & -\frac{1}{2}\sum_{j=1}^n \langle \alpha_m, \left(\nabla_{Y_j} Y_j\right)(m)\rangle \end{array} \quad \text{and}$$

$$\begin{array}{rccl} h_j : & T^*M & \longrightarrow & \mathbb{R} \\ & \alpha_m & \longmapsto & \langle \alpha_m, Y_j(m)\rangle. \end{array} \tag{2.33}$$

We will now study the projection onto M of Hamiltonian semimartingales Γ^h that have X as stochastic component and h as Hamiltonian function and will prove that they are M-valued Brownian motions. In order to do so we will be particularly interested in the ***projectable functions*** f of T^*M, that is, the functions $f \in C^\infty(T^*M)$ that can be written as $f = \overline{f} \circ \pi$ with $\overline{f} \in C^\infty(M)$ and $\pi : T^*M \to M$ the canonical projection.

We start by proving that for any projectable function $f = \overline{f} \circ \pi \in C^\infty(T^*M)$

$$\{f, h_0\} = g\left(\mathrm{grad}\, \overline{f}, -\frac{1}{2}\sum_{j=1}^n \nabla_{Y_j} Y_j\right) \quad \text{and} \quad \{f, h_j\} = g\left(\mathrm{grad}\, \overline{f}, Y_j\right), \tag{2.34}$$

and where $\{\cdot,\cdot\}$ is the Poisson bracket associated to the canonical symplectic form on T^*M. Indeed, let U a Darboux patch for T^*M with associated coordinates $\left(q^1,\ldots,q^n,p_1,\ldots,p_n\right)$ such that $\{q^i,p_j\}=\delta^i_j$. There exists functions $f^k_j\in C^\infty(\pi(U))$, with $k,j\in\{1,\ldots,n\}$ such that the vector fields may be locally written as $Y_j=\sum_{k=1}^n f^k_j\frac{\partial}{\partial q^k}$. Moreover, $h_j\left(q,p\right)=\sum_{k=1}^n f^k_j\left(q\right)p_k$ and

$$\begin{aligned}\{f,h_j\}&=\left\{\overline{f}\circ\pi,\sum_{k=1}^n f^k_j p_k\right\}=\sum_{k=1}^n f^k_j\left\{\overline{f}\circ\pi,p_k\right\}=\sum_{k,i=1}^n f^k_j\frac{\partial(\overline{f}\circ\pi)}{\partial q^i}\left\{q^i,p_k\right\}\\&=\sum_{k,i=1}^n f^k_j\delta^i_k\frac{\partial\overline{f}}{\partial q^i}=Y_j[\overline{f}]\circ\pi=g\left(\operatorname{grad}\overline{f},Y_j\right)\circ\pi,\end{aligned}$$

as required. The first equality in (2.34) is proved analogously. Notice that the formula that we just proved shows that if f is projectable then so is $\{f,h_j\}$, with $j\in\{1,\ldots,n\}$. Hence, using (2.34) again and (2.32) we obtain that

$$\left\{\{f,h_j\},h_j\right\}=Y_j\left[g\left(\operatorname{grad}\overline{f},Y_j\right)\right]\circ\pi=\operatorname{Hess}\overline{f}\left(Y_j,Y_j\right)\circ\pi+g\left(\operatorname{grad}\overline{f},\nabla_{Y_j}Y_j\right)\circ\pi,\quad(2.35)$$

for $j\in\{1,\ldots,n\}$. Now, using (2.34) and (2.35) in (2.10) we have shown that for any projectable function $f=\overline{f}\circ\pi$, the Hamiltonian semimartingale Γ^h satisfies that

$$\begin{aligned}\overline{f}\circ\pi(\Gamma^h)-\overline{f}\circ\pi(\Gamma^h_0)&=\sum_{j=1}^n\int g\left(\operatorname{grad}\overline{f},Y_j\right)(\pi\circ\Gamma^h)dB^j_s\\&\quad+\frac{1}{2}\sum_{j=1}^n\int\operatorname{Hess}\overline{f}\left(Y_j,Y_j\right)(\pi\circ\Gamma^h)dt,\end{aligned}\quad(2.36)$$

or equivalently

$$\overline{f}\circ\pi(\Gamma^h)-\overline{f}\circ\pi(\Gamma^h_0)-\frac{1}{2}\int\Delta_M(\overline{f})(\pi\circ\Gamma^h)dt=\sum_{j=1}^n\int g\left(\operatorname{grad}\overline{f},Y_j\right)(\pi\circ\Gamma^h)dB^j_s.\quad(2.37)$$

Since $\sum_{i=1}^n\int g\left(\operatorname{grad}\overline{f},Y_j\right)(\overline{\Gamma}^h)dB^i$ is a local martingale (see [P05, Theorem 20, page 63]), $\pi(\Gamma^h)$ is a Brownian motion.

Brownian motions on Lie groups. Let now G be a (finite dimensional) Lie group with Lie algebra $\mathfrak{g}$ and assume that G admits a bi-invariant metric g, for example when G is Abelian or compact. This metric induces a pairing in $\mathfrak{g}$ invariant with respect to the adjoint representation of G on $\mathfrak{g}$. Let $\{\xi_1,\ldots,\xi_n\}$ be an orthonormal basis of $\mathfrak{g}$ with respect to this invariant pairing and let $\{\nu_1,\ldots,\nu_n\}$ be the corresponding dual basis of $\mathfrak{g}^*$. The infinitesimal generator vector fields $\{\xi_{1G},\ldots,\xi_{nG}\}$ defined by $\xi_{iG}(h)=T_eL_h\cdot\xi$, with $L_h:G\to G$ the left translation map, $h\in G$, $i\in\{1,\ldots n\}$, are obviously an orthonormal parallelization of G, that is $g(\xi_{iG},\xi_{jG}):=\delta_{ij}$. Since g is bi-invariant then $\nabla_XY=\frac{1}{2}[X,Y]$, for any $X,Y\in\mathfrak{X}(G)$ (see [O83, Proposition 9, page 304]), and hence $\nabla_{\xi_{iG}}\xi_{iG}=0$. Therefore, in this particular case the first component h_0 of the Hamiltonian function introduced in (2.33) is zero and we can hence take $h_G=(h_1,...,h_n)$ and

$X_G = \left(B_t^1, ..., B_t^n\right)$ when we consider the Hamilton equations that define the Brownian motion with respect to g.

As a special case of the previous construction that serves as a particularly simple illustration, we are going to explicitly build the **Brownian motion on a circle**. Let $S^1 = \{e^{i\theta} \mid \theta \in \mathbb{R}\}$ be the unit circle. The stochastic Hamiltonian differential equation for the semimartingale Γ^h associated to $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$, given by $X_t(\omega) := B_t(\omega)$, and the Hamiltonian function $h : TS^1 \simeq S^1 \times \mathbb{R} \to \mathbb{R}$ given by $h(e^{i\theta}, \lambda) := \lambda$, is simply obtained by writing (2.36) down for the functions $f_1(e^{i\theta}) := \cos\theta$ and $f_2(e^{i\theta}) := \sin\theta$ which provide us with the equations for the projections X^h and Y^h of Γ^h onto the OX and OY axes, respectively. A straightforward computation yields

$$dX^h = -Y^h dB - \frac{1}{2}X^h dt \quad \text{and} \quad dY^h = X^h dB - \frac{1}{2}Y^h dt, \tag{2.38}$$

which, incidentally, coincides with the equations proposed in expression (5.1.13) of [O03]. A solution of (2.38) is $(X_t^h, Y_t^h) = (\cos B_t, \sin B_t)$, that is, $\Gamma_t^h = e^{iB_t}$.

Brownian motions on arbitrary manifolds. Let (M, g) be a not necessarily parallelizable Riemannian manifold. In this case we will reproduce the same strategy as in the previous paragraphs but replacing the cotangent bundle of the manifold by the cotangent bundle of its orthonormal frame bundle.

Let $\mathcal{O}_x(M)$ be the set of orthonormal frames for the tangent space T_xM. The orthonormal frame bundle $\mathcal{O}(M) = \bigcup_{x\in M}\mathcal{O}_x(M)$ has a natural smooth manifold structure of dimension $n(n+1)/2$. We denote by $\pi : \mathcal{O}(M) \to M$ the canonical projection. We recall that a curve $\gamma : (-\varepsilon, \varepsilon) \subset \mathbb{R} \to \mathcal{O}(M)$ is called horizontal if γ_t is the parallel transport of γ_0 along the projection $\pi(\gamma_t)$. The set of tangent vectors of horizontal curves that contain a point $u \in \mathcal{O}(M)$ defines the horizontal subspace $H_u\mathcal{O}(M) \subset T_u\mathcal{O}(M)$, with dimension n. The projection $\pi : \mathcal{O}(M) \to M$ induces an isomorphism $T_u\pi : H_u\mathcal{O}(M) \to T_{\pi(u)}M$. On the orthonormal frame bundle, we have n horizontal vector fields Y_i, $i = 1, ..., n$, defined as follows. For each $u \in \mathcal{O}(M)$, let $Y_i(u)$ be the unique horizontal vector in $H_u\mathcal{O}(M)$ such that $T_u\pi(Y_i) = u_i$, where u_i is the ith unit vector of the orthonormal frame u. Now, given a smooth function $F \in C^\infty(\mathcal{O}(M))$, the operator

$$\Delta_{\mathcal{O}(M)}(F) = \sum_{i=1}^n Y_i[Y_i[F]]$$

is called Bochner's horizontal Laplacian on $\mathcal{O}(M)$. At the same time, we recall that the Laplacian $\Delta_M(f)$, for any $f \in C^\infty(M)$, is defined as $\Delta_M(f) = \mathrm{Tr}(\mathrm{Hess}\, f)$. These two Laplacians are related by the relation

$$\Delta_{\mathcal{O}(M)}(\pi^* f) = \Delta_M(f), \tag{2.39}$$

for any $f \in C^\infty(M)$ (see [H02]).

The Eells-Elworthy-Malliavin construction of Brownian motion can be summarized as follows. Consider the following stochastic differential equation on $\mathcal{O}(M)$ (see [IW89]):

$$\delta U_t = \sum_{i=1}^n Y_i(U_t)\,\delta B_t^i \tag{2.40}$$

where B^j, $j = 1, ..., n$, are n-independent Brownian motions. Using the conventions introduced in Subsection 1.4.4 the expression (2.40) is the Stratonovich stochastic differential equation associated to the Stratonovich operator:

$$\begin{array}{rccc} e(v,u): & T_v\mathbb{R}^n & \longrightarrow & T_u\mathcal{O}(M) \\ & v = \sum_{i=1}^n v^i e_i & \longmapsto & \sum_{i=1}^n v^i Y_i(u), \end{array}$$

where $\{e_1, \ldots, e_n\}$ is a fixed basis for $\mathbb{R}^n$. A solution of the stochastic differential equation (2.40) is called a horizontal Brownian motion on $\mathcal{O}(M)$ since, by the Itô formula,

$$F(U) - F(U_0) = \sum_{i=1}^n \int Y_i[F](U_s)\,\delta B_s^i = \sum_{i=1}^n \int Y_i[F](U_s)\,dB_s^i + \frac{1}{2}\int \Delta_{\mathcal{O}(M)}(F)(U_s)\,ds,$$

for any $F \in C^\infty(\mathcal{O}(M))$. In particular, if $F = \pi^*(f)$ for some $f \in C^\infty(M)$, by (2.39)

$$f(X) - f(X_0) = \sum_{i=1}^n \int Y_i[\pi^*(f)](U_s)\,dB_s^i + \frac{1}{2}\int \Delta_M f(X_s)\,ds,$$

where $X_t = \pi(U_t)$, which implies precisely that X_t is a Brownian motion on M.

In order to generate (2.40) as a Hamilton equation, we introduce the functions $h_i : T^*\mathcal{O}(M) \to \mathbb{R}$, $i = 1, ..., n$, given by $h_i(\alpha) = \langle \alpha, Y_i \rangle$. Recall that $T^*\mathcal{O}(M)$ being a cotangent bundle it has a canonical symplectic structure. Mimicking the computations carried out in the parallelizable case it can be seen that the Hamiltonian vector field X_{h_i} coincides with Y_i when acting on functions of the form $F \circ \pi_{T^*\mathcal{O}(M)}$, where $F \in C^\infty(\mathcal{O}(M))$ and $\pi_{T^*\mathcal{O}(M)}$ is the canonical projection $\pi_{T^*\mathcal{O}(M)} : T^*\mathcal{O}(M) \to \mathcal{O}(M)$. By (2.8), the Hamiltonian semimartingale Γ^h associated to $h = (h_1, ..., h_n)$ and to the stochastic Hamiltonian equations on $T^*\mathcal{O}(M)$ with stochastic component $X = (B_t^1, ..., B_t^n)$ is such that

$$\begin{aligned} &F \circ \pi_{T^*\mathcal{O}(M)}(\Gamma^h) - F \circ \pi_{T^*\mathcal{O}(M)}(\Gamma_0^h) \\ &= \sum_{i=1}^n \int \left\{F \circ \pi_{T^*\mathcal{O}(M)}, h_i\right\}(\Gamma_s^h)\delta B_s^i = \sum_{i=1}^n \int Y_i[F]\left(\pi_{T^*\mathcal{O}(M)}(\Gamma_s^h)\right)\delta B_s^i \end{aligned}$$

for any $F \in C^\infty(\mathcal{O}(M))$. This expression obviously implies that $U^h = \pi_{T^*\mathcal{O}(M)}(\Gamma^h)$ is a solution of (2.40) and consequently $X^h = \pi(U^h)$ is a Brownian motion on M.

2.2.5 The inverted pendulum with stochastically vibrating suspension point

The equation of motion for small angles of a damped inverted unit mass pendulum of length l with a vertically vibrating suspension point is

$$\ddot{\phi} = \left(\frac{\ddot{y}}{l} + \frac{g}{l}\right)\phi - \lambda\dot{\phi}, \tag{2.41}$$

where ϕ is the angle that measures the separation of the pendulum from the vertical upright position, $y = y(t)$ is the height of the suspension point (externally controlled), λ is the friction coefficient, and g is the gravity constant. By construction, the point $(\phi, \dot{\phi}) = (0, 0)$ corresponds to the upright equilibrium position. It can be shown that if the function $y(t)$ is of the form $y(t) = az(\omega t)$, with z periodic, the amplitude a is sufficiently small, and the frequency ω is sufficiently high, then this equilibrium becomes nonlinearly stable.

We now consider the case in which the external forcing of the suspension point is given by a continuous stochastic process $\dot{z} : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ such that $\dot{z}^2$ is continuous and stationary. Under this assumptions, the equation (2.41) becomes the stochastic differential equation

$$d\phi = \dot{\phi} dt, \quad d\dot{\phi} = \left(\frac{g}{l} \phi - \lambda \dot{\phi} \right) dt + \varepsilon^2 \omega^2 \phi d\dot{z}_t, \tag{2.42}$$

where $\varepsilon := \sqrt{a/l}$. Observe that this equation is not Hamiltonian unless the friction term $-\lambda \dot{\phi}$ vanishes ($\lambda = 0$), in which case one obtains a Hamiltonian stochastic system with Hamiltonian function $h(\phi, \dot{\phi}) = (\frac{1}{2}(l^2 \dot{\phi}^2 - l\phi^2), \frac{1}{4}(\varepsilon^2 \omega^2 \phi l)^2, -\frac{1}{2}(\varepsilon \omega \phi l)^2)$ and noise semimartingale $X_t = (t, [\dot{z}, \dot{z}], \dot{z})$ (the symplectic form is obviously $l^2 d\phi \wedge d\dot{\phi}$).

The stability of the upright position of the stochastically forced pendulum has been studied in [O06, I01], and references therein. In [O06] it is assumed that the noise has the fairly strong mixing property. We recall that a continuous, adapted, stationary process $\Gamma : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ has the fairly strong mixing property if $E\left[\Gamma_t^2\right] < \infty$, there exists a real function c such that $\int_0^\infty c(s)\, ds < \infty$, and for any $t > s$

$$\left\| E\left[\Gamma_t - E\left[\Gamma_t\right] \middle| \mathcal{F}_s \right] \right\|_{L^2} \leq c(t-s) \left\| \Gamma_s - E\left[\Gamma_s\right] \right\|_{L^2},$$

where $\left\|\cdot\right\|_{L^2}$ stands for the L^2 norm. For example, if x is the unique stationary solution with zero mean of the Itô equations

$$dx_t = y_t dt, \quad dy_t = -\left(x_t + y_t\right) dt + dB_t,$$

where B_t is a standard Brownian motion, then $\dot{x}_t^2 - \frac{1}{2} = y_t^2 - \frac{1}{2}$ has the fairly strong mixing property. Using this hypothesis, it can be shown [O06, Theorem 1] that if $z : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is a continuously differentiable and stationary process such that, for any $t \in \mathbb{R}_+$, $E\left[z_t\right] = 0$, $E\left[\exp\left(\varepsilon \left|z_t\right|\right)\right] < \infty$ if $\varepsilon = \sqrt{a/l}$ is sufficiently small, and the process $\dot{z}^2$ has the fairly strong mixing property, then the solution $(\phi, \dot{\phi}) = (0, 0)$ of (2.42) is exponentially stable in probability, if ε is sufficiently small and $\frac{g}{l\varepsilon^4} < E\left[\dot{z}^2\right]$. Moreover, Ovseyevich shows in [O06, Section 4] that if we put $\lambda = 0$ in (2.42) and we consider hence the inverted pendulum as a Hamiltonian system, then the equilibrium point $(\phi, \dot{\phi}) = (0, 0)$ is unstable.

2.3 Critical action principles for the stochastic Hamilton equations

Our goal in this section is showing that the stochastic Hamilton equations can be characterized by a variational principle that generalizes the one used in the classical deterministic situation. In the following pages we shall consider an exact symplectic manifold (M, ω), that is, there exist a one-form $\theta \in \Omega(M)$ such that $\omega = -\mathbf{d}\theta$. The archetypical example of an exact symplectic manifold is the cotangent bundle T^*Q of any manifold Q, with θ the Liouville one-form.

In the following pages we will proceed in two stages. In the first subsection we will construct a critical action principle based on using variations of the solution semimartingale using the flow of a vector field on the manifold. Even though this approach is extremely natural and mathematically very tractable it yields a variational principle (Theorem 2.29) that does not fully characterize the stochastic Hamilton's equations. In order to obtain such a characterization one needs to use more general variations associated to the flows of vector fields defined on the solution semimartingale, that is, they depend on Ω. This complicates considerably the formulation and will be treated separately in the second subsection.

Definition 2.21 *Let $(M, \omega = -\mathbf{d}\theta)$ be an exact symplectic manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a semimartingale taking values on the vector space V, and $h : M \to V^*$ a Hamiltonian function. We denote by $\mathcal{S}(M)$ and $\mathcal{S}(\mathbb{R})$ the sets of M and real-valued semimartingales, respectively. We define the* ***stochastic action*** *associated to h as the map $S : \mathcal{S}(M) \to \mathcal{S}(\mathbb{R})$ given by*

$$S(\Gamma) = \int \langle \theta, \delta\Gamma \rangle - \int \left\langle \widehat{h}(\Gamma), \delta X \right\rangle,$$

where in the previous expression, $\widehat{h}(\Gamma) : \mathbb{R}_+ \times \Omega \to V \times V^*$ is given by $\widehat{h}(\Gamma)(t, \omega) := (X_t(\omega), h(\Gamma_t(\omega)))$.

2.3.1 Variations involving vector fields on the phase space

Definition 2.22 *Let M be a manifold, $F : \mathcal{S}(M) \to \mathcal{S}(\mathbb{R})$ a map, and $\Gamma \in \mathcal{S}(M)$. A local one-parameter group of diffeomorphisms $\varphi : \mathcal{D} \subset \mathbb{R} \times M \to M$ is said to be* ***complete with respect*** *to Γ if there exists $\epsilon > 0$ such that $\varphi_s(\Gamma)$ is a well-defined process for any $s \in (-\epsilon, \epsilon)$. We say that F is* ***differentiable*** *at Γ in the direction of a local one parameter group of diffeomorphisms φ complete with respect to Γ, if for any sequence $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ such that $s_n \underset{n\to\infty}{\longrightarrow} 0$, the family*

$$X_n = \frac{1}{s_n}\left(F\left(\varphi_{s_n}(\Gamma)\right) - F(\Gamma)\right)$$

converges uniformly on compacts in probability (ucp) to a process that we will denote by $\left.\frac{d}{ds}\right|_{s=0} F(\varphi_s(\Gamma))$ and that is referred to as the ***directional derivative*** *of F at Γ in the direction of φ_s.*

Remark 2.23 Note that global one-parameter groups of diffeomorphisms (for instance, flows of complete vector fields) are complete with respect to any semimartingale. Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a M-valued continuous and adapted stochastic process and $A \subset M$ a set. We will denote by $\tau_A = \inf\{t > 0 \mid \Gamma_t(\omega) \notin A\}$ the ***first exit time*** of Γ with respect to A. We recall that τ_A is a stopping time if A is a Borel set. Additionally, let Γ be a semimartingale and K a compact set such that $\Gamma_0 \subset K$. Then, any local one-parameter group of diffeomorphisms φ is complete with respect to the stopped process Γ^{τ_K}. Note that this conclusion could also hold for certain non-compact sets.

The proof of the following proposition can be found in Section 2.5.1.

Proposition 2.24 *Let M be a manifold, $\alpha \in \Omega(M)$ a one-form, and $F : \mathcal{S}(M) \to \mathcal{S}(\mathbb{R})$ the map defined by $F(\Gamma) := \int \langle \alpha, \delta\Gamma \rangle$. Then F is differentiable in all directions. Moreover, if $\Gamma : \mathbb{R}_+ \times \Omega \to M$ is a continuous semimartingale, φ is an arbitrary local one-parameter group of diffeomorphisms complete with respect to Γ, and $Y \in \mathfrak{X}(M)$ is the vector field associated to φ, then*

$$\left.\frac{d}{ds}\right|_{s=0} F(\varphi_s(\Gamma)) = \left.\frac{d}{ds}\right|_{s=0} \int \langle \alpha, \delta(\varphi_s \circ \Gamma) \rangle = \left.\frac{d}{ds}\right|_{s=0} \int \langle \varphi_s^* \alpha, \delta\Gamma \rangle = \int \langle \pounds_Y \alpha, \delta\Gamma \rangle. \quad (2.43)$$

The symbol $\pounds_Y \alpha$ denotes the Lie derivative of α in the direction given by Y.

Corollary 2.25 *In the setup of Definition 2.21 let $\alpha = \omega^\flat(Y) \in \Omega(M)$, with $\omega^\flat$ the inverse of the vector bundle isomorphism $\omega^\sharp : T^*M \to TM$ induced by ω. Let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a continuous adapted semimartingale. φ an arbitrary local one-parameter group of diffeomorphisms complete with respect to Γ, and $Y \in \mathfrak{X}(M)$ the associated vector field. Then, the action S is differentiable at Γ in the direction of φ and the directional derivative is given by*

$$\left.\frac{d}{ds}\right|_{s=0} S(\varphi_s(\Gamma)) = -\int \langle \alpha, \delta\Gamma \rangle - \int \left\langle \mathbf{d}h\left(\omega^{\#}(\alpha)\right)(\Gamma), \delta X \right\rangle + \mathbf{i}_Y\theta(\Gamma) - \mathbf{i}_Y\theta(\Gamma_0). \quad (2.44)$$

Proof. It is clear from Proposition 2.24 that

$$\frac{1}{s}\left[\int \langle \varphi_s^*\theta - \theta, \delta\Gamma \rangle\right] \xrightarrow{s \to 0} \int \langle \pounds_Y \theta, \delta\Gamma \rangle$$

in *ucp*. The proof of that result can be easily adapted to show that ucp

$$\frac{1}{s}\left[\int \left\langle \left(\varphi_s^*\widehat{h} - \widehat{h}\right)(\Gamma), \delta X \right\rangle\right] \xrightarrow{s \to 0} \int \left\langle \left(\pounds_Y \widehat{h}\right)(\Gamma), \delta X \right\rangle.$$

Thus, using (1.49a) and $\alpha = \omega^\flat(Y) \in \Omega(M)$,

$$\begin{aligned}
\left.\frac{d}{ds}\right|_{s=0} S(\varphi_s(\Gamma)) &= \int \langle \pounds_Y \theta, \delta\Gamma \rangle - \int \left\langle \left(\pounds_Y \widehat{h}\right)(\Gamma), \delta X \right\rangle \\
&= \int \langle \mathbf{i}_Y \mathbf{d}\theta + \mathbf{d}(\mathbf{i}_Y\theta), \delta\Gamma \rangle - \int \langle \mathbf{d}h(Y)(\Gamma), \delta X \rangle \\
&= -\int \langle \alpha, \delta\Gamma \rangle + \int \langle \mathbf{d}(\mathbf{i}_Y\theta), \delta\Gamma \rangle - \int \left\langle \mathbf{d}h\left(\omega^{\#}(\alpha)\right)(\Gamma), \delta X \right\rangle \\
&= -\int \langle \alpha, \delta\Gamma \rangle - \int \left\langle \mathbf{d}h\left(\omega^{\#}(\alpha)\right)(\Gamma), \delta X \right\rangle + (\mathbf{i}_Y\theta)(\Gamma) - (\mathbf{i}_Y\theta)(\Gamma_0).
\end{aligned}$$

■

Corollary 2.26 (Noether's theorem) *In the setup of Definition 2.21, let $\varphi : \mathbb{R} \times M \to M$ be a one parameter group of diffeomorphisms and $Y \in \mathfrak{X}(M)$ the associated vector field. If the action $S : \mathcal{S}(M) \to \mathcal{S}(\mathbb{R})$ is invariant by φ, that is, $S(\varphi_s(\Gamma)) = S(\Gamma)$, for any $s \in \mathbb{R}$, then the function $\mathbf{i}_Y\theta$ is a strongly conserved quantity of the stochastic Hamiltonian system associated to $h : M \to V^*$.*

Proof. Let Γ^h be the Hamiltonian semimartingale associated to h with initial condition Γ_0. Since φ_s leaves invariant the action we have that

$$\left.\frac{d}{ds}\right|_{s=0} S\left(\varphi_s(\Gamma^h)\right) = 0$$

and hence by (2.44) we have that

$$0 = -\int \left\langle \alpha, \delta\Gamma^h \right\rangle - \int \left\langle \mathbf{d}h\left(\omega^{\#}(\alpha)\right)(\Gamma^h), \delta X \right\rangle + \mathbf{i}_Y\theta(\Gamma^h) - \mathbf{i}_Y\theta(\Gamma_0).$$

As Γ^h is the Hamiltonian semimartingale associated to h we have that

$$-\int \left\langle \alpha, \delta\Gamma^h \right\rangle = \int \left\langle \mathbf{d}h\left(\omega^{\#}(\alpha)\right)(\Gamma^h), \delta X \right\rangle$$

and hence $\mathbf{i}_Y\theta(\Gamma^h) = \mathbf{i}_Y\theta(\Gamma_0)$, as required. ■

Remark 2.27 The hypotheses of the previous corollary can be modified by requiring, instead of the invariance of the action by φ_s, the existence of a function $F \in C^\infty(M)$ such that

$$\left.\frac{d}{ds}\right|_{s=0} S\left(\varphi_s(\Gamma^h)\right) = F(\Gamma) - F(\Gamma_0).$$

In that situation, the conserved quantity is $\mathbf{i}_Y\theta + F$.

Before we state the Critical Action Principle for the stochastic Hamilton equations we need one more definition.

Definition 2.28 *Let M be a manifold and A a set. We will say that a local one parameter group of diffeomorphisms $\varphi : \mathcal{D} \times M \to M$* ***fixes*** *$A$ if $\varphi_s(y) = y$ for any $y \in A$ and any $s \in \mathbb{R}$ such that $(s, y) \in \mathcal{D}$. The corresponding vector field $Y \in \mathfrak{X}(M)$ given by $Y(m) = \left.\frac{d}{ds}\right|_{s=0} \varphi_s(m)$ satisfies that $Y|_A = 0$.*

Theorem 2.29 (First Critical Action Principle) *Let $(M, \omega = -d\theta)$ be an exact symplectic manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a semimartingale taking values on the vector space V such that $X_0 = 0$, and $h : M \to V^*$ a Hamiltonian function. Let $m_0 \in M$ be a point in M and $\Gamma : \mathbb{R}_+ \times \Omega \to M$ a continuous semimartingale such that $\Gamma_0 = m_0$. Let K be a compact set that contains the point m_0. If the semimartingale Γ satisfies the stochastic Hamilton equations (2.7) (with initial condition $\Gamma_0 = m_0$) up to time τ_K then for any local one-parameter group of diffeomorphisms φ that fixes the set $\{m_0\} \cup \partial K$ we have*

$$\mathbf{1}_{\{\tau_K < \infty\}} \left[\left.\frac{d}{ds}\right|_{s=0} S\left(\varphi_s\left(\Gamma^{\tau_K}\right)\right) \right]_{\tau_K} = 0 \quad a.s.. \tag{2.45}$$

Proof. We start by emphasizing that when we write that Γ satisfies the stochastic Hamiltonian equations (2.7) up to time τ_K we mean that

$$\left(\int \langle \beta, \delta\Gamma \rangle + \int \left\langle \mathbf{d}h\left(\omega^{\#}(\beta)\right)(\Gamma), \delta X \right\rangle \right)^{\tau_K} = 0.$$

For the sake of simplicity in our notation we define the linear operator Ham : $\Omega(M) \to \mathcal{S}(\mathbb{R})$ given by

$$\mathrm{Ham}\,(\beta) := \left(\int \langle \beta, \delta\Gamma \rangle + \int \left\langle \mathbf{d}h\left(\omega^{\#}\left(\beta\right)\right)\left(\Gamma\right), \delta X \right\rangle\right), \qquad \beta \in \Omega\left(M\right).$$

Suppose now that the semimartingale Γ satisfies the stochastic Hamilton equations up to time τ_K. Let φ be a local one-parameter group of diffeomorphisms that fixes $\{m_0\} \cup \partial K$, and let $Y \in \mathfrak{X}(M)$ be the associated vector field. Then, taking $\alpha = \omega^{\flat}\left(Y\right)$, we have by Corollary 2.25,

$$\left.\frac{d}{ds}\right|_{s=0} S\left(\varphi_s\left(\Gamma^{\tau_K}\right)\right) = -\int \langle \alpha, \delta\Gamma^{\tau_K} \rangle - \int \left\langle \mathbf{d}h\left(\omega^{\#}\left(\alpha\right)\right)\left(\Gamma^{\tau_K}\right), \delta X \right\rangle + i_Y\theta\left(\Gamma^{\tau_K}\right), \tag{2.46}$$

since $Y\left(m_0\right) = 0$ and hence $\mathbf{i}_Y\theta\left(\Gamma_0\right) = 0$. Additionally, since Γ is continuous, $\mathbf{1}_{\{\tau_K<\infty\}}\Gamma_{\tau_K} \in \partial K$ and $Y|_{\partial K} = 0$. Hence,

$$\mathbf{1}_{\{\tau_K<\infty\}}\left[\left.\frac{d}{ds}\right|_{s=0} S\left(\varphi_s\left(\Gamma^{\tau_K}\right)\right)\right]_{\tau_K} = -\mathbf{1}_{\{\tau_K<\infty\}}\left[\int \langle \alpha, \delta\Gamma^{\tau_K} \rangle + \int \left\langle \mathbf{d}h\left(\omega^{\#}\left(\alpha\right)\right)\left(\Gamma^{\tau_K}\right), \delta X \right\rangle\right]_{\tau_K}.$$

Now, Proposition A.1 in the Appendix and the hypothesis on Γ satisfying Hamilton's equation guarantee that the previous expression equals

$$\begin{aligned}
&\mathbf{1}_{\{\tau_K<\infty\}}\left[\left.\frac{d}{ds}\right|_{s=0} S\left(\varphi_s\left(\Gamma^{\tau_K}\right)\right)\right]_{\tau_K} \\
&= -\mathbf{1}_{\{\tau_K<\infty\}}\left[\left[\int \langle \alpha, \delta\Gamma^{\tau_K} \rangle + \int \left\langle \mathbf{d}h\left(\omega^{\#}\left(\alpha\right)\right)\left(\Gamma^{\tau_K}\right), \delta X \right\rangle\right]^{\tau_K}\right]_{\tau_K} \\
&= -\mathbf{1}_{\{\tau_K<\infty\}}\left[\left[\int \langle \alpha, \delta\Gamma \rangle + \int \left\langle \mathbf{d}h\left(\omega^{\#}\left(\alpha\right)\right)\left(\Gamma\right), \delta X \right\rangle\right]^{\tau_K}\right]_{\tau_K} \\
&= -\mathbf{1}_{\{\tau_K<\infty\}}\left[\mathrm{Ham}(\alpha)_{\tau_K}\right] = 0 \ \text{ a.s.},
\end{aligned}$$

as required. ■

Remark 2.30 The relation between the Critical Action Principle stated in Theorem 2.29 and the classical one for Hamiltonian mechanics is not straightforward since the categories in which both are formulated are very much different; more specifically, the differentiability hypothesis imposed on the solutions of the deterministic principle is not a reasonable assumption in the stochastic context and this has serious consequences. For example, unlike the situation encountered in classical mechanics, Theorem 2.29 does not admit a converse within the set of hypotheses in which it is formulated. In order to elaborate a little bit more on this question let $(M, \omega = -d\theta)$ be an exact symplectic manifold, take the Hamiltonian function $h \in C^{\infty}(M)$, and consider the stochastic Hamilton equations with trivial stochastic component $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ given by $X_t(\omega) = t$. As we saw in Remark 2.7 the paths of the semimartingales that solve these stochastic Hamilton equations are the smooth curves that integrate the standard Hamilton equations. In this situation the action reads

$$S\left(\Gamma\right) = \int \langle \theta, \delta\Gamma \rangle - \int h\left(\Gamma_s\right) ds.$$

If the path $\Gamma_t(\omega)$ is differentiable then the integral $\left(\int \langle \theta, \delta\Gamma \rangle\right)(\omega)$ reduces to the Riemann integral $\int_{\Gamma_t(\omega)} \theta$ and $S(\Gamma)(\omega)$ coincides with the classical action. In particular, if Γ is a solution of the stochastic Hamilton equations then the paths $\Gamma_t(\omega)$ are necessarily differentiable (see Remark 2.7), they satisfy the standard Hamilton equations, and hence make the action critical. The following elementary example shows that the converse is not necessarily true: one may have semimartingales that satisfy (2.45) and that do not solve the Hamilton equations up to time τ_K. We will consider a deterministic example. Let m_0, $m_1 \in M$ be two points. Suppose there exists an integral curve $\gamma : [t_0, t_1] \to M$ of the Hamiltonian vector field X_h defined on some time interval $[t_0, t_1]$ such that $\gamma(t_0) = m_0$ and $\gamma(t_1) = m_1$. Define the continuous and piecewise smooth curve $\sigma : [0, t_1] \to M$ as follows:

$$\sigma(t) = \begin{cases} m_0 & \text{if } t \in [0, t_0] \\ \gamma(t) & \text{if } t \in [t_0, t_1]. \end{cases}$$

Let φ be a local one-parameter group of diffeomorphisms that fixes $\{m_0, m_1\}$. Then by (2.44)

$$\left[\left.\frac{d}{ds}\right|_{s=0} S(\varphi_s(\sigma))\right]_t = -\int_{\sigma|_{[0,t]}} \alpha + \int_0^t \langle \alpha, X_h \rangle (\sigma(t))\, dt + \langle \theta(\sigma(t)), Y(\sigma(t)) \rangle - \langle \theta(m_0), Y(m_0) \rangle,$$

where $Y(m) = \left.\frac{d}{ds}\right|_{s=0} \varphi_s(m)$, for any $m \in M$ and $\alpha = \omega^\flat(Y)$. Using that σ satisfies the Hamilton equations on $[t_0, t_1]$ and $\alpha(m_0) = 0$, it is easy to see that

$$\left[\left.\frac{d}{ds}\right|_{s=0} S(\varphi_s(\sigma))\right]_{t_1} = 0,$$

that is, σ makes the action critical. However, it does not satisfy the Hamilton equations on the interval $[0, t_1]$, because they do not hold on $(0, t_0)$. This shows that the converse of the statement in Theorem 2.29 is not necessarily true. In the following subsection we will obtain such a converse by generalizing the set of variations allowed in the variational principle.

2.3.2 Variations involving vector fields on the solution semimartingale

We start by spelling out the variations that we will use in order to obtain a converse to Theorem 2.29.

Definition 2.31 *Let M be a manifold and Γ a M-valued semimartingale. Let $s_0 > 0$; we say that the map $\Sigma : (-s_0, s_0) \times \mathbb{R}_+ \times \Omega \to M$ is a* ***pathwise variation*** *of Γ whenever $\Sigma_t^0 = \Gamma_t$ for any $t \in \mathbb{R}_+$ a.s.. We say that the pathwise variation Σ of Γ* ***converges uniformly*** *to Γ whenever the following properties are satisfied:*

(i) *For any $f \in C^\infty(M)$, $f(\Sigma^s) \to f(\Gamma)$ in ucp as $s \to 0$.*

(ii) *There exists a process $Y : \mathbb{R}_+ \times \Omega \to TM$ over Γ such that, for any $f \in C^\infty(M)$, the Stratonovich integral $\int Y[f]\, \delta X$ exists for any continuous real semimartingale X (this is for instance guaranteed if Y is a semimartingale) and, additionally, the increments $(f(\Sigma^s) - f(\Gamma))/\, s$ converge in ucp to $Y[f]$ as $s \to 0$. We will call such a Y the* ***infinitesimal generator*** *of Σ.*

We will say that Σ (respectively, Y) is ***bounded*** *when its image lies in a compact set of M (respectively, TM).*

The next proposition shows that, roughly speaking, there exist bounded pathwise variations that converge uniformly to a given semimartingale with prescribed bounded infinitesimal generator.

Proposition 2.32 *Let Γ be a continuous M-valued semimartingale Γ, $K \subseteq M$ a compact set, and τ_K the first exit time of Γ from K. Let $Y : \mathbb{R}_+ \times \Omega \to TM$ be a bounded process over Γ^{τ_K} such that $\int Y[f]\,\delta X$ exists for any continuous real semimartingale X and for any $f \in C^\infty(M)$. Then, there exists a bounded pathwise variation Σ that converges uniformly to Γ^{τ_K} whose infinitesimal generator is Y.*

Proof. Let $\{(V_k, \varphi_k)\}_{k\in\mathbb{N}}$ be a countable open covering of M by coordinate patches such that any V_k is contained in a compact set. This covering is always available by the second countability of the manifold and Lindelöf's Lemma. Let $\{U_k\}_{k\in\mathbb{N}}$ be an open subcovering such that, if $U_k \subseteq V_i$ for some k, $i \in \mathbb{N}$, then $U_k \subsetneq V_i$. Let $\{\tau_m\}_{m\in\mathbb{N}}$ be a sequence of stopping times (available by Lemma 3.5 in [E89]) such that, a.s., $\tau_0 = 0$, $\tau_m \le \tau_{m+1}$, $\sup_m \tau_m = \infty$, and that, on each of the sets $[\tau_m, \tau_{m+1}] \cap \{\tau_{m+1} > \tau_m\}$ the semimartingale Γ takes values in the open set $U_{k(m)}$, for some $k(m) \in \mathbb{N}$. Since K is compact, it can be covered by a finite number of these open sets, i.e. $K \subseteq \cup_{j\in J} U_{k_j}$, where $|J| < \infty$. Let $x_{k_j} \equiv (x^1_{k_j}, \ldots, x^n_{k_j})$, $n = \dim(M)$ be a set of coordinate functions on $U_{k_j(m)}$ and $(x_{k_j}, v_{k_j}) \equiv (x^1_{k_j}, \ldots, x^n_{k_j}, v^1_{k_j}, \ldots, v^n_{k_j})$ the corresponding adapted coordinates for TM on $\pi^{-1}_{TM}\left(U_{k_j(m)}\right)$. Since Y is bounded and covers Γ^{τ_K}, and on $[\tau_m, \tau_{m+1}] \cap \{\tau_{m+1} > \tau_m\}$ the semimartingale Γ takes values in the open set $U_{k_j(m)}$, there exist a $s_{k_j} > 0$ such that, on $[\tau_m, \tau_{m+1}] \cap \{\tau_{m+1} > \tau_m\}$, the points $(x^1_{k_j}(\Gamma) + s v^1_{k_j}(Y), \ldots, x^n_{k_j}(\Gamma) + s v^n_{k_j}(Y))$ lie in the image of some coordinate patch V_{k_j} containing $U_{k_j(m)}$ for all $s \in \left(-s_{k_j}, s_{k_j}\right)$. Let $s_0 = \min_{j\in J}\left\{s_{k_j}\right\}$. Now, since the sets of the form $I_m := [\tau_m, \tau_{m+1}) \cap \{\tau_{m+1} > \tau_m\} \subset \mathbb{R}_+ \times \Omega$, $m \in \mathbb{N}$ form a disjoint partition of $\mathbb{R}_+ \times \Omega$ we define Σ as the map that for any $m \in \mathbb{N}$ satisfies

$$
\begin{array}{rcl}
\Sigma|_{I_m} : (-s_0, s_0) \times [\tau_m, \tau_{m+1}) \cap \{\tau_{m+1} > \tau_m\} & \longrightarrow & V_{k_j} \\
(s, t, \omega) & \longmapsto & \varphi_k^{-1}\left(x_{k_j}(\Gamma_t(\omega)) + s v_{k_j}(Y_t(\omega))\right).
\end{array}
$$

Observe that by construction the image of Σ is covered by a finite number of coordinated patches and therefore, by hypothesis, contained in a compact set. Σ is hence bounded. More specifically

$$
\{\Sigma^s_t(\omega) \mid (s, t, \omega) \in (-s_0, s_0) \times \mathbb{R} \times \Omega\} \subseteq \bigcup_{j\in J} V_{k_j}. \tag{2.47}
$$

It is immediate to see that Σ is a pathwise variation which converges uniformly to Γ^{τ_K}. Indeed, if $f \in C^\infty(M)$ has compact support within one of the elements in the family $\left\{U_{k_j}\right\}_{j\in J}$, it can be easily checked that

$$
f(\Sigma^s) \underset{\substack{ucp \\ s\to 0}}{\longrightarrow} f(\Gamma) \quad \text{and} \quad \frac{f(\Sigma^s) - f(\Gamma)}{s} \underset{\substack{ucp \\ s\to 0}}{\longrightarrow} Y[f]. \tag{2.48}
$$

If, more generally, $f \in C^{\infty}(M)$ has not compact support contained in one of the $\{U_{k_j}\}_{j\in J}$, observe that, by (2.47), we only need to consider the restriction of f to $\bigcup_{j\in J} V_{k_j}$. Take now a partition of the unity $\{\phi_k\}_{k\in\mathbb{N}}$ subordinated to the covering $\{U_k\}_{k\in\mathbb{N}}$. Since $\{\operatorname{supp}(\phi_k)\}_{k\in\mathbb{N}}$ is a locally finite family and $\bigcup_{j\in J} V_{k_j}$ is contained in a compact set because, by hypothesis, so is each V_{k_j} for any $j \in J$, then among all the $\{\phi_k\}_{k\in\mathbb{N}}$ only a finite number of them have their supports in $\{U_{k_j}\}_{j\in J}$, say $\{\phi_{k_i}\}_{i\in I}$ with $|I| < \infty$. Thus,

$$f|_{\cup_{j\in J} V_{k_j}} = \sum_{i=1}^{|I|} \phi_{k_i} f$$

and since each $\phi_{k_i} f$ is a function similar to those considered in (2.48) it is straightforward to see that those implications also hold for f. ■

The following result generalizes Proposition 2.24 to pathwise variations of a semimartingale. The proof can be found in Subsection 2.5.2

Proposition 2.33 *Let Γ be a M-valued continuous semimartingale Γ, $K \subseteq M$ a compact set, and τ_K the first exit time of Γ from K. Let Σ be a bounded pathwise variation that converges uniformly to Γ^{τ_K} and $Y : \mathbb{R}_+ \times \Omega \to TM$ the infinitesimal generator of Σ that we will also assume to be bounded. Then, for any $\alpha \in \Omega(M)$,*

$$\lim_{\substack{ucp\\ s\to 0}} \frac{1}{s}\left[\int \langle \alpha, \delta\Sigma^s\rangle - \int \langle \alpha, \delta\Gamma^{\tau_K}\rangle\right] = \int \langle i_Y \mathbf{d}\alpha, \delta\Gamma^{\tau_K}\rangle + \langle \alpha(\Gamma^{\tau_K}), Y\rangle - \langle \alpha(\Gamma^{\tau_K}), Y\rangle_{t=0}.$$

The next theorem shows that the generalization of the Critical Action Principle in Theorem 2.29 to pathwise variations fully characterizes the stochastic Hamilton's equations.

Theorem 2.34 (Second Critical Action Principle) *Let $(M, \omega = -d\theta)$ be an exact symplectic manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a semimartingale that takes values in the vector space V, and $h : M \to V^*$ a Hamiltonian function. Let m_0 be a point in M and $\Gamma : \mathbb{R}_+ \times \Omega \to M$ a continuous adapted semimartingale defined on $[0, \zeta_\Gamma)$ such that $\Gamma_0 = m_0$. Let $K \subseteq M$ be a compact set that contains m_0 and let τ_K be the first exit time of Γ from K. Suppose that $\tau_K < \infty$ a.s.. Then,*

(i) *For any bounded pathwise variation Σ with bounded infinitesimal generator Y which converges to Γ^{τ_K} uniformly, the action has a directional derivative that equals*

$$\left.\frac{d}{ds}\right|_{s=0} S(\Sigma^s) := \lim_{\substack{ucp\\ s\to 0}} \frac{1}{s}\left[S(\Sigma^s) - S(\Gamma^{\tau_K})\right] = \int \langle \mathbf{i}_Y d\theta, \delta\Gamma^{\tau_K}\rangle - \int \left\langle \widehat{Y[h]}(\Gamma^{\tau_K}), \delta X\right\rangle + \langle \theta(\Gamma^{\tau_K}), Y\rangle - \langle \theta(\Gamma^{\tau_K}), Y\rangle_{t=0}, \tag{2.49}$$

where the symbol $\widehat{Y[h]}(\Gamma^{\tau_K})$ is consistent with the notation introduced in Definition 2.21

(ii) *The semimartingale Γ satisfies the stochastic Hamiltonian equations (2.7) with initial condition $\Gamma_0 = m_0$ up to time τ_K if and only if, for any bounded pathwise variation*

$\Sigma : (-s_0, s_0) \times \mathbb{R}_+ \times \Omega \to M$ with bounded infinitesimal generator which converges uniformly to Γ^{τ_K} and such that $\Sigma_0^s = m_0$ and $\Sigma_{\tau_K}^s = \Gamma_{\tau_K}$ a.s. for any $s \in (-s_0, s_0)$,

$$\left[\left. \frac{d}{ds} \right|_{s=0} S\left(\Sigma^s\right) \right]_{\tau_K} = 0 \quad a.s..$$

Proof. We first show that the limit (2.49) exist. Let Σ be an arbitrary bounded pathwise variation converging to Γ uniformly and $Y : \mathbb{R}_+ \times \Omega \to TM$ its infinitesimal generator, that we also assume to be bounded. We have

$$\begin{aligned} \frac{1}{s}\left[S\left(\Sigma^s\right) - S\left(\Gamma^{\tau_K}\right)\right] &= \frac{1}{s}\left[\int \langle \theta, \delta\Sigma^s \rangle - \int \langle \theta, \delta\Gamma^{\tau_K} \rangle\right] \\ &\quad -\frac{1}{s}\left[\int \left\langle \widehat{h}\left(\Sigma^s\right) - \widehat{h}\left(\Gamma^{\tau_K}\right), \delta X \right\rangle\right]. \end{aligned} \tag{2.50}$$

By Proposition 2.33, the first summand in the right hand side of (2.50) converges *ucp* to

$$\int \langle \mathbf{i}_Y d\theta, \delta\Gamma^{\tau_K} \rangle + \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle - \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{t=0}.$$

as $s \to 0$. An argument similar to the one leading to Proposition 2.33 shows that the second summand converges to $\int \left\langle \widehat{Y[h]}(\Gamma^{\tau_K}), \delta X \right\rangle$. Hence,

$$\begin{aligned} \lim_{s\to 0}\frac{1}{s}\left[S\left(\Sigma^s\right) - S\left(\Gamma^{\tau_K}\right)\right] &= \int \langle \mathbf{i}_Y d\theta, \delta\Gamma^{\tau_K} \rangle - \int \left\langle \widehat{Y[h]}(\Gamma^{\tau_K}), \delta X \right\rangle \\ &\quad + \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle - \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{t=0}. \end{aligned}$$

If we denote by $\eta := -\mathbf{i}_Y d\theta = \mathbf{i}_Y \omega$ the one-form over Γ^{τ_K} built using the vector field Y over Γ^{τ_K}, the previous relation may be rewritten as

$$\begin{aligned} \left[\left. \frac{d}{ds} \right|_{s=0} S\left(\Sigma^s\right) \right] &= -\int \langle \eta, \delta\Gamma^{\tau_K} \rangle - \int \left\langle \mathbf{d}h\left(\Gamma^{\tau_K}\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X \right\rangle \\ &\quad + \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle - \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{t=0}. \end{aligned} \tag{2.51}$$

We are now going to prove the assertion in part **(ii)**. Recall that the hypothesis that Γ satisfies the stochastic Hamilton equations up to time τ_K means that

$$\left(\int \langle \beta, \delta\Gamma \rangle + \int \left\langle \left(\mathbf{d}h \cdot \omega^{\#}\left(\beta\right)\right)\left(\Gamma\right), \delta X \right\rangle\right)^{\tau_K} = 0, \tag{2.52}$$

for any $\beta \in \Omega\left(M\right)$. We now show that this expression is also true if we replace β with any process $\eta : \mathbb{R}_+ \times \Omega \to T^*M$ over Γ such that the two Stratonovich integrals involved in (2.52) are well-defined (for instance if β is a semimartingale). Indeed, invoking ([E89, 7.7]) and Whitney's embedding theorem, there exist an integer $p \in \mathbb{N}$ such that the manifold M can be seen as an embedded submanifold of $\mathbb{R}^p$. In this embedded picture, there exists a family of functions $\left\{f^1, ..., f^p\right\} \subset C^{\infty}(\mathbb{R}^p)$ such that the one-form η may be written as

$$\eta = \sum_{j=1}^{p} Z_j \mathbf{d} f^j,$$

where the $Z_j : \mathbb{R}_+ \times \Omega \to \mathbb{R}$, $j \in \{1, ..., p\}$, are real processes. Moreover, using the properties of the Stratonovich integral (see [E89, Proposition 7.4]),

$$\begin{aligned}
&\left(\int \langle \eta, \delta\Gamma \rangle + \int \left\langle \left(\mathbf{d}h \cdot \omega^{\#}(\eta) \right)(\Gamma), \delta X \right\rangle \right)^{\tau_K} \\
&\quad = \left(\sum_{j=1}^{p} \int Z_j \delta \left(\int \left\langle \langle \mathbf{d} f^j, \delta\Gamma \rangle + \int \left\langle \left(\mathbf{d}h \cdot \omega^{\#}\left(\mathbf{d} f^j\right)\right)(\Gamma), \delta X \right\rangle \right\rangle \right) \right)^{\tau_K} \\
&\quad = \sum_{j=1}^{p} \int Z_j \delta \left(\int \left\langle \langle \mathbf{d} f^j, \delta\Gamma \rangle + \int \left\langle \left(\mathbf{d}h \cdot \omega^{\#}\left(\mathbf{d} f^j\right)\right)(\Gamma), \delta X \right\rangle \right\rangle \right)^{\tau_K},
\end{aligned}$$

where the last equality follows from Proposition A.1. Therefore, since $\mathbf{d} f^j$ is a deterministic one-form we can conclude that $\left(\int \langle \langle \mathbf{d} f^j, \delta\Gamma \rangle + \int \langle \left(\mathbf{d}h \cdot \omega^{\#}\left(\mathbf{d} f^j\right)\right)(\Gamma), \delta X \rangle \rangle \right)^{\tau_K} = 0$, which justifies why (2.52) also holds if we replace $\beta \in \Omega(M)$ by an arbitrary integrable one-form η over Γ. Suppose now that Γ satisfies the stochastic Hamilton equations up to τ_K and let $\Sigma : (-s_0, s_0) \times \mathbb{R}_+ \times \Omega \to M$ be a pathwise variation like in the statement of the theorem. We want to show that

$$\left[\left. \frac{d}{ds} \right|_{s=0} S(\Sigma^s) \right]_{\tau_K} = 0 \text{ a.s..}$$

Due to (2.51), we have that

$$\begin{aligned}
\left[\left. \frac{d}{ds} \right|_{s=0} S(\Sigma^s) \right]_{\tau_K} = & -\left(\int \langle \eta, \delta\Gamma^{\tau_K} \rangle + \int \left\langle \mathbf{d}h\left(\Gamma^{\tau_K}\right)\left(\omega^{\#}(\eta)\right), \delta X \right\rangle \right)_{\tau_K} \\
& + \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{\tau_K} - \langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{t=0}.
\end{aligned}$$

Since $\Sigma_0^s = m_0$ and $\Sigma_{\tau_K}^s = \Gamma_{\tau_K}$ a.s. for any $s \in (-s_0, s_0)$, then $Y_0 = Y_{\tau_K} = 0$ a.s. and both $\langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{\tau_K}$ and $\langle \theta\left(\Gamma^{\tau_K}\right), Y \rangle_{t=0}$ vanish. Moreover,

$$\begin{aligned}
&\left(\int \langle \eta, \delta\Gamma^{\tau_K} \rangle + \int \left\langle \mathbf{d}h\left(\Gamma^{\tau_K}\right)\left(\omega^{\#}(\eta)\right), \delta X \right\rangle \right)_{\tau_K} \\
&= \left(\left(\int \langle \eta, \delta\Gamma^{\tau_K} \rangle + \int \left\langle \mathbf{d}h\left(\Gamma^{\tau_K}\right)\left(\omega^{\#}(\eta)\right), \delta X \right\rangle \right)^{\tau_K} \right)_{\tau_K} \\
&= \left(\left(\int \langle \eta, \delta\Gamma \rangle + \int \left\langle \mathbf{d}h(\Gamma)\left(\omega^{\#}(\eta)\right), \delta X \right\rangle \right)^{\tau_K} \right)_{\tau_K}
\end{aligned} \tag{2.53}$$

which is zero because of (2.52). In the last equality we have used Proposition A.1. Conversely, suppose that $\left[\left. \frac{d}{ds} \right|_{s=0} S(\Sigma^s) \right]_{\tau_K} = 0$ a.s. for arbitrary bounded pathwise variations tending to

Γ^{τ_K} uniformly, like in the statement. We want to show that (2.52) holds. Since our pathwise variations satisfy that $Y_0 = Y_{\tau_K} = 0$ a.s., we obtain that

$$\left[\left.\frac{d}{ds}\right|_{s=0} S\left(\Sigma^s\right)\right]_{\tau_K} = -\left(\int \langle \eta, \delta\Gamma^{\tau_K}\rangle + \int \left\langle \mathbf{d}h\left(\Gamma^{\tau_K}\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X\right\rangle\right)_{\tau_K} = 0 \tag{2.54}$$

where η is an arbitrary bounded one form over Γ. Suppose now that η is a semimartingale. Then $\mathbf{1}_{[0,t]}\eta : \mathbb{R}_+ \times \Omega \to T^*M$ is again bounded and expressions

$$\int \left\langle \mathbf{1}_{[0,t]}\eta, \delta\Gamma^{\tau_K}\right\rangle \quad \text{and} \quad \int \left\langle \mathbf{d}h\left(\Gamma^{\tau_K}\right)\left(\omega^{\#}\left(\mathbf{1}_{[0,t]}\eta\right)\right), \delta X\right\rangle$$

are well-defined by Proposition A.5 because both Γ^{τ_K} and X are continuos semimartingales. We already saw in (2.53) that (2.54) is equivalent to

$$\left(\int \langle \eta, \delta\Gamma\rangle + \int \left\langle \mathbf{d}h\left(\Gamma\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X\right\rangle\right)_{\tau_K} = 0.$$

Replacing η by $\mathbf{1}_{[0,t]}\eta$ in (2.54) and using again the Proposition A.5, we write

$$\begin{aligned}
0 &= \left(\int \left\langle \mathbf{1}_{[0,t]}\eta, \delta\Gamma\right\rangle + \int \left\langle \mathbf{d}h\left(\Gamma\right)\left(\omega^{\#}\left(\mathbf{1}_{[0,t]}\eta\right)\right), \delta X\right\rangle\right)_{\tau_K} \\
&= \left(\left(\int \langle \eta, \delta\Gamma\rangle + \int \left\langle \mathbf{d}h\left(\Gamma\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X\right\rangle\right)^t\right)_{\tau_K} \\
&= \left(\int \langle \eta, \delta\Gamma\rangle + \int \left\langle \mathbf{d}h\left(\Gamma\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X\right\rangle\right)_{t\wedge\tau_K} \\
&= \left(\left(\int \langle \eta, \delta\Gamma\rangle + \int \left\langle \mathbf{d}h\left(\Gamma\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X\right\rangle\right)^{\tau_K}\right)_t .
\end{aligned}$$

Since t is arbitrary this implies that the process $\left(\int \langle \eta, \delta\Gamma\rangle + \int \left\langle \mathbf{d}h\left(\Gamma\right)\left(\omega^{\#}\left(\eta\right)\right), \delta X\right\rangle\right)^{\tau_K}$ is identically zero, as required. ■

2.4 Stochastic Hamilton-Jacobi equation

Hamilton-Jacobi theory is an important part of classical mechanics that provides a characterization of the generating functions of certain time-dependent canonical transformations that put a given Hamiltonian system in such a form that its solutions are extremely easy to find; this is the so called solution by *reduction to the equilibrium*. In this respect, the fact that the classical action satisfies the Hamilton-Jacobi equation is a very relevant result. Hamilton-Jacobi theory also plays a fundamental role in the study of the quantum-classical relationship, in integrable systems, or in the development of structure preserving numerical integrators. For all these reasons it is desirable to have at hand similar tools in the stochastic Hamiltonian context; this is the main goal of this work. The Hamilton-Jacobi equation was already studied by Bismut [B81] in the context of Hamiltonian diffusions and, as we will see, most of the ideas in that piece of work are still valid at our degree of generality; at some level, this section can be seen as a completion of Bismut's work in which complete proofs are provided and where the results have been adapted to our framework using a more modern geometric language; this makes them more palatable to a growing community interested both in geometric mechanics and stochastics.

2.4.1 The stochastic action on Lagrangian submanifolds and the Hamilton-Jacobi equation

It is a classical result in mechanics that the action, when written as a function of the configuration space and time, satisfies the Hamilton-Jacobi equation (see for instance [A89]). The main goal of this subsection is showing that an analogous result holds for the stochastic action.

To start with, we now state some of the basic properties of the flow defined by (2.7). Let $\varphi\left(\cdot,z\right):[0,\zeta(z))\subseteq\mathbb{R}_{+}\times\Omega\rightarrow M$ denote the unique solution of (2.5) with initial condition $\Gamma_0=z\in M$ a.s.. The map φ will be referred to as the ***stochastic flow*** associated to (2.5). For any $(t,\eta)\in\mathbb{R}_{+}\times\Omega$, let $\mathbb{D}_t\left(\eta\right)=\{z\in M\mid\zeta\left(z,\eta\right)>t\}$. Observe that $\mathbb{D}_t\left(\eta\right)\subseteq\mathbb{D}_s\left(\eta\right)$ if $s\leq t$. By [K90, Lemma 4.8.3] $\mathbb{D}_t\left(\eta\right)$ is an open set for any $t\in\mathbb{R}_{+}$ a.s. and

$$\begin{array}{rcl}\varphi_t\left(\eta\right):\mathbb{D}_t\left(\eta\right) & \longrightarrow & M\\ z & \longmapsto & \varphi_t\left(z,\eta\right)\end{array}$$

is a continuously differentiable diffeomorphism ([K90, Theorem 4.8.4]). Additionally,

$$\begin{array}{rcl}\varphi\left(\eta\right):[0,t]\times\mathbb{D}_t\left(\eta\right) & \longrightarrow & M\\ (s,z) & \longmapsto & \varphi_s\left(z,\eta\right)\end{array}$$

is continuous and its partial derivatives with respect to $z\in\mathbb{D}_t\left(\eta\right)$ are also continuous on $[0,t]\times\mathbb{D}_t\left(\eta\right)$. The local version of these results, that is, the case $M=\mathbb{R}^{2n}$, can be also found in [P05, Chapter V Theorem 39]. Furthermore, the stochastic flow φ acts naturally on tensor fields and in particular on differential forms. Hence, by [K81, Theorem 3.3] and [K90, Section 4.9], if $\alpha\in\Omega^k\left(M\right)$ is a k-form, $k\in\mathbb{N}$, then

$$\varphi_t\left(\eta\right)^*\alpha=\alpha+\sum_{i=1}^{r}\left(\int_0^t\varphi_s^*\left(\pounds_{X_{h_i}}\alpha\right)\delta X_s^i\right)\left(\eta\right) \tag{2.55}$$

on $\mathbb{D}_t\left(\eta\right)$, $(t,\eta)\in\mathbb{R}_{+}\times\Omega$. In particular, if $\alpha=\omega$ is the symplectic form, then $\pounds_{X_{h_i}}\omega=0$ for any $i=1,...,r$ and $\varphi^*\omega=\omega$ which is the stochastic version of the Liouville's Theorem (Theorem 2.11).

Let $\varphi_t(\eta):\mathbb{D}_t\left(\eta\right)\rightarrow M$ be the flow associated to the stochastic Hamilton equations (2.5), $(t,\eta)\in\mathbb{R}_{+}\times\Omega$. We define the function $R_t\left(\eta\right):\mathbb{D}_t\left(\eta\right)\rightarrow\mathbb{R}$ as

$$R_t\left(\eta,z\right):=S\left(\varphi\left(z\right)\right)_t\left(\eta\right).$$

The next proposition provides the differential of $R_t\left(\eta\right)$.

Proposition 2.35 *Let $t \in \mathbb{R}_+$ be a fixed time instant and $\eta \in \Omega$. Then $R_t(\eta) : \mathbb{D}_t(\eta) \to \mathbb{R}$ is differentiable and*

$$\mathbf{d}R_t(\eta) = \varphi_t(\eta)^* \theta - \theta, \tag{2.56}$$

where θ is the one form of the exact symplectic manifold $(M, \omega = -\mathbf{d}\theta)$.

Proof. We will proceed by showing that for any pair of points $x, y \in \mathbb{D}_t(\eta)$ we can write

$$R_t(\eta, x) - R_t(\eta, y) = \int_\gamma \left(\varphi_t(\eta)^*(\theta) - \theta\right),$$

where $\gamma : (a,b) \subseteq \mathbb{R} \to \mathbb{D}_t(\eta)$ is any smooth curve in $\mathbb{D}_t(\eta)$ that links x and y. This expression immediately implies that R_t has continuous directional derivatives and it is hence Fréchet differentiable. Indeed, using first (2.55), we have

$$\begin{aligned} \int_\gamma \left(\varphi_t(\eta)^*(\theta) - \theta\right) &= \int_\gamma \left(\sum_{i=1}^r \int_0^t \varphi_s^* \left(\pounds_{X_{h_i}} \theta\right) \delta X_s^i\right)(\eta) \\ &= \left(\sum_{i=1}^r \int_0^t \left(\int_\gamma \varphi_s^* \left(\pounds_{X_{h_i}} \theta\right)\right) \delta X_s^i\right)(\eta), \end{aligned} \tag{2.57}$$

where in the second equality we used Fubini's Theorem. Now, since $\mathbf{i}_{X_{h_i}} \omega = \mathbf{d}h_i$, for any $i = 1, \ldots, r$, (2.57) equals

$$\begin{aligned} &\sum_{i=1}^r \left(\int_0^t \left(\int_\gamma \varphi_s^* \mathbf{d}\left(\mathbf{i}_{X_{h_i}} \theta\right)\right) \delta X_s^i - \int_0^t \left(\int_\gamma \varphi_s^* \mathbf{d}h_i\right) \delta X_s^i\right)(\eta) \\ &= \sum_{i=1}^r \left(\int_0^t \left(\int_\gamma \mathbf{d}\left(\varphi_s^*(\mathbf{i}_{X_{h_i}} \theta)\right)\right) \delta X_s^i - \int_0^t \left(\int_\gamma \mathbf{d}(\varphi_s^* h_i)\right) \delta X_s^i\right)(\eta) \\ &= \sum_{i=1}^r \left(\int_0^t \left[\mathbf{i}_{X_{h_i}} \theta\left(\varphi_s(\gamma_b)\right) - \mathbf{i}_{X_{h_i}} \theta\left(\varphi_s(\gamma_a)\right)\right] \delta X_s^i - \int_0^t \left[h_i\left(\varphi_s(\gamma_b)\right) - h_i\left(\varphi_s(\gamma_a)\right)\right] \delta X_s^i\right)(\eta) \\ &= \left(\int_0^t \langle \theta, \delta\varphi_s(\gamma_b)\rangle - \int_0^t \left\langle \hat{h}\left(\varphi_s(\gamma_b)\right), \delta X_s \right\rangle\right)(\eta) \\ &\quad - \left(\int_0^t \langle \theta, \delta\varphi_s(\gamma_a)\rangle - \int_0^t \left\langle \hat{h}\left(\varphi_s(\gamma_a)\right), \delta X_s \right\rangle\right)(\eta) = R_t(\eta, x) - R_t(\eta, y). \end{aligned}$$

Given that $\gamma : (a,b) \to \mathbb{D}_t(\eta)$ and the points $x, y \in \mathbb{D}_t(\eta)$ are arbitrary, the result follows. ■

Later on in this section we will need the composition of R with the inverse of the stochastic flow φ. More specifically, let $(t, \eta) \in \mathbb{R}_+ \times \Omega$ and let $\varphi_t^{-1}(\eta) : \varphi_t(\eta)(\mathbb{D}_t(\eta)) \to \mathbb{D}_t(\eta)$ the inverse of $\varphi_t(\eta)$. We define $\hat{R}_t(\eta) : \varphi_t(\eta)(\mathbb{D}_t(\eta)) \to \mathbb{D}_t(\eta)$ as $\hat{R}_t(\eta) := R_t(\eta) \circ \varphi_t^{-1}(\eta) = \varphi_t^{-1}(\eta)^*(R_t(\eta))$. Consequently,

$$\mathbf{d}\hat{R}_t(\eta) = \varphi_t^{-1}(\eta)^*(\mathbf{d}R_t(\eta)) = \varphi_t^{-1}(\eta)^*\left(\varphi_t(\eta)^*(\theta) - \theta\right) = \theta - \varphi_t^{-1}(\eta)^*(\theta). \tag{2.58}$$

In order to get closer to the classical deterministic result on the Hamilton-Jacobi equation we are first going to visualize it, using the map R, as a process depending on M through the initial condition of the flow φ generated by (2.5). Second, we will restrict R to a Lagrangian submanifold of M; this encodes mathematically the writing of the action as a function of the configuration space. Recall that a submanifold $\iota : L \hookrightarrow M$ of a symplectic manifold (M, ω) is called ***Lagrangian*** if $\dim(L) = \dim(M)/2$ and $\iota^*\omega = 0$. Observe that since $\varphi_t(\eta)$ is a symplectomorphism a.s. for any $t \in \mathbb{R}_+$ and $\mathbb{D}_t(\eta)$ is an open set, if L is a Lagrangian submanifold so are $L \cap \mathbb{D}_t(\eta)$ and $\varphi_t(\eta)(L \cap \mathbb{D}_t(\eta))$.

From now on we are going to assume that the underlying symplectic manifold (M, ω) is actually a cotangent bundle endowed with its canonical symplectic structure. More specifically, $M = T^*Q$ for some manifold Q. In this case, a point $y \in L \subset T^*Q$ in a Lagrangian submanifold L is said to be a ***regular point*** of L, if the restriction $\pi|_L : L \to Q$ of the canonical projection $\pi : T^*Q \to Q$ to L is a local diffeomorphism at y (that is, $T_y\, \pi|_L : T_yL \to T_{\pi(y)}Q$ is an isomorphism). In a neighborhood $U \subset L$ of a regular point $y \in L$ we can obviously describe the Lagrangian submanifold L using local coordinates on the base manifold Q, which we will generally denote by $(q^1, \ldots, q^n)$. On the other hand, since $\iota^*\omega = \mathbf{d}(\iota^*\theta) = 0$, there exists by the Poincaré lemma (shrinking U if necessary) a smooth function $f \in C^\infty(U)$ such that $\iota^*\theta = \mathbf{d}f$. Conversely, if $\left(q^1, \ldots, q^n, p_1, \ldots, p_n\right)$ are local Darboux coordinates in a neighborhood $V \subseteq T^*Q$ and $f \in C^\infty(\pi(V))$ is a function with no critical points, then the set

$$L_f = \left\{ (q,p) \in V \mid p_i = \frac{\partial f}{\partial q^i},\ i = 1, ..., n \right\} \tag{2.59}$$

is a local Lagrangian submanifold such that

$$\iota_f^*\theta = \pi|_{L_f}^*\mathbf{d}f, \tag{2.60}$$

with $\iota_f : L_f \hookrightarrow V$ the inclusion and $\pi|_{L_f} : L_f \subset T^*Q \to \pi(V)$ the local diffeomorphism obtained by restriction of the canonical projection.

Theorem 2.36 *Let Q be a manifold and let $L \subset T^*Q$ a Lagrangian submanifold. Let $y_0 \in L$ be a regular point and let $x_0 = \pi(y_0)$, where $\pi : T^*Q \to Q$ is the canonical projection. Then, there exist two neighborhoods $V_{y_0} \subseteq L$ and $V_{x_0} \subseteq Q$ of y_0 and of x_0, respectively and a map $\xi : \Omega \times V_{x_0} \to \mathbb{R}_+$ with the property that $\xi(x) : \Omega \to \mathbb{R}_+$ is a stopping time, such that the equation*

$$\pi(\varphi_s(\eta, y)) = x \tag{2.61}$$

has a unique solution in $V_{y_0} \subseteq L$ for any $\eta \in \Omega$, any $x \in V_{x_0}$, and any $s \in [0, \xi(\eta, x)]$. We are going to denote this solution by $\psi_s(\eta, x)$. Moreover, $\psi(x) : [0, \xi(x)) \to V_{y_0}(\eta)$ is a semimartingale for any $x \in V_{x_0}$ and $\psi_s(\eta) : V_{x_0} \to V_{y_0}$ is a diffeomorphism for any $s \in [0, \xi(x))$ which depends continuously on s.

Proof. Let $U_{y_0} \subset L$ be an open neighborhood of $y_0 \in L$. We pick U_{y_0} small enough so that $\pi|_{U_{y_0}}$ is a diffeomorphism onto its image and a set of local coordinates $\left(q^i; i = 1, ..., n\right)$ can be chosen on $U_{x_0} := \pi(U_{y_0})$. Let $\left(y^i = q^i \circ \pi|_L; i = 1, ..., n\right)$ be the corresponding induced coordinates on U_{y_0}. Denote by $\hat{q} : U_{x_0} \to \mathbb{R}^n$ and $\hat{y} : U_{y_0} \to \mathbb{R}^n$ the local chart maps associated to these coordinates. For any $y \in U_{y_0}$, let $\tau_{U_{x_0}}(y, \eta) = \inf\{t > 0 \mid \pi \circ \varphi_t(\eta, y) \notin U_{x_0}\}$ be the first exit time at which the semimartingale $\pi \circ \varphi(y)$ leaves U_{x_0}. Let F be the restriction of $\pi \circ \varphi$ to the set $A := \{(s, \eta, y) \in \mathbb{R}_+ \times \Omega \times U_{y_0} \mid s \in [0, \tau_{U_{x_0}}(y, \eta))\}$. In local coordinates, $F : A \to U_{x_0}$ is expressed as

$$F_s^j(\eta)\left(y^1, ..., y^n\right) = q^j \circ \pi \circ \varphi_s(\eta) \circ \hat{y}^{-1}\left(y^1, ..., y^n\right), \ \ j = 1, ..., n.$$

Now, remark that $\det\left(\frac{\partial F_0^j(\eta)}{\partial y^i}(y_0)\right) \neq 0$ a.s. because $y_0 \in L$ is a regular point. The continuity of the derivative of $F_0(\eta) : U_{y_0} \to U_{x_0}$ implies that there exists a neighborhood $V_{y_0} \subseteq U_{y_0}$ such that $\det\left(\frac{\partial F_0^j(\eta)}{\partial y^i}(y)\right) > 0$ a.s., for any $y \in V_{y_0}$. For any of these $y \in V_{y_0}$, let

$$\begin{array}{rcl} Z(y) := \det\left(\frac{\partial F^j}{\partial y^i}(y)\right) : \left[0, \tau_{U_{x_0}}(y)\right) & \longrightarrow & \mathbb{R} \\ (s, \eta) & \longmapsto & \det\left(\frac{\partial F_s^j(\eta)}{\partial y^i}(y)\right), \end{array}$$

which is a well defined and continuous semimartingale, by the continuity of the differential of the flow φ. Observe that $Z_0(y) > 0$ for any $y \in V_{y_0}$. Let $T(y, \eta) := \inf\{\tau_{U_{x_0}}(y) \geq t > 0 \mid Z_t(y, \eta) \notin \mathbb{R}_+\}$.

Now, recall that we want to see that the equation $\pi(\varphi_s(\eta, y)) = x$ has a unique solution in $y \in L$, for any $x \in V_{x_0}$ in a suitable V_{x_0} and up to a suitable stopping time $\xi(x)$. Therefore, it suffices to solve the equation

$$\pi\left(\varphi_s^{T(y)}(\eta, y)\right) = x, \tag{2.62}$$

where $\varphi^{T(y)}(y)$ denotes the process $\varphi(y)$ stopped at time $T(y)$, that is, $\varphi^{T(y)}(y)(s, \eta) = \varphi_{T(y,\eta)\wedge s}(\eta, y)$. Observe that $\varphi^{T(y)}(y)$ is always in U_{y_0} if y was already in V_{y_0}. Consequently, $\varphi^{T(y)}(y)$ may be described using the local coordinates introduced above. Moreover, if we set $\xi(x) := T\left(\pi|_L^{-1}(x)\right)$, $V_{x_0} := \pi(V_{y_0})$, the equation (2.62) admits by construction a unique solution $\psi_s(\eta, x)$ via the Implicit Function Theorem. Additionally, if we apply the Stratonovich differentiation rules to

$$\pi\left(\varphi_s^{T(y)}(\eta, \psi_s(\eta, x))\right) = x, \ \ s \in [0, \xi(x, \eta))$$

we obtain that $\psi_s(\eta, x)$ satisfies up to time $\xi(x)$ the Stratonovich differential equation

$$\delta\psi_s(x) = \sum_{i=1}^{r} \left[T_{\psi_s(x))}F\right]^{-1}\left(T_{\varphi_s^{T(y)}(\psi_s(x))}(\hat{q} \circ \pi)\left(X_{h_i}(\varphi_s^{T(y)}(\psi_s(x)))\right)\right)\delta X_s^i \tag{2.63}$$

with initial condition $\psi_{s=0}(x) = y(x) \in V_{y_0}$ a.s. such that $\pi(y(x)) = x \in V_{x_0}$. That is, we can visualize $\psi_s(\eta, x)$ as the unique stochastic flow associated to the stochastic differential equation (2.63). This guarantees that the properties claimed in the statement hold. ■

We proceed now by considering the stochastic action R not as a semimartingale parametrized by T^*Q through the initial condition of the stochastic flow φ defined by (2.5), but as a process depending on the base manifold Q. More specifically, we will restrict to the open neighborhood $V_{x_0} \subset Q$ introduced in the statement of Theorem 2.36 and which is mapped onto $V_{y_0} \subset L$ using the map ψ that solves (2.61). Furthermore, since we are always going to work around regular points of the Lagrangian submanifold, we will always consider Lagrangian submanifolds of the type L_f (see (2.59)) for some $f \in C^\infty(Q)$.

Definition 2.37 *Let $L_f \subseteq T^*Q$ be a Lagrangian submanifold, $f \in C^\infty(Q)$. Let $V_{x_0} \subseteq Q$ be the open neighborhood of x_0 introduced in Theorem 2.36 and $\psi(x) : [0, \xi(x)) \to V_{y_0}$ the semimartingale solution of (2.63) with initial condition $x \in V_{x_0}$ a.s.. We define the **projected stochastic action** $\widetilde{S}(x) : [0, \xi(x)) \to \mathbb{R}$ as*

$$\widetilde{S}_t(\eta, x) = R_t(\eta, \psi_t(\eta, x)) + f(\pi(\psi_t(\eta, x))) = (R_t(\eta) + f \circ \pi) \circ \psi_t(\eta, x).$$

Notice that the differentiability properties of the maps R, $f \in C^\infty(Q)$, and ψ imply that the map

$$\begin{array}{rcl} \widetilde{S}_t(\eta) : \mathbb{D}_t^\psi(\eta) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \widetilde{S}_t(\eta, x) \end{array} \tag{2.64}$$

is continuously differentiable for any $(t, \eta) \in \mathbb{R}_+ \times \Omega$ such that $t \in [0, \xi(x, \eta))$. In this expression $\mathbb{D}_t^\psi(\eta) := \{x \in V_{x_0} \mid t < \xi(x, \eta)\}$. The following theorem provides an explicit expression for the spatial derivatives of the projected stochastic action $\widetilde{S}$.

Theorem 2.38 *Let L_f be a Lagrangian submanifold of T^*Q, $f \in C^\infty(Q)$. Then, on the open set $\mathbb{D}_t^\psi(\eta)$, $(t, \eta) \in \mathbb{R}_+ \times \Omega$,*

$$\mathbf{d}\widetilde{S}_t(\eta) = (\varphi_t(\eta) \circ \psi_t(\eta))^* \theta. \tag{2.65}$$

If $(q^i, p_i; i = 1, ..., n)$ are local Darboux coordinates of T^*Q on an open neighborhood of a regular point $y_0 \in L_f$, the expression (2.65) can be locally written as

$$\frac{\partial \widetilde{S}_t(\eta)}{\partial q^i}(q) = p_i(\varphi_t(\eta, \psi_t(\eta, q))), \quad i = 1, ..., n.$$

Proof. First of all observe that $\widetilde{S}_t(\eta)$ can be expressed in terms of $\hat{R}_t(\eta)$ as follows:

$$\widetilde{S}_t(\eta, q) = \hat{R}_t(\eta) \circ \varphi_t(\eta) \circ \psi_t(\eta, q) + f \circ \pi \circ \psi_t(\eta, q).$$

Then, for any smooth curve $\gamma : [a, b] \to \mathbb{D}_t^\psi(\eta)$

$$\widetilde{S}_t(\eta, \gamma_b) - \widetilde{S}_t(\eta, \gamma_a) = \int_\gamma \mathbf{d}\widetilde{S}_t(\eta) = \int_\gamma \mathbf{d}\left[\hat{R}_t(\eta) \circ \varphi_t(\eta) \circ \psi_t(\eta)\right] + \int_\gamma \mathbf{d}(f \circ \pi \circ \psi_t(\eta)). \tag{2.66}$$

Given that $\mathbb{D}_t^\psi(\eta) \subset L_f$, the curve γ takes values in the Lagrangian submanifold L_f and hence (2.66) can be rewritten as

$$\begin{aligned} \widetilde{S}_t(\eta, \gamma_b) - \widetilde{S}_t(\eta, \gamma_a) &= \int_\gamma \iota_{L_f}^* \mathbf{d}\left[\hat{R}_t(\eta) \circ \varphi_t(\eta) \circ \psi_t(\eta)\right] + \int_\gamma \mathbf{d}(f \circ \pi \circ \psi_t(\eta)) \\ &= \int_\gamma \mathbf{d}\left[\hat{R}_t(\eta) \circ \varphi_t(\eta) \circ \psi_t(\eta) \circ \iota_{L_f}\right] + \int_\gamma \mathbf{d}(f \circ \pi \circ \psi_t(\eta)). \end{aligned} \tag{2.67}$$

On the other hand, we saw in (2.58) that

$$\mathbf{d}\hat{R}_t = \theta - \varphi_t^{-1}(\eta)^*(\theta).$$

Moreover, since $\iota_f^*\theta = \pi|_{L_f}^* \mathbf{d}f$ we have that

$$\mathbf{d}\left[\hat{R}_t(\eta) \circ \varphi_t(\eta) \circ \psi_t(\eta) \circ \iota_{L_f}\right] = \left(\varphi_t(\eta) \circ \psi_t(\eta) \circ \iota_{L_f}\right)^* \theta - \mathbf{d}\left(f \circ \pi \circ \psi_t(\eta)\right),$$

which substituted in (2.67) yields

$$\widetilde{S}_t(\eta, \gamma_b) - \widetilde{S}_t(\eta, \gamma_a) = \int_\gamma \left(\varphi_t(\eta) \circ \psi_t(\eta) \circ \iota_{L_f}\right)^* \theta = \int_\gamma \left(\varphi_t(\eta) \circ \psi_t(\eta)\right)^* \theta.$$

Since γ is an arbitrary smooth curve, we can conclude that

$$\mathbf{d}\widetilde{S}_t(\eta) = \left(\varphi_t(\eta) \circ \psi_t(\eta)\right)^* \theta,$$

as required. ■

We conclude this section by proving that the projected stochastic action $\widetilde{S}_t$ satisfies a specific stochastic differential equation which generalizes the classical Hamilton-Jacobi equation. For obvious reasons, this equation will be referred to as the ***stochastic Hamilton-Jacobi equation***.

Theorem 2.39 (Stochastic Hamilton-Jacobi equation) *Using the same notation as in Theorem 2.36, the projected stochastic action $\widetilde{S}(q) : [0, \xi(q)) \to \mathbb{R}$ associated to the Lagrangian submanifold L_f defined by the function $f \in C^\infty(Q)$ satisfies*

$$\widetilde{S}(q) = f(q) - \int \left\langle \hat{h}\left(q, \frac{\partial \widetilde{S}_s}{\partial q}(q)\right), \delta X_s \right\rangle$$

for any $q \in V_{x_0}$.

In order to prove this theorem we need the following auxiliary result.

Proposition 2.40 ([K90, Theorem 3.3.2]) *Let $F(x) : \mathbb{R}_+ \times \Omega \to \mathbb{R}$, $x \in \mathbb{R}^n$, be a family of continuous semimartingales parametrized by $\mathbb{R}^n$. Suppose that the dependence of this family on the $\mathbb{R}^n$parameter is at least three times differentiable. In addition, suppose that there exists a process $f : \mathbb{R}_+ \times \Omega \times \mathbb{R}^n \to \mathbb{R}^d$ that satisfies sufficient regularity conditions and a semimartingale $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ such that*

$$F(x) = \sum_{j=1}^{r} \int f_j(t, x)\, \delta X_t^j.$$

Let $g : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n$ be a continuous $\mathbb{R}^n$-valued semimartingale. Then $F(g) : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ defined as $F(g)(t, \eta) := F(g_t(\eta), t, \eta)$ satisfies

$$F(g_t, t) - F(g_0, 0) = \sum_{j=1}^{r} \int_0^t f_j(s, g_s)\, \delta X_s^j + \sum_{i=1}^{n} \int_0^t \frac{\partial F}{\partial x^i}(s, g_s)\, \delta g_s^i.$$

Proof of Theorem 2.39. First of all observe that using the definition of the function R_t the semimartingale $\widetilde{S}(q) : [0, \xi(q)) \to \mathbb{R}$ may be expressed as

$$\begin{aligned}\widetilde{S}(q) &= f \circ \pi \circ \psi_t(\eta, q) + R_t(\eta, \psi_t(\eta, q)) \\ &= f \circ \pi \circ \psi_t(\eta, q) + \sum_{j=1}^{r} \left(\int \left(\mathbf{i}_{X_{h_j}} \theta - h_j \right) (\varphi_s(z))\, \delta X_s^j \right) \Bigg|_{z=\psi_t(\eta,q)} .\end{aligned}$$

If we use Proposition 2.40 in the second summand of this expression, we obtain

$$\widetilde{S}(q) = f \circ \pi \circ \psi(q) + \sum_{j=1}^{r} \left(\int \left(\mathbf{i}_{X_{h_j}} \theta - h_j \right) (\varphi_s(\psi_s(q)))\, \delta X_s^j \right) + \int \langle \mathbf{d} R_s, \delta \psi_s(q) \rangle . \tag{2.68}$$

We now separately study the summands in the right hand side of this equation in order to prove the statement of the theorem. We start by recalling that by Proposition 2.35, $\mathbf{d}R_s = \varphi_s^* \theta - \theta$ and hence

$$\int \langle \mathbf{d} R_s, \delta \psi_s(q) \rangle = \int \langle \varphi_s^* \theta - \theta, \delta \psi_s(q) \rangle . \tag{2.69}$$

Furthermore, since $\iota_f^* \theta = \pi|_{L_f}^* \mathbf{d} f$ and the semimartingale $\psi(q)$ takes values in $V_{y_0} \subseteq L_f$,

$$\int_0^t \langle \theta, \delta \psi_s(q) \rangle = \int_0^t \langle \mathbf{d}(f \circ \pi), \delta \psi_s(q) \rangle = f \circ \pi \circ \psi_t(q) - f(q) . \tag{2.70}$$

We now recall that the semimartingale $\varphi(\psi(q)) : [0, \xi(q)) \to T^*Q$ takes values in the fiber $\pi^{-1}(q)$. Indeed, by the construction in Theorem 2.36, $\psi(q)$ is the semimartingale starting at q such that

$$\pi(\varphi_s(\eta, \psi_s(\eta, q))) = q$$

for any $(s, \eta) \in [0, \xi(q))$. Then, since θ is a semibasic form we necessarily have that

$$\int \langle \theta, \delta(\varphi_s(\psi_s(q))) \rangle = 0.$$

But, using the fact that φ is the flow of the stochastic Hamilton equations (2.5), by Proposition 2.40, we have that for any $g \in C^\infty(M)$

$$g(\varphi(\psi(q))) = g(y(q)) + \sum_{j=1}^{r} \int X_{h_j}[g](\varphi_s(\psi_s(q)))\delta X_s^j + \int \langle \mathbf{d}(g \circ \varphi_s), \delta \psi_s(q) \rangle \tag{2.71}$$

where $y(q) \in L_f$ is the unique point such that $\pi|_L(y(q)) = q$. We claim that

$$0 = \int \langle \theta, \delta(\varphi_s(\psi_s(q))) \rangle = \sum_{j=1}^{r} \int \left(\mathbf{i}_{X_{h_j}} \theta \right) (\varphi_s(\psi_s(q)))\, \delta X_s^j + \int \langle \varphi_s^* \theta, \delta \psi_s(q) \rangle . \tag{2.72}$$

Indeed, since we are working at a local level we can use Darboux coordinates and we can replace θ by $\sum_{i=1}^{n} p_i \mathbf{d}q^i$; (2.72) is a straightforward consequence of (2.71). If we now plug (2.69), (2.70), and (2.72) into (2.68) we obtain

$$\widetilde{S}\left(q\right)=f\left(q\right)-\sum_{j=1}^{r}\int h_j\left(\varphi_s\left(\psi_s(q)\right)\right)\delta X_s^j. \tag{2.73}$$

Finally, we saw in Theorem 2.38 that

$$p_i\left(\varphi_t\circ\psi_t\left(\eta,q\right)\right)=\frac{\partial\widetilde{S}_t\left(\eta\right)}{\partial q^i}\left(q\right),\quad i=1,...,n,$$

on $\mathbb{D}_t^{\psi}\left(\eta\right)=\left\{x\in V_{x_0}\mid\xi\left(x,\eta\right)>t\right\}$, $(t,\eta)\in\mathbb{R}_+\times\Omega$. For any $\eta\in\Omega$, the time parameter s in the integrand of (2.73) is always smaller than $\xi\left(q,\eta\right)$ and hence as $\frac{\partial\widetilde{S}_s}{\partial q^i}\left(q\right)$ and $p_i\left(\varphi_s\circ\psi_s\left(q\right)\right)$ coincide a.s. on $[0,\xi(q))$ for any $i=1,...,n$, the result follows. ■

Example 2.41 Let $Q=\mathbb{R}^n$ and $T^*Q=\mathbb{R}^n\times\mathbb{R}^n$ with global coordinates $\left(q^i,p_i;i=1,...,n\right)$. Let $f\in C^{\infty}\left(\mathbb{R}^n\right)$, $h_0\in C^{\infty}\left(\mathbb{R}^{2n}\right)$, and $h_i=p_i$ for any $i=1,...,n$. Consider the semimartingale $X:\mathbb{R}_+\times\Omega\to\mathbb{R}^{n+1}$ given by $(t,\omega)\mapsto(t,B_t^1,\ldots,B_t^n)$, where $(B^1,\ldots,B^n)$ is a n-dimensional Brownian motion with diffusion coeficient $\nu=2$. That is, $[B_t^i,B_t^j]=\delta^{ij}2t$, where $[\cdot,\cdot]$ denotes the quadratic variation. Then, the projected stochastic action $\widetilde{S}:\mathbb{R}_+\times\Omega\times\mathbb{R}^n\to\mathbb{R}$ built from the stochastic Hamiltonian system on $\mathbb{R}^{2n}$ with Hamiltonian fuction $h=(h_0,h_1,\ldots,h_n)$ and stochastic component X satisfies, by Theorem 2.39

$$\widetilde{S}_t\left(q\right)=f(q)-\int_0^t h_0\left(q,\frac{\partial\widetilde{S}_s}{\partial q}(q)\right)ds-\sum_{i=1}^{n}\int\frac{\partial\widetilde{S}_s}{\partial q^i}(q)\delta B_s^i. \tag{2.74}$$

If we transform the Stratonovich integrals in this expression into Itô integrals, (2.74) reads

$$\widetilde{S}_t\left(q\right)=f(q)-\int_0^t h_0\left(q,\frac{\partial\widetilde{S}_s}{\partial q}(q)\right)ds-\sum_{i=1}^{n}\int_0^t\frac{\partial\widetilde{S}_s}{\partial q^i}(q)dB_s^i-\frac{1}{2}\sum_{i=1}^{n}\left[\frac{\partial\widetilde{S}}{\partial q^i}(q),B^i\right]_t,$$

A lengthy but straightforward computation shows that

$$\frac{\partial\widetilde{S}_t}{\partial q^i}\left(q\right)=\frac{\partial f}{\partial q^i}(q)-\int_0^t\frac{\partial}{\partial q^i}\left(h_0\left(q,\frac{\partial\widetilde{S}_s}{\partial q}(q)\right)\right)ds-\sum_{r=1}^{n}\int_0^t\frac{\partial}{\partial q^i}\left(h_r\left(q,\frac{\partial\widetilde{S}_s}{\partial q}(q)\right)\right)\delta B_s^r.$$

Since $h_r=p_r$ for any $r=1,...,n$,

$$\frac{\partial\widetilde{S}_t}{\partial q^i}\left(q\right)=\frac{\partial f}{\partial q^i}(q)-\int_0^t\frac{\partial}{\partial q^i}\left(h_0\left(q,\frac{\partial\widetilde{S}_s}{\partial q}(q)\right)\right)ds-\sum_{r=1}^{n}\int_0^t\frac{\partial^2\widetilde{S}_s}{\partial q^i\partial q^r}(q)\delta B_s^r.$$

Therefore, disregarding all the finite variation terms in this last expression, we have

$$\begin{aligned}\left[\frac{\partial\widetilde{S}}{\partial q^i}(q),B^i\right]_t&=-\sum_{r=1}^{n}\left[\int\frac{\partial^2\widetilde{S}_s}{\partial q^i\partial q^r}(q)dB_s^r,\int dB_s^i\right]_t=-\sum_{r=1}^{n}\int_0^t\frac{\partial^2\widetilde{S}_s}{\partial q^i\partial q^r}(q)d[B^r,B^i]_s\\&=-2\sum_{r=1}^{n}\int_0^t\frac{\partial^2\widetilde{S}_s}{\partial q^i\partial q^r}(q)\delta^{ir}ds=-2\int_0^t\frac{\partial^2\widetilde{S}_s}{(\partial q^i)^2}(q)ds.\end{aligned}$$

In the previous equalities we have used the property

$$\left[\int HdX, \int KdY\right]_t = \int_0^t H_s K_s d[X,Y]_s$$

that holds for arbitrary real semimartingales H, K, X, and Y ([P05, Chapter II Theorem 29]). Thus

$$\widetilde{S}_t\,(q) = f(q) + \int_0^t \left(\Delta \widetilde{S}_s(q) - h_0\left(q, \frac{\partial \widetilde{S}_s}{\partial q}(q)\right)\right) ds - \sum_{i=1}^n \int_0^t \frac{\partial \widetilde{S}_s}{\partial q^i}(q) dB_s^i.$$

Let now $\Phi_t(q) := E[\exp(-\widetilde{S}_t\,(q))]$. By the Itô formula,

$$\mathrm{e}^{-\widetilde{S}_t(q)} - \mathrm{e}^{-f(q)} = -\int_0^t \mathrm{e}^{-\widetilde{S}_s(q)}\, d\widetilde{S}_s\,(q) + \frac{1}{2}\int_0^t \mathrm{e}^{-\widetilde{S}_s(q)}\, d[\widetilde{S}\,(q)\,, \widetilde{S}\,(q)]_s$$

$$= \int_0^t \mathrm{e}^{-\widetilde{S}_s(q)} \left(h_0\left(q, \frac{\partial \widetilde{S}_s}{\partial q}(q)\right) - \Delta \widetilde{S}_s(q) + \frac{1}{2}\sum_{i=1}^n \left(\frac{\partial \widetilde{S}_s}{\partial q^i}(q)\right)^2\right) ds + \int_0^t \mathrm{e}^{-\widetilde{S}_s(q)}\, \frac{\partial \widetilde{S}_s}{\partial q^i}(q) dB_s^i. \tag{2.75}$$

Taking expectations in both sides of (2.75), assuming that all the processes involved are regular enough so that Fubini's Theorem may be invoked, and imposing $h_0 = \frac{1}{2}\sum_{i=1}^n p_i^2 + V(q)$, $V \in C^\infty(\mathbb{R}^n)$, we obtain

$$\begin{aligned} \frac{\partial}{\partial t}\Phi_t(q) &= \Phi_t(q)V(q) + E\left[\mathrm{e}^{-\widetilde{S}_t(q)}\left(\sum_{i=1}^n \left(\frac{\partial \widetilde{S}_t}{\partial q^i}(q)\right)^2 - \Delta \widetilde{S}_t(q)\right)\right] \\ &= V(q)\Phi_t(q) + \Delta \Phi_t(q). \end{aligned}$$

This shows that the projected stochastic action $\widetilde{S}_t$ can be used to construct solutions of the heat equation modified with a potential term V, with initial condition given by the function $\exp(f) \in C^\infty(\mathbb{R}^n)$.

2.4.2 The Hamilton-Jacobi equation and generating functions

One of the main features of the Hamilton-Jacobi equation is that its solutions can be used as generating functions of time-dependent symplectomorphisms that transform the original Hamiltonian system in such a way that its solutions can be easily written down. The natural framework for carrying this out is that of time-dependent Hamiltonian systems; that is why we have included this subsection that briefly recalls the classical theory of non-autonomous Hamiltonian systems and presents it in a form that is suitable for generalization in the stochastic context. Some of the statements there are either inspired or are a direct generalization of analogous results in [B81]; we have nevertheless included them in order to have a complete and self-contained presentation of the theory.

The deterministic case

We start by recalling the relation between the Hamilton-Jacobi equation and the generating functions for integrating canonical transformations in the classical deterministic case. In the next paragraphs we will write down some classical results in a form that is well adapted for the subsequent generalization to the stochastic case. All along this section we will consider Hamiltonian systems on cotangent bundles $(T^*Q, \omega = -\mathbf{d}\theta)$ endowed with their canonical symplectic forms.

Consider the manifold $T^*Q \times T^*Q$ endowed with the symplectic form $\Omega := \tau_1^*\omega - \tau_2^*\omega$, where $\tau_i : T^*Q \times T^*Q \to T^*Q$, $i = 1, 2$, denote the canonical projections onto the first and the second factors, respectively. Let now $\psi : T^*Q \to T^*Q$ be a smooth function. It is easy to verify that the map ψ is a symplectomorphism if and only if $\iota_\psi^*\Omega = 0$, where $\iota_\psi : L^\psi \hookrightarrow T^*Q \times T^*Q$ is the inclusion of the graph L^ψ of ψ ([AM78, Proposition 5.2.1]), in which case is a Lagrangian submanifold of $T^*Q \times T^*Q$. Given that $\Omega = -\mathbf{d}\Theta$, with $\Theta = \tau_1^*\theta - \tau_2^*\theta$, we have that $0 = \iota_\psi^*\Omega = -\iota_\psi^*(\mathbf{d}\Theta) = -\mathbf{d}(\iota_\psi^*\Theta)$ and hence by Poincaré's Lemma, we can locally write $\iota_\psi^*\Theta = \mathbf{d}S$, for some function $S \in C^\infty\left(L^\psi\right)$. We will say that S is a local ***generating function*** for the symplectic map ψ. In addition, suppose that

$$\tau : T^*Q \times T^*Q \to Q \times Q, \quad \tau = \pi \circ \tau_1 \times \pi \circ \tau_2 \tag{2.76}$$

with $\pi : T^*Q \to Q$ the canonical projection, is a local diffeomorphism when restricted to L^ψ and denote its (local) inverse by $\tau^{-1} : Q \times Q \to L^\psi$. We will suppose throughout this section that this is the case and we will think of the generating function $S \in C^\infty\left(L^\psi\right)$ as a function defined on $Q \times Q$; that is, we will not distinguish between S and $\left(\tau^{-1}\right)^* S$. With this convention, we can write

$$\mathbf{d}_{Q\times Q}S = \left(\tau^{-1}\right)^* \circ \iota_\psi^*(\Theta). \tag{2.77}$$

Let now $\{\psi_t\}_{t\in\mathbb{R}}$ be a family of symplectomorphisms depending smoothly on $t \in \mathbb{R}$ (for example $\{\psi_t\}_{t\in\mathbb{R}}$ could be the flow of a Hamiltonian vector field) and let $S : \mathbb{R} \times Q \times Q \to \mathbb{R}$ be the corresponding generating functions associated to this family. We will say that ψ_t ***transforms a vector field*** $X \in \mathfrak{X}(T^*Q)$ ***to equilibrium*** if $T\psi_t(X) = 0$ for any $t \in \mathbb{R}$. For example, if $X = X_h$ is the Hamiltonian vector field associated to a Hamiltonian function $h \in C^\infty(T^*Q)$ and ψ_t transforms X_h to equilibrium, then the integral curve γ of X_h with initial condition z is

$$\gamma_t = \hat{\psi}^{-1}(\psi_0(z), t)$$

where $\hat{\psi}^{-1}$ is the inverse of the diffeomorphism $\hat{\psi} : T^*Q \times \mathbb{R} \to T^*Q \times \mathbb{R}$ given by $(z,t) \mapsto (\psi_t(z), t)$. The main goal of the classical Hamilton-Jacobi theory in this context is proving that ψ transforms X_h to equilibrium if, roughly speaking, its generating function S satisfies the (deterministic) Hamilton-Jacobi equation. As we deal with time-dependent transformations ψ_t of the phase space, the time-dependent Hamiltonian formalism is more convenient.

Time-dependent Hamiltonian systems

Recall that, for time-dependent Hamiltonian systems, the phase space T^*Q is replaced with the *extended* phase space $\mathbb{R}\times T^*Q$. Given a time-dependent Hamiltonian function $h \in C^\infty(\mathbb{R} \times T^*Q)$, one introduces $\Omega_h \in \Omega^2(\mathbb{R} \times T^*Q)$ as $\Omega_h = \mathbf{d}h \wedge \mathbf{d}t + \omega$, where $\omega \in \Omega^2(T^*Q)$ is the canonical symplectic form and t denotes the global time coordinate in $\mathbb{R}$. Observe that Ω_h is exact, $\Omega_h = -\mathbf{d}\theta_h$, where $\theta_h = \theta - h\mathbf{d}t$ and θ is the canonical Liouville one form on the cotangent bundle. Then, the Hamiltonian vector field $X_h \in \mathfrak{X}(\mathbb{R} \times T^*Q)$ is characterized by the two equations

$$i_{X_h}\Omega_h = 0, \quad T\pi_{\mathbb{R}}(X_h) = \frac{\partial}{\partial t},$$

where $\pi_{\mathbb{R}} : \mathbb{R} \times T^*Q \to \mathbb{R}$ is the projection onto the first factor.

Sometimes it is more convenient to encode time-dependent Hamiltonian systems as autonomous Hamiltonian systems on the symplectic manifold $E := T^*(\mathbb{R} \times Q) = T^*\mathbb{R} \times T^*Q$: let (t, u) be global coordinates for $T^*\mathbb{R}$, that is u is the conjugate momentum associated to the time t, and denote by $\pi_{\mathbb{R}\times T^*Q} : T^*\mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ the projection $((t, u), z) \mapsto (t, z)$, with $z \in T^*Q$. It is straightforward to check that the Hamiltonian vector field $X_{h^\star}$ associated to the function $h^\star := u + \pi^*_{\mathbb{R}\times T^*Q}(h) \in C^\infty(E)$ is such that $T\pi_{\mathbb{R}\times T^*Q}(X_{h^\star}) = X_h$. In other words, any time-dependent Hamiltonian system may be visualized as an autonomous Hamiltonian system by replacing $\mathbb{R} \times T^*Q$ by E and h by $h^\star$; the integral curves of the original system X_h are simply obtained form the integral curves of the autonomous system $X_{h^\star}$ by dropping the additional degree of freedom u, which is irrelevant as far as the dynamical description of the system is concerned. The following proposition deals with a time-dependent family of symplectomorphisms $\{\psi_t\}_{t\in\mathbb{R}}$ of T^*Q in the enlarged phase space E and will be useful in order to transform time-dependent Hamiltonian systems.

Proposition 2.42 *Let $\{\psi_t\}_{t\in\mathbb{R}}$ be a family of symplectomorphisms of T^*Q and $S \in C^\infty(\mathbb{R}\times Q \times Q)$ its generating function. Define*

$$\begin{array}{rcl} \bar{\psi} \ : E & \longrightarrow & E \\ (t, u, z) & \longmapsto & (t, u, \psi_t(z)), \end{array}$$

where $t \in \mathbb{R}$, $u \in \mathbb{R}$, $z \in T^*Q$, and

$$\begin{array}{rcl} J_t : T^*Q & \longrightarrow & Q \times Q \\ z & \longmapsto & (\pi(z), \pi(\psi_t(z))). \end{array} \tag{2.78}$$

Then,

(i) *$\omega_E = \bar{\psi}^*(\omega_E) + \mathbf{d}\left(\frac{\partial S}{\partial t} \circ J \circ \pi_{\mathbb{R}\times T^*Q}\right) \wedge \mathbf{d}t$, where ω_E denotes the canonical symplectic two form of $E = T^*(\mathbb{R} \times Q)$.*

(ii) *$\bar{\psi}^*(\omega_E)$ is non-degenerate and, for any $\alpha \in \Omega(\mathbb{R} \times T^*Q)$ and any $h \in C^\infty(\mathbb{R} \times T^*Q)$,*

$$\mathbf{d}h^\star\left(\omega_E^\# \circ \bar{\psi}^* \circ \pi_{\mathbb{R}\times T^*Q}(\alpha)\right) = \mathbf{d}\left(h \circ \hat{\psi}^{-1} + \frac{\partial S}{\partial t} \circ J_t \circ \hat{\psi}^{-1}\right)^\star \left(\omega_E^\# \circ \pi_{\mathbb{R}\times T^*Q}(\alpha)\right) \circ \bar{\psi}$$

Proof. (i) Let $((t, u), (q^i, p_i; i = 1, ..., n))$ be local coordinates on a suitable open neighborhood $U \subseteq E$. It is immediate to see from (2.77) that for any $z \in T^*Q$

$$p_i(z) = \frac{\partial S}{\partial q_1^i}(t, J_t(z)) \quad \text{and} \quad p_i(\psi_t(z)) = -\frac{\partial S}{\partial q_2^i}(t, J_t(z)),$$

$i=1,...,n$ (see, for instance, (7.9.1) in [MR99]), which implies that the canonical one-form $\theta_E := u\mathbf{d}t + \sum_{i=1}^{n} p_i \mathbf{d}q^i$ locally equals

$$\bar{\psi}^*(\theta_E) + \mathbf{d}S \circ J \circ \pi_{\mathbb{R}\times T^*Q} - \frac{\partial S}{\partial t} \circ J \circ \pi_{\mathbb{R}\times T^*Q}\mathbf{d}t$$

(see, for instance, (7.9.5) in [MR99]). Applying $-\mathbf{d}$ to this expression, the result follows.

(ii) By (i), $(\bar{\psi}^{-1})^*\omega_E = \omega_E + \mathbf{d}\left(\frac{\partial S}{\partial t} \circ J \circ \pi_{\mathbb{R}\times T^*Q} \circ \bar{\psi}^{-1}\right) \wedge \mathbf{d}t$. In order to simplify our notation let $F := \frac{\partial S}{\partial t} \circ J_t \circ \hat{\psi}^{-1}$. Then, using $\{\mathbf{d}t, \mathbf{d}u, \mathbf{d}q^i, \mathbf{d}p_i\}_{i=1,\ldots,n}$ and $\{\frac{\partial}{\partial t}, \frac{\partial}{\partial u}, \frac{\partial}{\partial q^i}, \frac{\partial}{\partial p_i}\}_{i=1,\ldots,n}$ as bases of $T^*_{\bar{\psi}(m)}U$ and $T_{\bar{\psi}(m)}U$ respectively, we have the relations

$$\begin{array}{ll}
\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}(\mathbf{d}t) = -\frac{\partial}{\partial u}, & \omega_E^{\#}(\mathbf{d}t) = -\frac{\partial}{\partial u}, \\
\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}(\mathbf{d}u) = \frac{\partial}{\partial t} + \sum_{i=1}^{n}\left(\frac{\partial F}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial F}{\partial q^i}\frac{\partial}{\partial p_i}\right), & \omega_E^{\#}(\mathbf{d}u) = \frac{\partial}{\partial t}, \\
\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}(\mathbf{d}q^i) = -\frac{\partial F}{\partial p_i}\frac{\partial}{\partial u} - \frac{\partial}{\partial p_i}, & \omega_E^{\#}(\mathbf{d}q^i) = -\frac{\partial}{\partial p_i}, \\
\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}(\mathbf{d}p_i) = \frac{\partial F}{\partial q^i}\frac{\partial}{\partial u} + \frac{\partial}{\partial q^i}, & \omega_E^{\#}(\mathbf{d}p_i) = \frac{\partial}{\partial q^i},
\end{array} \tag{2.79}$$

which easily shows the non-degeneracy of $\left(\bar{\psi}^{-1}\right)^*\omega_E$.

Let now $g \in C^\infty(\mathbb{R}\times T^*Q)$, $\alpha \in \Omega(\mathbb{R}\times T^*Q)$, and $g^\star = u + \pi^*_{\mathbb{R}\times T^*Q}(g)$. Using (2.79), it is straightforward to check that

$$\mathbf{d}g^\star\left[\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}\left(\pi^*_{\mathbb{R}\times T^*Q}(\alpha)\right)\right] = \mathbf{d}(g+F)^\star\left[\omega_E^{\#}\left(\pi^*_{\mathbb{R}\times T^*Q}(\alpha)\right)\right]. \tag{2.80}$$

Additionally, for any $m \in U \subseteq E$, the following diagram commutes:

$$\begin{array}{ccc}
T_m^*E & \overset{\omega_E^{\#}(m)}{\longrightarrow} & T_mE \\
{\scriptstyle T_m^*\bar{\psi}}\uparrow & & \downarrow{\scriptstyle T_m\bar{\psi}} \\
T^*_{\bar{\psi}(m)}E & \overset{\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}(\bar{\psi}(m))}{\longrightarrow} & T_{\bar{\psi}(m)}E.
\end{array} \tag{2.81}$$

Therefore, by (2.81), for any $\beta \in \Omega(E)$ and any $h \in C^\infty(\mathbb{R}\times T^*Q)$,

$$\begin{aligned}
\mathbf{d}h^\star\left[\omega_E^{\#}\circ\bar{\psi}^*(\beta)\right](m) &= \mathbf{d}h^\star(m)\left[\omega_E^{\#}(m)\left[T_m^*\bar{\psi}\left(\beta\left(\bar{\psi}(m)\right)\right)\right]\right] \\
&= \mathbf{d}h^\star(m)\left[T_{\bar{\psi}(m)}\bar{\psi}^{-1}\left[\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}\left(\bar{\psi}(m)\right)\left[\beta\left(\bar{\psi}(m)\right)\right]\right]\right] \\
&= \mathbf{d}\left((\bar{\psi}^{-1})^*h^\star\right)\left(\bar{\psi}(m)\right)\left[\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}\left(\bar{\psi}(m)\right)\left[\beta\left(\bar{\psi}(m)\right)\right]\right] \\
&= \mathbf{d}\left((\bar{\psi}^{-1})^*h^\star\right)\left(\bar{\psi}(m)\right)\left[\left((\bar{\psi}^{-1})^*\omega_E\right)^{\#}\left(\bar{\psi}(m)\right)\left[\beta\left(\bar{\psi}(m)\right)\right]\right].
\end{aligned}$$

In addition, if β is of the form $\pi^*_{\mathbb{R}\times T^*Q}(\alpha)$ for some $\alpha \in \Omega(\mathbb{R}\times T^*Q)$, by (2.80) with $g = (\hat{\psi}^{-1})^* h$ we have

$$\mathbf{d}h^\star\left[\omega_E^\# \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha)\right](m) = \mathbf{d}\left(\left((\hat{\psi}^{-1})^* h + F\right)^\star\right)(\bar{\psi}(m))\left[\left(\omega_E^\# \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha)\right)(\bar{\psi}(m))\right].$$

Since $F = \frac{\partial S}{\partial t} \circ J_t \circ \hat{\psi}^{-1}$, the expression in (ii) follows. ■

Proposition 2.43 *Let $h \in C^\infty(\mathbb{R}\times T^*Q)$. With the same notation as in Proposition 2.42, a curve $\gamma : [0,T] \to \mathbb{R}\times T^*Q$ is a solution of the Hamiltonian system defined by h if and only if, for any family of symplectomorphisms $\{\psi_t\}_{t\in\mathbb{R}}$ of T^*Q, the curve $\psi \circ \gamma : [0,T] \to \mathbb{R}\times T^*Q$ such that $(\hat{\psi}\circ\gamma)(t) := (t, \psi_t(\gamma(t)))$ is a solution of a Hamiltonian system with Hamiltonian function*

$$h' = h \circ \hat{\psi}^{-1} + \frac{\partial S}{\partial t} \circ J \circ \hat{\psi}^{-1} \tag{2.82}$$

where $S \in C^\infty(\mathbb{R}\times Q\times Q)$ is the generating function of $\{\psi_t\}_{t\in\mathbb{R}}$.

Proof. Let $\gamma : [0,T] \to \mathbb{R}\times T^*Q$ be a solution of the time-dependent Hamiltonian system defined by h. Let $\bar{\gamma} : [0,T] \to E = T^*(\mathbb{R}\times Q)$ be the curve such that $\gamma = \pi^*_{\mathbb{R}\times T^*Q}(\bar{\gamma})$ and $\dot{u} = \frac{\partial h}{\partial t}(\gamma)$, u being the conjugate momenta of the time coordinate t. Then γ is a solution of the time-dependent Hamiltonian system defined by $h \in C^\infty(\mathbb{R}\times T^*Q)$ if and only if $\bar{\gamma}$ is a solution of the autonomous Hamilton system on the phase space E with Hamiltonian function $h^\star = u + \pi^*_{\mathbb{R}\times T^*Q}(h)$. By (2.2), this means that for any $\beta \in \Omega(E)$,

$$\int_{\bar{\gamma}|_{[0,t]}} \beta = -\int_0^t \mathbf{d}h^\star\left(\omega_E^\#(\beta)\right)\circ\gamma(s)ds \tag{2.83}$$

for any $t \in [0,T]$. However, since we are not interested in the evolution of u, the conjugate momentum of the time, verifying that γ is a solution of the time-dependent Hamilton equations is equivalent to taking any curve $\bar{\gamma}$ such that $\gamma = \pi^*_{\mathbb{R}\times T^*Q}(\bar{\gamma})$ and checking that (2.83) holds for any differential form of the type $\pi^*_{\mathbb{R}\times T^*Q}(\alpha)$, $\alpha \in \Omega(\mathbb{R}\times T^*Q)$.

Let now $\{\psi_t\}_{t\in\mathbb{R}}$ be a time-dependent family of symplectomorphisms of T^*Q and consider $\hat{\psi} : \mathbb{R}\times T^*Q \to \mathbb{R}\times T^*Q$ such that $\hat{\psi}(t,z) = (t, \psi_t(z))$, $(t,z) \in \mathbb{R}\times T^*Q$, and $\bar{\psi} : E \to E$ such that $\bar{\psi}(t,u,z) = (t,u,\psi_t(z))$ as in Proposition 2.42. Let $\bar{\psi}\circ\bar{\gamma} : [0,T] \to E$ be defined as $(\bar{\psi}\circ\bar{\gamma})(s) := \bar{\psi}_s(\bar{\gamma}(s))$. Then

$$\begin{aligned}\int_{\bar{\psi}\circ\bar{\gamma}|_{[0,t]}} \pi^*_{\mathbb{R}\times T^*Q}(\alpha) &= \int_{\bar{\gamma}|_{[0,t]}} \bar{\psi}^*\left(\pi^*_{\mathbb{R}\times T^*Q}(\alpha)\right) = -\int_0^t \mathbf{d}h^\star\left(\omega_E^\# \circ \bar{\psi}^*(\pi^*_{\mathbb{R}\times T^*Q}(\alpha))\right)\circ\bar{\gamma}(s)ds \\ &= -\int_0^t \mathbf{d}\left(h\circ\hat{\psi}^{-1} + \frac{\partial S}{\partial t}\circ J\circ\hat{\psi}^{-1}\right)^\star\left(\omega_E^\#(\pi^*_{\mathbb{R}\times T^*Q}(\alpha))\right)\circ\left(\bar{\psi}\circ\bar{\gamma}\right)(s)ds\end{aligned}$$

where Proposition 2.42 (ii) have been used in the last equality. Hence, we conclude that $\pi_{\mathbb{R}\times T^*Q}(\bar{\psi}\circ\bar{\gamma}) = \hat{\psi}\circ\gamma$ is a solution of the time-dependent Hamiltonian system given by $(\hat{\psi}^{-1})^*(h + \frac{\partial S}{\partial t}\circ J)$. The converse is left to the reader. ■

The content of Proposition 2.43 can be restated as follows. Given $h \in C^{\infty}(\mathbb{R} \times T^*Q)$ and a family of symplectomorphisms $\{\psi_t\}_{t\in\mathbb{R}}$, there exists a smooth function $h' \in C^{\infty}(\mathbb{R} \times T^*Q)$ such that $\Omega_h = \hat{\psi}^* \Omega_{h'}$, where $\Omega_{h'} = \mathbf{d}h' \wedge \mathbf{d}t + \omega$ and h' is given by (2.82) (see [MR99, Section 7.9]). Furthermore $T\hat{\psi}^{-1}(X_h)$ is the Hamiltonian vector field related to h' and the flow of $X_{h'}$ restricted to the phase space T^*Q is $\hat{\varphi}_t = \psi_t^{-1} \circ \varphi_t \circ \psi_0$ where, as usual, φ denotes the flow of symplectomorphisms of the Hamiltonian vector field $X_h \in \mathfrak{X}(T^*Q)$. However, as we will be interested in transforming X_h using $T\psi$ rather than $T\psi^{-1}$ we will rewrite (2.82) in the form

$$h'(t, \psi_t(z)) := h(z) + \frac{\partial S}{\partial t}(t, J_t \circ z). \tag{2.84}$$

Definition 2.44 *Let $h \in C^{\infty}(T^*Q)$ be a Hamiltonian function and let $\left(q^i, p_i; i = 1, ..., n\right)$ be local Darboux coordinates on T^*Q. Regarding h as a function of these coordinates, we will say that the generating function $S : \mathbb{R} \times Q \times Q \to \mathbb{R}$ satisfies the (deterministic) Hamilton-Jacobi equation if the function $K : \mathbb{R} \times Q \times Q \to \mathbb{R}$*

$$K_t(q_1, q_2) := h\left(q_1, \frac{\partial S}{\partial q_1}(t, q_1, q_2)\right) + \frac{\partial S}{\partial t}(t, q_1, q_2), \quad (q_1, q_2) \in Q \times Q \tag{2.85}$$

does not depend on the first entry $q_1 \in Q$.

Observe that in the right hand side of (2.85) we have carried out the substitution $(p_1)_i = \frac{\partial S}{\partial q_1^i}(t, q_1, q_2)$, $i = 1, ..., n$. We could also write (2.85) more intrinsically as

$$h(\mathbf{d}_{Q_1} S(t, q_1, q_2)) + \frac{\partial S}{\partial t}(t, q_1, q_2)$$

where, for a fixed value $(t, q_2) \in \mathbb{R} \times Q$, we consider $\mathbf{d}_{Q_1} S(t, q_1, q_2)$ as an element in $T^*_{q_1} Q$.

Notice that the map J_t introduced in (2.78) is a local diffeomorphism for any $t \in \mathbb{R}$ because we required the projection τ defined in (2.76) to be a local diffeomorphism when restricted to the graph of ψ_t. We may therefore (locally) write any $z \in T^*Q$ as $z = J_t^{-1}(q_1, q_2)$ for some suitable $(q_1, q_2) \in Q \times Q$. The important point is that $J_t^{-1}(q_1, q_2) = \mathbf{d}_{Q_1} S(t, q_1, q_2)$ ([MR99, (7.9.1)]) and, consequently, the transformed Hamiltonian h' in (2.84) can be seen as a function on $\mathbb{R} \times Q \times Q$. Explicitly, if $\bar{z} = \psi_t(z) \in T^*Q$,

$$h'(t, \bar{z}) = h(\mathbf{d}_{Q_1} S(t, q_1, q_2)) + \frac{\partial S}{\partial t}(t, q_1, q_2), \tag{2.86}$$

so $h'(t, \bar{z})$ equals the function $K_t(q_1, q_2)$ introduced in Definition 2.44. Suppose now that $S : \mathbb{R} \times Q \times Q \to \mathbb{R}$ is a solution to the Hamilton-Jacobi equation. In other words, $K_t(q_1, q_2) \equiv K_t(q_2)$. Since $q_2 = \pi(\psi_t(z))$ is the base point in the configuration space of the transformed point $\psi_t(z)$, $z \in T^*Q$, we conclude that h' does not depend on the fiber coordinates. Hence, removing the subindices, the Hamilton equations associated to the new Hamiltonian h' are

$$\dot{q}^i = 0, \quad \dot{p}_i = -\frac{\partial K}{\partial q^i}(t, q), \quad i = 1, ..., n,$$

which are easily integrable. In particular, if K is independent of both q_1 and q_2, then ψ_t transforms X_h to equilibrium.

The stochastic case

We are now going to see that the classical Hamilton-Jacobi that we just outlined has a stochastic counterpart. More specifically, one may use a time-dependent family of symplectomorphisms and their generating function to transform a stochastic Hamiltonian system into another one in much the same fashion as in the deterministic case. The strategy consists of finding and characterizing a suitable generating function so that the new Hamiltonian system is easier to solve.

Let T^*Q be the cotangent bundle of the configuration space manifold Q and let $\{h_0, h_1, ..., h_r\} \subset C^\infty(T^*Q)$ be a family of functions. Take a $\mathbb{R}^{r+1}$-valued semimartingale $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{r+1}$ such that

$$X = \left(X^0, X^1, ..., X^r\right), \quad \text{with} \quad X^0 = t \quad \text{a.s.}, \tag{2.87}$$

and consider the stochastic Hamiltonian system on T^*Q with Hamiltonian function $h := (h_0, h_1, ..., h_r)$ and stochastic component X. If we want to remove the assumption that there is a Hamiltonian vector field, i.e. X_{h_0}, playing the role of a deterministic drift, we may simply choose $h_0 = 0$.

Using an approach similar to the one in 2.4.2, we will work in the extended phase space $E := T^*(\mathbb{R} \times Q)$. Indeed, it is easy to check that the solution semimartingales of the stochastic Hamiltonian system can be obtained out of the solutions of the stochastic Hamiltonian system on E with Hamiltonian function $\bar{h} = (h_0^\star, \pi^*_{\mathbb{R}\times T^*Q}(h_1), ..., \pi^*_{\mathbb{R}\times T^*Q}(h_r))$ and stochastic component X; notice that the functions $h_0, h_1, ..., h_r$ have already been considered as functions on $\mathbb{R} \times T^*Q$ instead of only T^*Q. The solutions of the original system can be recovered by composing the solutions of the Hamiltonian system on E with $\pi_{\mathbb{R}\times T^*Q}$. When instead of working on the space E one uses directly $\mathbb{R} \times T^*Q$ instead of T^*Q then a T^*Q-valued semimartingale Γ is a solution of the corresponding stochastic Hamiltonian system when for any $\alpha \in \Omega(T^*Q)$,

$$\int \langle \alpha, \delta\Gamma_s \rangle = -\int \left\langle \mathbf{d}h \left(\tau^*_{T^*Q} \circ \omega^\#(\alpha)\right)(s, \Gamma_s), \delta X_s \right\rangle,$$

where $\tau_{T^*Q} : \mathbb{R} \times T^*Q \to T^*Q$ is the canonical projection onto the second factor.

Proposition 2.45 *Let $\{\psi_t\}_{t\in\mathbb{R}}$ be a time-dependent family of symplectomorphisms of T^*Q with generating function $S \in C^\infty(\mathbb{R} \times Q \times Q)$. Consider $\hat{\psi} : \mathbb{R} \times T^*Q \to \mathbb{R} \times T^*Q$ and $\bar{\psi} : E \to E$ the natural diffeomorphisms extending ψ to $\mathbb{R} \times T^*Q$ and E respectively. Then the semimartingale $\Gamma : \mathbb{R}_+ \times \Omega \to T^*Q$ is a solution of the Hamiltonian system with Hamiltonian function $h : T^*Q \to \mathbb{R}^{r+1}$, $h = (h_0, h_1, ..., h_r)$, and stochastic component $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{r+1}$ as in (2.87), if and only if $\psi(\Gamma)$ is a solution of the Hamiltonian system with Hamiltonian function $h' : \mathbb{R} \times T^*Q \to \mathbb{R}^{r+1}$ with components given by*

$$\begin{aligned} h'_0 &= \tau^*_{T^*Q}(h_0) \circ \hat{\psi}^{-1} + \frac{\partial S}{\partial t} \circ J \circ \hat{\psi}^{-1}, \\ h'_1 &= \tau^*_{T^*Q}(h_1) \circ \hat{\psi}^{-1}, \\ &\vdots \\ h'_r &= \tau^*_{T^*Q}(h_r) \circ \hat{\psi}^{-1}. \end{aligned} \tag{2.88}$$

and stochastic component X.

Proof. Suppose that $\Gamma : \mathbb{R}_+ \times \Omega \to T^*Q$ is a solution of the Hamiltonian system with Hamiltonian function h and stochastic component X and let $\bar{\Gamma} : \mathbb{R}_+ \times \Omega \to E$ be a semimartingale such that $\pi_{\mathbb{R}\times T^*Q}(\bar{\Gamma}_t) = (t, \Gamma_t) \in \mathbb{R} \times T^*Q$, $t \in \mathbb{R}_+$. We want to check that $\bar{\psi}(\bar{\Gamma})$ is a solution of the stochastic Hamiltonian system given by the Hamiltonian function (2.88). Let $\alpha \in \Omega\,(\mathbb{R} \times T^*Q)$. Since Γ is a solution, we may write

$$\int \left\langle \pi^*_{\mathbb{R}\times T^*Q}(\alpha), \delta\bar{\psi}(\bar{\Gamma}) \right\rangle = \int \left\langle \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha), \delta\bar{\Gamma} \right\rangle = -\int \left\langle \mathbf{d}\bar{h}(\omega^{\#}_E \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\Gamma}), \delta X \right\rangle$$
$$= -\int \mathbf{d}h^{\star}_0(\omega^{\#}_E \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\Gamma})dt - \sum_{i=1}^{r} \int \mathbf{d}(\pi^*_{T^*Q} h_i)(\omega^{\#}_E \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\Gamma})\delta X^i,$$

where $\pi_{T^*Q} : E = T^*\mathbb{R} \times T^*Q \to T^*Q$ is the projection onto the second factor. Now, by Proposition 2.43 we have

$$\int \mathbf{d}h^{\star}_0(\omega^{\#}_E \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\Gamma})dt =$$
$$\int \mathbf{d}\left(\tau_{T^*Q}(h_0) \circ \hat{\psi}^{-1} + \frac{\partial S}{\partial t} \circ J_t \circ \hat{\psi}^{-1}\right)^{\star} \left(\omega^{\#}_E \circ \pi_{\mathbb{R}\times T^*Q}\,(\alpha)\right) \left(\bar{\psi}(\bar{\Gamma})\right) dt. \tag{2.89}$$

On the other hand, using (2.79) and (2.81) it is easy to see that for any $g \in C^\infty\,(T^*Q)$

$$\mathbf{d}(\pi^*_{T^*Q} g)(\omega^{\#}_E \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(m) = \mathbf{d}(\pi^*_{T^*Q}(g) \circ \bar{\psi}^{-1})(\omega^{\#}_E \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\psi}(m)).$$

Consequently,

$$\int \mathbf{d}(\pi^*_{T^*Q} h_i)(\omega^{\#}_E \circ \bar{\psi}^* \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\Gamma})\delta X^i$$
$$= \int \mathbf{d}(\pi^*_{T^*Q}(h_i) \circ \bar{\psi}^{-1})(\omega^{\#}_E \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\psi}(\bar{\Gamma}))\delta X^i, \tag{2.90}$$

for any $i = 1, ..., r$. Combining (2.89) and (2.90) we obtain that

$$\int \left\langle \pi^*_{\mathbb{R}\times T^*Q}(\alpha), \delta\bar{\psi}(\bar{\Gamma}) \right\rangle =$$
$$-\int \mathbf{d}\left(\tau_{T^*Q}(h_0) \circ \hat{\psi}^{-1} + \frac{\partial S}{\partial t} \circ J_t \circ \hat{\psi}^{-1}\right)^{\star} \left(\omega^{\#}_E \circ \pi_{\mathbb{R}\times T^*Q}\,(\alpha)\right) \left(\bar{\psi}(\bar{\Gamma})\right) dt$$
$$-\int \mathbf{d}(\pi^*_{T^*Q}(h_i) \circ \bar{\psi}^{-1})(\omega^{\#}_E \circ \pi^*_{\mathbb{R}\times T^*Q}(\alpha))(\bar{\psi}(\bar{\Gamma}))\delta X^i,$$

which means that $\psi_t\,(\Gamma_t)$ is a solution of the time-dependent stochastic Hamiltonian system with stochastic component X and Hamiltonian function (2.88). The converse is left to the reader. ■

The system (2.88) may be written as

$$\begin{aligned} h_0'(t,\psi_t(z)) &= \tau_{T^*Q}^*(h_0)\left(\mathbf{d}_{Q_1}S\left(t,q_1,q_2\right)\right)+\frac{\partial S}{\partial t}\left(t,q_1,q_2\right) \\ h_1'(t,\psi_t(z)) &= \tau_{T^*Q}^*(h_1)\left(\mathbf{d}_{Q_1}S\left(t,q_1,q_2\right)\right) \\ &\vdots \\ h_r'(t,\psi_t(z)) &= \tau_{T^*Q}^*(h_r)\left(\mathbf{d}_{Q_1}S\left(t,q_1,q_2\right)\right) \end{aligned} \tag{2.91}$$

where, as in (2.86) we have written $z \in T^*Q$ as $z = J_t^{-1}(q_1, q_2)$ for some suitable $(q_1, q_2) \in Q \times Q$. In addition, if the generating function S is such that the right hand side of (2.91) is independent of the variable q_1, that is,

$$\begin{aligned} \tau_{T^*Q}^*(h_0)\left(\mathbf{d}_{Q_1}S\left(t,q_1,q_2\right)\right)+\frac{\partial S}{\partial t}\left(t,q_1,q_2\right) &=: K_0\left(t,q_2\right), \\ \tau_{T^*Q}^*(h_1)\left(\mathbf{d}_{Q_1}S\left(t,q_1,q_2\right)\right) &=: K_1\left(t,q_2\right), \\ &\vdots \\ \tau_{T^*Q}^*(h_r)\left(\mathbf{d}_{Q_1}S\left(t,q_1,q_2\right)\right) &=: K_r\left(t,q_2\right), \end{aligned} \tag{2.92}$$

then the stochastic Hamilton equations of the transformed system may be expressed in local coordinates as

$$\begin{aligned} \delta q^i &= 0 \\ \delta p_i &= -\frac{\partial K_0}{\partial q}\left(t,q\right)dt-\sum_{i=1}^{r}\frac{\partial K_i}{\partial q}\left(t,q\right)\delta X^i. \end{aligned}$$

The next result is basically due to Bismut (see [B81, Théorème 7.6, page 349]).

Proposition 2.46 *In the conditions of the previous proposition, if (2.92) holds then*

$$\begin{aligned} \left\{h_i,h_j\right\}(z) &= 0 \\ \{h_0,h_i\}(z)+\frac{\partial K_i}{\partial t}\left(t,\pi\left(\psi_t(z)\right)\right) &= 0 \end{aligned}$$

locally for any $1 \le i, j \le r$.

Proof. Suppose that there exists a generating function $S \in C^{\infty}(\mathbb{R} \times Q \times Q)$ such that the equalities (2.92) are satisfied. We take a fixed point $q_2 \in Q$ and write $K_i^{q_2}(t)$ instead of $K_i(t, q_2)$, $i = 0, ..., r$, and $S^{q_2}(t, q)$ instead of $S(t, q, q_2)$. Consider the following family of functions of the extended phase space $E = T^*(\mathbb{R} \times Q)$:

$$\begin{aligned} g_0 &= u+\pi_{T^*Q}^*(h_0)-K_0^{q_2}\left(t\right) \\ g_1 &= \pi_{T^*Q}^*(h_1)-K_1^{q_2}\left(t\right) \\ &\vdots \\ g_r &= \pi_{T^*Q}^*(h_r)-K_r^{q_2}\left(t\right), \end{aligned}$$

where u denotes the conjugate momentum of the time coordinate t in E. The functions $g_0, ..., g_r \subset C^\infty(E)$ vanish on the Lagrangian submanifold $L_S \subset E$ locally defined by

$$L_S = \left\{ (t, u, q, p) \in E \mid p_i = \frac{\partial S^{q_2}}{\partial q^i}(t, q),\ u = \frac{\partial S^{q_2}}{\partial t}(t, q) \right\}.$$

Given that if a family of functions is locally constant on a Lagrangian submanifold, then their Poisson brackets must vanish on it, we have that $\{g_i, g_j\} = 0$ for any $0 \leq i, j \leq r$. Equivalently,

$$\begin{aligned} 0 &= \left.\{\pi_{T^*Q}^* h_i, \pi_{T^*Q}^* h_j\}\right|_{L_S} = \left.\pi_{T^*Q}^* (\{h_i, h_j\})\right|_{L_S}, \\ 0 &= \left.\pi_{T^*Q}^* (\{h_0, h_i\})\right|_{L_S} + \left.\frac{\partial K_i^{q_2}}{\partial t}\right|_{L_S}, \end{aligned} \tag{2.93}$$

for any $i, j = 1, ..., r$. In particular, since the inverse $J_t^{-1} : Q \times Q \to T^*Q$ of the local diffeomorphism introduced in (2.78) is such that $z = J_t^{-1}(q_1, q_2) = (q_1, \mathbf{d}S^{q_2}(t, q_1))$, we have the freedom to chose q_2 so that $z = J_t^{-1}(q_1, q_2)$ is a point in the fiber of $q_1 \in Q$. With this choice (2.93) implies that

$$\begin{gathered} \{h_i, h_j\}(z) = 0 \\ \{h_0, h_i\}(z) + \frac{\partial K_i^{q_2}}{\partial t}(t) = \{h_0, h_j\}(z) + \frac{\partial K_i}{\partial t}(t, \pi(\psi_t(z))) = 0 \end{gathered}$$

for any $z \in T^*Q$. ■

2.5 Proofs and auxiliary results

2.5.1 Proof of Proposition 2.24

Before proving the proposition, we recall a technical lemma dealing with the convergence of sequences in a metric space.

Lemma 2.47 *Let (E, d) be a metric space. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of functions $x_n : (0, \delta) \to E$ where $(0, \delta) \subset \mathbb{R}$ is an open interval of the real line. Suppose that x_n converges uniformly on $(0, \delta)$ to a function x. Additionally, suppose that for any n, the limits $\lim_{s\to 0} x_n(s) = x_n^* \in E$ exist and so does $\lim_{n\to\infty} x_n^*$. Then*

$$\lim_{s\to 0} x(s) = \lim_{n\to\infty} x_n^*.$$

Proof. Let $\varepsilon > 0$ be an arbitrary real number. We have

$$d\left(x(s), \lim_{n\to\infty} x_n^*\right) \leq d(x(s), x_k(s)) + d(x_k(s), x_k^*) + d\left(x_k^*, \lim_{n\to\infty} x_n^*\right).$$

From the definition of limit and since $x_k(s)$ converges uniformly to x on $(0, \delta)$, we can choose k_0 such that $d(x_k^*, \lim_{n\to\infty} x_n^*) < \frac{\varepsilon}{3}$ and $d(x(s), x_k(s)) < \frac{\varepsilon}{3}$, simultaneously for any $k \geq k_0$.

Additionally, since $\lim_{s \to 0} x_k(s) = x_k^*$ we choose s_0 small enough such that $d(x_k(s), x_k^*) < \frac{\varepsilon}{3}$, for any $s < s_0$. Thus,

$$d\left(x(s), \lim_{n \to \infty} x_n^*\right) < \varepsilon$$

for any $s < s_0$. Since $\varepsilon > 0$ is arbitrary, we conclude that $\lim\limits_{s \to 0} x(s) = \lim\limits_{n \to \infty} x_n^*$. ▼

Proof of Proposition 2.24. First of all, the second equality in (2.43) is a straightforward consequence of [E89, page 93]. Now, let $\{U_k\}_{k \in \mathbb{N}}$ be a countable open covering of M by coordinate patches. By [E89, Lemma 3.5] there exists a sequence $\{\tau_m\}_{m \in \mathbb{N}}$ of stopping times such that $\tau_0 = 0$, $\tau_m \leq \tau_{m+1}$, $\sup_m \tau_m = \infty$, a.s., and that, on each of the sets

$$[\tau_m, \tau_{m+1}] \cap \{\tau_m < \tau_{m+1}\} := \{(t, \omega) \in \mathbb{R}_+ \times \Omega \mid \tau_{m+1}(\omega) > \tau_m(\omega) \text{ and } t \in [\tau_m(\omega), \tau_{m+1}(\omega)]\}$$

the semimartingale Γ takes its values in one of the elements of the family $\{U_k\}_{k \in \mathbb{N}}$. Second, the statement of the proposition is formulated in terms of Stratonovich integrals. However, the proof will be carried out in the context of Itô integration since we will use several times the notion of uniform convergence on compacts in probability (ucp) which behaves well only with respect to this integral. Regarding this point we recall that by the very definition of the Stratonovich integral of a 1-form α along a semimartingale Γ we have that

$$\int \langle \varphi_s^* \alpha, \delta\Gamma \rangle = \int \langle d_2(\varphi_s^* \alpha), d\Gamma \rangle \quad \text{and} \quad \int \langle \pounds_Y \alpha, \delta\Gamma \rangle = \int \langle d_2(\pounds_Y \alpha), d\Gamma \rangle . \tag{2.94}$$

The proof of the proposition follows directly from Lemma 2.47 by applying it to the sequence of functions given by

$$x_n(s) := \left(\int \left\langle \frac{1}{s} \left[d_2(\varphi_s^* \alpha) - d_2(\alpha)\right], d\Gamma \right\rangle\right)^{\tau_n}.$$

This sequence lies in the space $\mathbb{D}$ of càglàd processes endowed with the topology of the ucp convergence. We recall that this space is metric [P05, page 57] and hence we are in the conditions of Lemma 2.47. In the following points we verify that the rest of the hypotheses of this result are satisfied.

(i) The sequence of functions $\{x_n(s)\}_{n \in \mathbb{N}}$ converges uniformly to

$$x(s) := \int \left\langle \frac{1}{s} \left[d_2(\varphi_s^* \alpha) - d_2(\alpha)\right], d\Gamma \right\rangle.$$

The pointwise convergence is a consequence of part **(i)** in Proposition A.2. Moreover, in the proof of that result we saw that if $d : \mathbb{D} \times \mathbb{D} \to \mathbb{R}_+$ is a distance function function associated to the ucp convergence, then for any $t \in \mathbb{R}_+$ and any $s \in (0, \epsilon)$, $d(x_n(s), x(s)) \leq P(\{\tau_n < t\})$. Since the right hand side of this inequality does not depend on s and $P(\{\tau_n < t\}) \to 0$ as $n \to \infty$, the uniform convergence follows.

(ii)

$$\lim_{\substack{ucp \\ s \to 0}} x_n(s) = \left(\int \langle d_2(\pounds_Y \alpha), d\Gamma \rangle\right)^{\tau_n} =: x_n^*.$$

By the construction of the covering $\{U_k\}_{k\in\mathbb{N}}$ and of the stopping times $\{\tau_m\}_{m\in\mathbb{N}}$, there exists a $k(m)\in\mathbb{N}$ such that the semimartingale Γ takes its values in $U_{k(m)}$ when evaluated in the stochastic interval $(\tau_n,\tau_{n+1}]\subset[\tau_n,\tau_{n+1}]\cap\{\tau_n<\tau_{n+1}\}$. Now, since d_2 is a linear operator and $\frac{1}{s}\left((\varphi_s^*\alpha)-\alpha\right)(m)\stackrel{s\to 0}{\longrightarrow}\pounds_Y\alpha(m)$, for any $m\in M$, we have that $\frac{1}{s}\left(d_2\left(\varphi_s^*\alpha\right)-d_2\alpha\right)(m)\stackrel{s\to 0}{\longrightarrow} d_2\left(\pounds_Y\alpha\right)(m)$. Moreover, a straightforward application of Taylor's theorem shows that

$$\frac{1}{s}\left(d_2\left(\varphi_s^*\alpha\right)-d_2\alpha\right)|_{U_{k(m)}}\stackrel{s\to 0}{\longrightarrow} d_2\left(\pounds_Y\alpha\right)|_{U_{k(m)}}$$

uniformly, using a Euclidean norm in $\tau^*U_{k(m)}$ (we recall that $U_{k(m)}$ is a coordinate patch). This fact immediately implies that

$$\mathbf{1}_{(\tau_n,\tau_{n+1}]}\frac{1}{s}\left(d_2\left(\varphi_s^*\alpha\right)-d_2\alpha\right)(\Gamma)\stackrel{s\to 0}{\longrightarrow}\mathbf{1}_{(\tau_n,\tau_{n+1}]}d_2\left(\pounds_Y\alpha\right)(\Gamma)$$

in ucp. As by construction the Itô integral behaves well when we apply it to a ucp convergent sequence of processes we have that

$$\lim_{\substack{ucp\\ s\to 0}}\int\mathbf{1}_{(\tau_n,\tau_{n+1}]}\left\langle\frac{1}{s}\left(d_2\left(\varphi_s^*\alpha\right)-d_2\alpha\right)(\Gamma),d\Gamma\right\rangle=\int\mathbf{1}_{(\tau_n,\tau_{n+1}]}\left\langle d_2\left(\pounds_Y\alpha\right)(\Gamma),d\Gamma\right\rangle. \tag{2.95}$$

Consequently,

$$\begin{aligned}
&\lim_{\substack{ucp\\ s\to 0}}\left(\int\left\langle\frac{1}{s}\left[d_2\left(\varphi_s^*\alpha\right)-d_2\left(\alpha\right)\right],d\Gamma\right\rangle\right)^{\tau_n}\\
&=\lim_{\substack{ucp\\ s\to 0}}\sum_{m=0}^{n-1}\left[\left(\int\left\langle\frac{1}{s}\left[d_2\left(\varphi_s^*\alpha\right)-d_2\left(\alpha\right)\right],d\Gamma\right\rangle\right)^{\tau_{m+1}}-\left(\int\left\langle\frac{1}{s}\left[d_2\left(\varphi_s^*\alpha\right)-d_2\left(\alpha\right)\right],d\Gamma\right\rangle\right)^{\tau_m}\right]\\
&=\lim_{\substack{ucp\\ s\to 0}}\sum_{m=0}^{n-1}\int\mathbf{1}_{(\tau_m,\tau_{m+1}]}\left\langle\frac{1}{s}\left(d_2\left(\varphi_s^*\alpha\right)-d_2\alpha\right),d\Gamma\right\rangle=\sum_{m=0}^{n-1}\int\mathbf{1}_{(\tau_m,\tau_{m+1}]}\left\langle d_2\left(\pounds_Y\alpha\right),d\Gamma\right\rangle\\
&=\left(\int\left\langle d_2\left(\pounds_Y\alpha\right),d\Gamma\right\rangle\right)^{\tau_n},
\end{aligned}$$

where in the second equality we have used Proposition A.1 and the third one follows from (2.95).

(iii)

$$\lim_{n\to\infty}x_n^*=\int\left\langle d_2\left(\pounds_Y\alpha\right),d\Gamma\right\rangle.$$

It is a straightforward consequence of Proposition A.2. The equation (2.43) follows from Lemma 2.47 applied to the sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{x_n^*\}_{n\in\mathbb{N}}$, and using the statements in **(i)**, **(ii)**, and **(iii)**. ■

2.5.2 Proof of Proposition 2.33

We will start the proof by a preparatory result.

Lemma 2.48 *Let $\{X_n\}_{n\in\mathbb{N}}$ and $\{Y_n\}_{n\in\mathbb{N}}$ be two sequences of real valued processes converging in ucp to a couple of processes X and Y respectively. Suppose that, for any $t \in \mathbb{R}_+$, the random variables $\sup_{n\in\mathbb{N}}\sup_{0\le s\le t}\left|(X_n)_s\right|$ and $\sup_{0\le s\le t}|Y_s|$ are bounded (their images lie in a compact set of $\mathbb{R}$). Then, the sequence X_nY_n converges in ucp to XY as $n\to\infty$.*

Proof. We need to prove that for any $\varepsilon > 0$ and any $t \in \mathbb{R}_+$,

$$P\left(\left\{\sup_{0\le s\le t}\left|(X_nY_n)_s - (XY)_s\right| \le \varepsilon\right\}\right) \underset{n\to\infty}{\longrightarrow} 1.$$

First of all, note that

$$\sup_{0\le s\le t}\left|(X_nY_n)_s - (XY)_s\right| \le \sup_{0\le s\le t}|X_n|\,|Y_n - Y| + \sup_{0\le s\le t}|Y|\,|X_n - X|.$$

Hence, we have

$$\begin{aligned}\left\{\sup_{0\le s\le t}\left|(X_nY_n)_s - (XY)_s\right| \le \varepsilon\right\} &\supseteq \left\{\sup_{0\le s\le t}|X_n|\,|Y_n - Y| + \sup_{0\le s\le t}|Y|\,|X_n - X| \le \varepsilon\right\}\\ &\supseteq \left\{\sup_{0\le s\le t}|X_n|\,|Y_n - Y| \le \frac{\varepsilon}{2}\right\} \cap \left\{\sup_{0\le s\le t}|Y|\,|X_n - X| \le \frac{\varepsilon}{2}\right\}.\end{aligned}$$

Denote

$$A_n := \left\{\sup_{0\le s\le t}|X_n|\,|Y_n - Y| \le \frac{\varepsilon}{2}\right\}, \quad \text{and} \quad B_n := \left\{\sup_{0\le s\le t}|Y|\,|X_n - X| \le \frac{\varepsilon}{2}\right\},$$

and let c be a constant such that $\sup_{n\in\mathbb{N}}\sup_{0\le s\le t}\left|(X_n)_s\right| < c$ and $\sup_{0\le s\le t}|Y_s| < c$, available by the boundedness hypothesis. Then,

$$\begin{aligned}1 \ge P(A_n) &\ge P\left(\left\{\sup_{0\le s\le t}|Y_n - Y| \le \frac{\varepsilon}{2c}\right\}\right) \underset{n\to\infty}{\longrightarrow} 1,\\ 1 \ge P(B_n) &\ge P\left(\left\{\sup_{0\le s\le t}|X_n - X| \le \frac{\varepsilon}{2c}\right\}\right) \underset{n\to\infty}{\longrightarrow} 1.\end{aligned}$$

Thus, $P(A_n) \to 1$ and $P(B_n) \to 1$ as $n\to\infty$. But as $P(A_n\cap B_n) = P(A_n) + P(B_n) - P(A_n\cup B_n)$, we conclude that

$$P(A_n\cap B_n) \underset{n\to\infty}{\longrightarrow} 1.$$

Since $A_n\cap B_n \subseteq \left\{\sup_{0\le s\le t}\left|(X_nY_n)_s - (XY)_s\right| \le \varepsilon\right\}$, we obtain

$$P\left(\left\{\sup_{0\le s\le t}\left|(X_nY_n)_s - (XY)_s\right| \le \varepsilon\right\}\right) \underset{n\to\infty}{\longrightarrow} 1. \qquad \blacktriangledown$$

Proof of Proposition 2.33.

We now proceed with the proof of the proposition. We will start by using Whitney's Embedding Theorem and the remarks in [E89, §7.7] to visualize M as an embedded submanifold of $\mathbb{R}^p$, for some $p \in \mathbb{N}$, and to write down our Stratonovich integrals as *real* Stratonovich integrals. Indeed, there exists a family of functions $\left\{h^1, ..., h^p\right\} \subset C^\infty\left(\mathbb{R}^p\right)$ such that, in the embedded picture, the one form α can be written as $\alpha = \sum_{j=1}^p Z_j \mathbf{d} h^j$, where $Z_j \in C^\infty\left(\mathbb{R}^p\right)$ for $j \in \{1, ..., p\}$. Therefore, using the properties of the Stratonovich integral (see [E89, Proposition 7.4]),

$$\begin{aligned} &\frac{1}{s}\left[\int \left\langle \alpha, \delta \Sigma^s \right\rangle - \int \left\langle \alpha, \delta \Gamma^{\tau_K} \right\rangle\right] \\ &= \sum_{j=1}^p \frac{1}{s}\left[\int Z_j\left(\Sigma^s\right)\delta\left(h^j\left(\Sigma^s\right)\right) - \int Z_j\left(\Gamma^{\tau_K}\right)\delta\left(h^j\left(\Gamma^{\tau_K}\right)\right)\right]. \end{aligned} \tag{2.96}$$

Adding and subtracting the term $\sum_{j=1}^p \int Z_j\left(\Sigma^s\right)\delta h^j\left(\Gamma^{\tau_K}\right)$ in the right hand side of (2.96), we have

$$\begin{aligned} \frac{1}{s}\left[\int \left\langle \alpha, \delta \Sigma^s \right\rangle - \int \left\langle \alpha, \delta \Gamma^{\tau_K} \right\rangle\right] &= \sum_{j=1}^p \underbrace{\frac{1}{s}\left[\int Z_j\left(\Sigma^s\right)\delta h^j\left(\Sigma^s\right) - \int Z_j\left(\Sigma^s\right)\delta h^j\left(\Gamma^{\tau_K}\right)\right]}_{(1)} \\ &+ \sum_{j=1}^p \underbrace{\frac{1}{s}\left[\int \left(Z_j\left(\Sigma^s\right) - Z_j\left(\Gamma^{\tau_K}\right)\right)\delta h^j\left(\Gamma^{\tau_K}\right)\right]}_{(2)}. \end{aligned} \tag{2.97}$$

We are going to study the terms (1) and (2) separately. We start by considering

$$\sigma_n = \left\{0 = T_0^n \leq T_1^n \leq \ldots \leq T_{k_n}^n < \infty\right\},$$

a sequence of random partitions that tends to the identity (in the sense of [P05, page 64]).

The expression (1):

We want to study the *ucp* convergence of $\frac{1}{s}\left[\int Z_j\left(\Sigma^s\right)\delta h^j\left(\Sigma^s\right) - \int Z_j\left(\Sigma^s\right)\delta h^j\left(\Gamma^{\tau_K}\right)\right]$ as $s \to 0$. Define

$$\begin{aligned} x_n(s) :=& \frac{1}{s}\left(\sum_{i=0}^{k_n-1}\frac{1}{2}\left(Z_j\left(\Sigma^s\right)_{T_{i+1}^n} + Z_j\left(\Sigma^s\right)_{T_i^n}\right)\left(h^j\left(\Sigma^s\right)^{T_{i+1}^n} - h^j\left(\Sigma^s\right)^{T_i^n}\right)\right. \\ &\left. - \sum_{i=0}^{k_n-1}\frac{1}{2}\left(Z_j\left(\Sigma^s\right)_{T_{i+1}^n} + Z_j\left(\Sigma^s\right)_{T_i^n}\right)\left(h^j\left(\Gamma^{\tau_K}\right)^{T_{i+1}^n} - h^j\left(\Gamma^{\tau_K}\right)^{T_i^n}\right)\right) \\ =& \sum_{i=0}^{k_n-1}\frac{1}{2}\left(Z_j\left(\Sigma^s\right)_{T_{i+1}^n} + Z_j\left(\Sigma^s\right)_{T_i^n}\right)\left(\frac{h^j\left(\Sigma^s\right)^{T_{i+1}^n} - h^j\left(\Gamma^{\tau_K}\right)^{T_{i+1}^n}}{s}\right. \\ &\left. - \frac{h^j\left(\Sigma^s\right)^{T_i^n} - h^j\left(\Gamma^{\tau_K}\right)^{T_i^n}}{s}\right). \end{aligned}$$

which corresponds to the discretization of the Stratonovich integrals $\frac{1}{s}\left[\int Z_j\left(\Sigma^s\right)\delta h^j\left(\Sigma^s\right) - \int Z_j\left(\Sigma^s\right)\delta h^j\left(\Gamma^{\tau_K}\right)\right]$ using the random partitions of σ_n. Indeed, by [P05, Corollary 1, page

291],

$$x_n(s) \underset{n\to\infty}{\underset{ucp}{\longrightarrow}} \frac{1}{s}\left[\int Z_j(\Sigma^s)\,\delta h^j(\Sigma^s) - \int Z_j(\Sigma^s)\,\delta h^j(\Gamma^{\tau_K})\right].$$

On the other hand, as $T_i^n < \infty$ a.s. for any $i \in \{1, ..., k_n\}$, part **(i)** in Definition 2.31 and Lemma A.3 imply that

$$\frac{1}{2}\left(Z_j(\Sigma^s)_{T_{i+1}^n} + Z_j(\Sigma^s)_{T_i^n}\right) \underset{s\to 0}{\underset{ucp}{\longrightarrow}} \frac{1}{2}\left(Z_j(\Gamma^{\tau_K})_{T_{i+1}^n} + Z_j(\Gamma^{\tau_K})_{T_i^n}\right)$$

The convergence above is in probability but, for convenience, we prefer to regard these random variables as trivial processes. Furthermore, part **(ii)** in Definition 2.31 and Lemma A.4 imply that

$$\frac{h^j(\Sigma^s)^{T_{i+1}^n} - h^j(\Gamma^{\tau_K})^{T_{i+1}^n}}{s} = \left(\frac{h^j(\Sigma^s) - h^j(\Gamma^{\tau_K})}{s}\right)^{T_{i+1}^n} \underset{s\to 0}{\underset{ucp}{\longrightarrow}} Y\left[h^j\right]^{T_{i+1}^n},$$

$$\frac{h^j(\Sigma^s)^{T_i^n} - h^j(\Gamma^{\tau_K})^{T_i^n}}{s} = \left(\frac{h^j(\Sigma^s) - h^j(\Gamma^{\tau_K})}{s}\right)^{T_i^n} \underset{s\to 0}{\underset{ucp}{\longrightarrow}} Y\left[h^j\right]^{T_i^n}$$

Now, since by hypothesis Σ and Y are bounded then so are $\frac{1}{2}\left(Z_j(\Sigma^s)_{T_{i+1}^n} + Z_j(\Sigma^s)_{T_i^n}\right)$ and $Y\left[h^j\right] = \mathbf{i}_Y \mathbf{d}h^j$ ($\mathbf{d}h^j$ is only evaluated on the compact K since Y is a vector field over Γ^{τ_K}) and hence by Lemma 2.48

$$x_n(s) \underset{s\to 0}{\underset{ucp}{\longrightarrow}} \sum_{i=0}^{k_{n-1}} \frac{1}{2}\left(Z_j(\Gamma^{\tau_K})_{T_{i+1}^n} + Z_j(\Gamma^{\tau_K})_{T_i^n}\right)\left(Y\left[h^j\right]^{T_{i+1}^n} - Y\left[h^j\right]^{T_i^n}\right) =: x_n^*$$

In addition, by [P05, Corollary 1, page 291],

$$x_n^* \underset{n\to\infty}{\underset{ucp}{\longrightarrow}} \int Z_j(\Gamma^{\tau_K})\,\delta\left(Y\left[h^j\right]\right).$$

Hence, by Lemma 2.47 we conclude that

$$\frac{1}{s}\left[\int Z_j(\Sigma^s)\,\delta h^j(\Sigma^s) - \int Z_j(\Sigma^s)\,\delta h^j(\Gamma^{\tau_K})\right] \underset{s\to 0}{\underset{ucp}{\longrightarrow}} \int Z_j(\Gamma^{\tau_K})\,\delta\left(Y\left[h^j\right]\right). \tag{2.98}$$

The expression (2):

We want to study now the *ucp* convergence of $\frac{1}{s}\int\left(Z_j\left(\Sigma^s\right)-Z_j\left(\Gamma^{\tau_K}\right)\right)\delta h^j\left(\Gamma^{\tau_K}\right)$ as $s\to 0$. As in the previous paragraphs, we define

$$\begin{aligned}
y_n(s) := \frac{1}{s}\Bigg(&\sum_{i=0}^{k_n-1}\frac{1}{2}\left(Z_j\left(\Sigma^s\right)_{T_{i+1}^n}+Z_j\left(\Sigma^s\right)_{T_i^n}\right)\left(h^j\left(\Gamma^{\tau_K}\right)^{T_{i+1}^n}-h^j\left(\Gamma^{\tau_K}\right)^{T_i^n}\right)\\
&-\sum_{i=0}^{k_n-1}\frac{1}{2}\left(Z_j\left(\Gamma^{\tau_K}\right)_{T_{i+1}^n}+Z_j\left(\Gamma^{\tau_K}\right)_{T_i^n}\right)\left(h^j\left(\Gamma^{\tau_K}\right)^{T_{i+1}^n}-h^j\left(\Gamma^{\tau_K}\right)^{T_i^n}\right)\Bigg)\\
&=\sum_{i=0}^{k_n-1}\frac{1}{2}\Bigg(\frac{Z_j\left(\Sigma^s\right)_{T_{i+1}^n}-Z_j\left(\Gamma^{\tau_K}\right)_{T_{i+1}^n}}{s}\\
&\qquad+\frac{Z_j\left(\Sigma^s\right)_{T_i^n}-Z_j\left(\Gamma^{\tau_K}\right)_{T_i^n}}{s}\Bigg)\left(h^j\left(\Gamma^{\tau_K}\right)^{T_{i+1}^n}-h^j\left(\Gamma^{\tau_K}\right)^{T_i^n}\right)
\end{aligned}$$

as a discretization of the Stratonovich integral $\frac{1}{s}\int\left(Z_j\left(\Sigma^s\right)-Z_j\left(\Gamma^{\tau_K}\right)\right)\delta h^j\left(\Gamma^{\tau_K}\right)$ using σ_n. Then, by construction,

$$y_n(s)\underset{\substack{ucp\\ n\to\infty}}{\longrightarrow}\frac{1}{s}\int\left(Z_j\left(\Sigma^s\right)-Z_j\left(\Gamma^{\tau_K}\right)\right)\delta h^j\left(\Gamma^{\tau_K}\right).$$

On the other hand, invoking Definition 2.31 and Lemma A.3 we have that

$$\begin{aligned}
\frac{Z_j\left(\Sigma^s\right)_{T_{i+1}^n}-Z_j\left(\Gamma^{\tau_K}\right)_{T_{i+1}^n}}{s}&=\left(\frac{Z_j\left(\Sigma^s\right)-Z_j\left(\Gamma^{\tau_K}\right)}{s}\right)_{T_{i+1}^n}\underset{\substack{ucp\\ s\to 0}}{\longrightarrow}Y\left[Z_j\right]_{T_{i+1}^n}\\
\frac{Z_j\left(\Sigma^s\right)_{T_i^n}-Z_j\left(\Gamma^{\tau_K}\right)_{T_i^n}}{s}&=\left(\frac{Z_j\left(\Sigma^s\right)-Z_j\left(\Gamma^{\tau_K}\right)}{s}\right)_{T_i^n}\underset{\substack{ucp\\ s\to 0}}{\longrightarrow}Y\left[Z_j\right]_{T_i^n}.
\end{aligned}$$

We now use again the boundedness of Σ and Y to guarantee the boundedness of $Y\left[Z_j\right]_{T_{i+1}^n}=\left(\mathbf{i}_Y\mathbf{d}Z_j\right)_{T_{i+1}^n}$ and $Y\left[Z_j\right]_{T_i^n}=\left(\mathbf{i}_Y\mathbf{d}Z_j\right)_{T_i^n}$ (notice that $\mathbf{d}Z_j$ is only evaluated on the compact set K because Y is a vector field over $\Gamma^{\tau_K}\subseteq K$). Therefore, by Lemma 2.48,

$$x_n(s)\underset{\substack{ucp\\ s\to 0}}{\longrightarrow}\sum_{i=0}^{k_n-1}\frac{1}{2}\left(Y\left[Z_j\right]_{T_{i+1}^n}+Y\left[Z_j\right]_{T_i^n}\right)\left(h^j\left(\Gamma^{\tau_K}\right)^{T_{i+1}^n}-h^j\left(\Gamma^{\tau_K}\right)^{T_i^n}\right):=x_n^*.$$

Additionally, the sequence $\{x_n^*\}_{n\in\mathbb{N}}$ obviously converge in *ucp* to $\int Y\left[Z_j\right]\delta\left(h^j\left(\Gamma^{\tau_K}\right)\right)$ as $n\to\infty$. Hence, by Lemma 2.47, we conclude that

$$\frac{1}{s}\left[\int\left(Z_j\left(\Sigma^s\right)-Z_j\left(\Gamma^{\tau_K}\right)\right)\delta h^j\left(\Gamma^{\tau_K}\right)\right]\underset{\substack{ucp\\ s\to 0}}{\longrightarrow}\int Y\left[Z_j\right]\delta\left(h^j\left(\Gamma^{\tau_K}\right)\right).\tag{2.99}$$

To sum up, if we substitute (2.98) and (2.99) in (2.97) we obtain that

$$\frac{1}{s}\left[\int\left\langle\alpha,\delta\Sigma^s\right\rangle-\int\left\langle\alpha,\delta\Gamma^{\tau_K}\right\rangle\right]\underset{\substack{ucp\\ s\to 0}}{\longrightarrow}\sum_{j=1}^{p}\int Z_j\left(\Gamma^{\tau_K}\right)\delta\left(Y\left[h^j\right]\right)+\int Y\left[Z_j\right]\delta\left(h^j\left(\Gamma^{\tau_K}\right)\right).$$

Using the integration by parts formula,

$$\begin{aligned}\int Z_j\left(\Gamma^{\tau_K}\right)\delta\left(Y\left[h^j\right]\right) &= Z_j\left(\Gamma^{\tau_K}\right)Y\left[h^j\right]-\left(Z_j\left(\Gamma^{\tau_K}\right)Y\left[h^j\right]\right)_{t=0}-\int Y\left[h^j\right]\delta\left(Z_j\left(\Gamma^{\tau_K}\right)\right)\\ &= \left\langle\alpha\left(\Gamma^{\tau_K}\right),Y\right\rangle-\left\langle\alpha\left(\Gamma^{\tau_K}\right),Y\right\rangle_{t=0}-\int Y\left[h^j\right]\delta\left(Z_j\left(\Gamma^{\tau_K}\right)\right)\end{aligned}$$

and, consequently,

$$\begin{aligned}\tfrac{1}{s}\left[\int\langle\alpha,\delta\Sigma^s\rangle-\int\left\langle\alpha,\delta\Gamma^{\tau_K}\right\rangle\right]\underset{\substack{ucp\\ s\to 0}}{\longrightarrow}&\int Y\left[Z_j\right]\delta\left(h^j\left(\Gamma^{\tau_K}\right)\right)-\int Y\left[h^j\right]\delta\left(Z_j\left(\Gamma^{\tau_K}\right)\right)\\ &+\left\langle\alpha\left(\Gamma^{\tau_K}\right),Y\right\rangle-\left\langle\alpha\left(\Gamma^{\tau_K}\right),Y\right\rangle_{t=0}.\end{aligned}$$

In order to conclude the proof, we claim that

$$\int Y\left[Z_j\right]\delta\left(h^j\left(\Gamma^{\tau_K}\right)\right)-\int Y\left[h^j\right]\delta\left(Z_j\left(\Gamma^{\tau_K}\right)\right)=\int\left\langle\mathbf{i}_Y\mathbf{d}\alpha,\delta\Gamma^{\tau_K}\right\rangle. \tag{2.100}$$

Indeed,

$$\mathbf{d}\alpha=\mathbf{d}\left(\sum_{j=1}^{p}Z_j\mathbf{d}h^j\right)=\sum_{j=1}^{p}\mathbf{d}Z_j\wedge\mathbf{d}h^j,\quad\text{and}\quad\mathbf{i}_Y\mathbf{d}\alpha=\sum_{j=1}^{p}\left(Y\left[Z_j\right]\mathbf{d}h^j-Y\left[h^j\right]\mathbf{d}Z_j\right)$$

which proofs (2.100), as required. ■

3
Reduction and reconstruction of symmetric SDEs

Symmetries have historically played a role of paramount importance in the study of dynamical systems in general (see [GS85, GS02, ChL00], and references therein) and of physical, mechanical, and Hamiltonian systems in particular (see for instance [AM78, MR99, OR04] for general presentations of the subject, historical overviews, and references). The presence of symmetries in a system usually brings in its wake the occurrence of degeneracies, conservation laws, and invariance properties that can be used to simplify or *reduce* the system and hence its analysis. In trying to pursue this strategy, researchers have developed powerful mathematical tools that optimize the benefit of this approach in specific situations.

The impressive volume of work that has been done in this field over the centuries does not have a counterpart in the context of stochastic dynamics, probably because most symmetry based mathematical tools are formulated using global analysis and Lie theory in an essential way, and this machinery has been adapted to the stochastic context relatively recently [M81, M82, S82, E82, E89]. As we will show in this chapter, *most of the symmetry based techniques available for dynamical systems can be formulated and taken advantage of when studying stochastic differential equations*.

In a first approach, symmetry based techniques can be roughly grouped into two separate procedures, namely, *reduction* and *reconstruction*. Reduction is explicitly implemented by combining the restriction of the system to dynamically invariant submanifolds whose existence is implied by its symmetries and by eliminating the remaining symmetry degeneracies through projection to an appropriate orbit space. Even if the space in which the system is originally formulated is Euclidean, the resulting reduced space is most of the time a non-Euclidean manifold hence showing the importance of global analysis in this context. The reduction procedure yields a dimensionally smaller space in which the symmetry degeneracies have been eliminated and that should, in principle, be easier to study; in the stochastic context, reduction has the added value of being able in some instances to isolate the non-stochastic part of the dynamics (see the example on collective motion in Subsection 3.6.1).

If once the reduced system has been solved we want to come back to the original one, we need to reconstruct the reduced solutions. In practice, this is obtained by horizontally lifting the reduced motion using a connection and then correcting the result with a curve in the group that satisfies a certain first order differential equation. The strategy of combining reduction and reconstruction in the search for the solutions of a symmetric dynamical system, splits the task into two parts, which most of the time simplifies greatly the problem.

Another approach used to take advantage of the symmetries of a problem consists of using the Slice Theorem [P61] and the tangent-normal decomposition [K90, F91] available for proper group actions to locally split the dynamics into a direction tangent to the group orbits and another one transversal to them. We will see that this tool, that is used in a standard fashion in the context of deterministic equivariant dynamics and equivariant bifurcation theory, yields in the stochastic case *skew-product splittings* that have already been extensively studied in the equivariant diffusions literature (see for instance [PR88, L89, T92], and references therein) to construct decompositions of the associated second order differential operators.

It must be noticed that the mathematical value of the results obtained with the two approaches that we just briefly discussed, that is, the one based on reduction-reconstruction and the one based on the tangent-normal decomposition, is morally the same. However, there are important technical conditions that make them different and preferable over one another in different specific situations:

(i) The reduction-reconstruction technique uses very strongly the orbit space of the symmetry group in question; this space could be geometrically convoluted and we may need to use only its strata if we want to face regular quotient manifolds where the standard calculus on manifolds is valid. The main advantage of this technique is that it yields *global results.*

(ii) The use of the Slice Theorem and the tangent-normal decomposition makes unnecessary the use of quotient manifolds and the entire analysis takes place in the original manifold. However, the results obtained are *local and are limited to a tubular neighborhood of the orbits.*

In this chapter we show how the symmetries of stochastic differential equations can be used by implementing techniques similar to those available for their deterministic counterparts. We start in Section 3.1 by introducing the notion of group of symmetries of a stochastic differential equation and by studying the associated invariant submanifolds as well as the implied degeneracies in the solutions. The reduction and reconstruction procedures are presented in Section 3.2; reconstruction is carried out using the horizontal lifts for semimartingales introduced in [S82, C01] and references therein.

The skew-product decomposition of second order differential operators is a factorization technique that has been used in the stochastic processes literature in order to split the semielliptic and, in particular, the diffusion operators, associated to certain stochastic differential equations (see, for instance, [PR88, L89, T92], and references therein). This splitting has important consequences as to the properties of the solutions of these equations, like certain factorization properties of their probability laws and of the associated stochastic flows. In Section 3.3

we show that symmetries are a natural way to obtain this kind of decompositions. Our work extends the existing results in two ways: first, we generalize the notion of skew-product to arbitrary stochastic differential equations by working with the notion of skew-product decomposition of the Stratonovich operator. Obviously, our approach coincides with the traditional one in the case of diffusions. Second, we use the Slice Theorem [P61] and the tangent-normal decomposition [K90, F91] to construct local skew-product decompositions in the presence of arbitrary proper symmetries (not necessarily free) in a neighborhood of any point in the open and dense principal orbit type. This result generalizes the skew-product decompositions presented in [ELL04] for regular free actions. Section 3.4 studies stochastic differential equations on associated bundles; in this situation the local skew-product splitting induced by the Slice Theorem is globally available.

Section 3.5 is dedicated to reduction and reconstruction of the stochastic Hamiltonian systems introduced in Chapter 2. It is worth mentioning that, as it was already the case for deterministic Hamiltonian systems, stochastic Hamiltonian systems are stable with respect to symplectic and Poisson reduction; in short, the reduction of a stochastic Hamiltonian system is again a stochastic Hamiltonian system. In Section 3.6 we present several (Hamiltonian) examples. The first one (Subsection 3.6.1) has to do with deterministic systems in which a stochastic perturbation is added using the conserved quantities associated to the symmetry (collective perturbation); such systems share the remarkable feature that symplectic reduction eliminates the stochastic part of the equation making the reduced system deterministic. In Subsection 3.6.2 we study the symmetries of stochastic mechanical systems on the cotangent bundles of Lie groups. In this situation, the reduction and reconstruction equations can be written down in a particularly explicit fashion that has to do with the Lie-Poisson structure in the dual of the Lie algebra of the group in question. A particular case of this is presented in Subsection 3.6.3 where we analyze two different stochastic perturbations of the free rigid body: one of them models the dynamics of a free rigid body subjected to small random impacts and the other one an "unbolted" rigid body that is not completely rigid.

The content of this chapter is a transcription of the paper [LO08] written by the author of this thesis in collaboration with Juan Pablo Ortega.

3.1 Symmetries and conservation laws of stochastic differential equations

Let M and N be two finite dimensional manifolds and let $(\Omega, \mathcal{F}, \{\mathcal{F}_t \,|\, t \geq 0\}, P)$ be a filtered probability space. Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale. According to Subsection 1.4.4, a Stratonovich operator from N to M is a family $\{S(x,y)\}_{x \in N, y \in M}$ such that $S(x,y) : T_xN \to T_yM$ is a linear mapping that depends smoothly on its two entries. We recall that a M-valued semimartingale Γ is a solution of the the Stratonovich stochastic differential equation

$$\delta\Gamma = S(X, \Gamma)\delta X \tag{1.53}$$

associated to X and S, if for any $\alpha \in \Omega(M)$, the following equality between Stratonovich integrals holds:

$$\int \langle \alpha, \delta\Gamma \rangle = \int \langle S^*(X,\Gamma)\alpha, \delta X \rangle.$$

If we prefer it, differential equations can be formulated using Itô integration by associating a natural Schwartz operator $\mathcal{S} : \tau_x N \to \tau_y M$ on the second order tangent bundles, to the Stratonovich operator S.

Definition 3.1 *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale and let $S : TN \times M \to TM$ be a Stratonovich operator. Let $\phi : M \to M$ be a diffeomorphism. We say that ϕ is a **symmetry** of the stochastic differential equation (1.53) if for any $x \in N$ and $y \in M$*

$$S\left(x, \phi\left(y\right)\right) = T_y\phi \circ S\left(x, y\right). \tag{3.1}$$

As it was already the case in standard deterministic context, the symmetries of a stochastic differential equation imply degeneracies at the level of its solutions, as we spell out in the following proposition.

Proposition 3.2 *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale, $S : TN \times M \to TM$ a Stratonovich operator, and let $\phi : M \to M$ be a symmetry of the corresponding stochastic differential equation (1.53). If Γ is solution of (1.53) then so is $\phi\left(\Gamma\right)$.*

Proof. Let Γ be a solution of (1.53). We need to show that for any $\alpha \in \Omega\left(M\right)$,

$$\int \langle \alpha, \delta\phi\left(\Gamma\right) \rangle = \int \langle S^*\left(X, \phi\left(\Gamma\right)\right)\alpha, \delta X \rangle.$$

Since ϕ is a diffeomorphism, $\int \langle \alpha, \delta\phi\left(\Gamma\right)\rangle = \int \langle \phi^*\alpha, \Gamma \rangle$ (see, for instance, [E89, §7.5]). Now, since Γ is a solution of (1.53), $\int \langle \phi^*\alpha, \Gamma \rangle = \int \langle S^*\left(X, \Gamma\right)\left(\phi^*\alpha\right), \delta X \rangle$. Since ϕ is a symmetry, we have that $S^*\left(x, \phi\left(y\right)\right) = S^*\left(x, y\right) \circ T_y^*\phi$, for any $x \in N$, $y \in M$ and hence,

$$\int \langle \phi^*\alpha, \Gamma \rangle = \int \langle S^*\left(X, \Gamma\right)\left(\phi^*\alpha\right), \delta X \rangle = \int \langle S^*\left(X, \phi\left(\Gamma\right)\right)\left(\alpha\right), \delta X \rangle,$$

which shows that $\phi\left(\Gamma\right)$ is a solution of (1.53). ■

The symmetries that we are mostly interested in are induced by the action of a Lie group G on the manifold M via the map $\Phi : G \times M \to M$. Given $(g, z) \in G \times M$, we will usually write $g \cdot z$ to denote $\Phi\left(g, z\right)$. We also introduce the maps

$$\begin{array}{rccc} \Phi_z : G & \longrightarrow & M \\ g & \longmapsto & g \cdot z \end{array}, \qquad \begin{array}{rccc} \Phi_g : M & \longrightarrow & M \\ z & \longmapsto & g \cdot z \end{array}.$$

The Lie algebra of G will be usually denoted by $\mathfrak{g}$ and we will write the tangent space to the orbit $G \cdot m$ that contains $m \in M$ as $\mathfrak{g} \cdot m := T_m(G \cdot m)$.

Definition 3.3 *We will say that the stochastic differential equation (1.53) is G-**invariant** if, for any $g \in G$, the diffeomorphism $\Phi_g : M \to M$ is a symmetry in the sense of Definition 3.1. In this situation we will also say that the Stratonovich operator S is G-**invariant**.*

Remark 3.4 Given a solution Γ of a G-invariant stochastic differential equation, Proposition 3.2 provides an entire orbit of solutions since for any $g \in G$, the semimartingale $\Phi_g(\Gamma)$ is also a solution. This degeneracy has also a reflection in the probability laws of the solutions in a form that we spell out in the following lines. Let $\Gamma : \{0 \leq t < \zeta\} \to M$ be a solution of the G-invariant system (M, S, X, N) defined up to the explosion time ζ, which may be finite if M is not compact. In such case, Γ can be actually understood as a process that takes values in the Alexandroff one-point compactification $\hat{M} := M \cup \{\infty\}$ of M and it is hence defined in the whole space $\mathbb{R}_+ \times \Omega$ ([IW89, Chapter V]). In this picture, the process Γ is continuous and with the property that $\Gamma_t(\omega) = \{\infty\}$, for any $(t, \omega) \in \mathbb{R}_+ \times \Omega$ such that $t \geq \zeta(\omega)$.

Let now $\hat{W}(M)$ be the path space defined by

$$\hat{W}(M) = \{w : [0, \infty] \to \hat{M} \text{ continuous such that } w(0) \in M \text{ and} \\ \text{if } w(t) = \{\infty\} \text{ then } w(t') = \{\infty\} \text{ for any } t' \geq t\}.$$

Let $\{P_z \mid z \in M\}$ be the family of probability measures on $\hat{W}(M)$ defined by the solutions of (M, S, X, N), that is, P_z is the law of the random variable $\Gamma^z : \Omega \to \hat{W}(M)$, where Γ^z is the solution of (M, S, X, N) with initial condition $\Gamma^z_{t=0} = z$ a.s.. The action $\Phi : G \times M \to M$ may be extended to $\hat{M}$ just putting $\Phi_g(\{\infty\}) = \{\infty\}$ for any $g \in G$. Since $\Phi_g(\Gamma^z)$ is the unique solution of the system (M, S, X, N) with initial condition $g \cdot z$ by Proposition 3.2 then $P_{g \cdot z} = \Phi_g^* P_z$. More explicitly, for any measurable set $A \subset \hat{W}(M)$, $P_{g \cdot z}(A) = P_z(\Phi_g(A))$.

The equivariance property of the probabilities $\{P_z \mid z \in M\}$ can be found in [ELL04] formulated in the context of equivariant diffusions on principal bundles. In that setup, the authors replace the path space $\hat{W}(M)$ by $C(l, r, M) = \{\sigma : [l, r] \to M \mid \sigma \text{ is continuous}\}$, $0 \leq l < r < \infty$ and prove [ELL04, Theorem 2.5] that the probability laws $\{P_z^{l,r} \mid z \in M\}$ admit a factorization through probability kernels $\{P_z^{H,l,r} \mid z \in M\}$ from M to $C(l, r, M)$ and $\{Q_w^{l,r} \mid w \in C(l, r, M)\}$ from $C(l, r, M)$ to $C_e(l, r, G) = \{\sigma : [l, r] \to G \mid \sigma \text{ is continuous}, \sigma(l) = e\}$ such that

$$P_z^{l,r}(U) = \int \int \mathbf{1}_U(g \cdot w) Q_w^{l,r}(dg) P_z^{H,l,r}(dw)$$

for any Borel set $U \subseteq C(l, r, M)$. The proof of this fact uses a technique very close to the reduction-reconstruction scheme that we will introduce in the next section.

Apart from degeneracies, the presence of symmetry in a stochastic differential equation is also associated with the occurrence of conserved quantities and, more generally, with the appearance of invariant submanifolds.

Definition 3.5 *Let Γ be a solution of the stochastic differential equation (1.53) and let L be an injectively immersed submanifold of M. Let ζ be the maximal stopping time of Γ and suppose that $\Gamma_0(\omega) = Z_0$, where Z_0 is a random variable such that $Z_0(\omega) \in L$, for all $\omega \in \Omega$. We say that L is an **invariant submanifold** (respectively, a **locally invariant submanifold**) of the stochastic differential equation if for any stopping time $\tau < \zeta$ (respectively, if there exists a nontrivial stopping time $\zeta_L \leq \zeta$ such that for any stopping time $\tau < \zeta_L$) we have that $\Gamma_\tau \in L$.*

Proposition 3.6 *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale and let $S : TN \times M \to TM$ be a Stratonovich operator. Let L be an injectively immersed submanifold of M and suppose* *that the Stratonovich operator S is such that $\operatorname{Im}(S(x,y)) \subset T_yL$, for any $y \in L$ and any $x \in N$. Then, L is a locally invariant submanifold of the stochastic differential equation (1.53) associated to X and S. If L is closed in M, then it is an invariant submanifold.*

Proof. The proof is similar to the proof of Proposition 2.9. By hypothesis, the Stratonovich operator $S : TN \times M \to TM$ induces another Stratonovich operator $S_L : TN \times L \to TL$, obtained from S by restriction. It is clear that if $i : L \hookrightarrow M$ is the inclusion then

$$S_L^*(x,y) \circ T_y^* i = S^*(x,y), \tag{3.2}$$

for any $x \in N$ and $y \in L$. Let Γ_L be the semimartingale in L that is a solution of the Stratonovich stochastic differential equation

$$\delta\Gamma_L = S_L(X, \Gamma_L)\delta X \tag{3.3}$$

with initial condition Γ_0 in L. We now show that $\overline{\Gamma} := i \circ \Gamma_L$ is a solution of

$$\delta\overline{\Gamma} = S(X, \overline{\Gamma})\delta X.$$

Since the maximal stopping times ζ_L and ζ of, respectively, Γ_L and of the stochastic differential equation associated to S with the same initial condition, are such that $\zeta_L \leq \zeta$, this will prove the statement. Indeed, for any $\alpha \in \Omega(M)$,

$$\int \langle \alpha, \delta\overline{\Gamma} \rangle = \int \langle \alpha, \delta(i \circ \Gamma_L) \rangle = \int \langle i^*\alpha, \delta\Gamma_L \rangle.$$

Since Γ_L satisfies (3.3) and $i^*\alpha \in \Omega(L)$, by (3.2) this equals

$$\int \langle S_L^*(X, \Gamma_L)(i^*\alpha), \delta X \rangle = \int \langle S^*(X, i \circ \Gamma_L)(\alpha), \delta X \rangle = \int \langle S^*(X, \overline{\Gamma})(\alpha), \delta X \rangle,$$

that is, $\delta\overline{\Gamma} = S(X, \overline{\Gamma})\delta X$, as required. When L is closed in M, one can show (see [E82, Theorem 3 page 123]) that $\zeta_L = \zeta$ and hence L is an invariant submanifold. ■

We now use Proposition 3.6 to show that the invariant manifolds that can be associated to deterministic symmetric systems are also available in the stochastic context. Let M be a manifold acted properly upon by a Lie group G via the map $\Phi : G \times M \to M$. We recall that the action Φ is said to be proper when for any two convergent sequences $\{m_n\}$ and $\{g_n \cdot m_n := \Phi(g_n, m_n)\}$ in M, there exists a convergent subsequence $\{g_{n_k}\}$ in G. The properness hypothesis on the action implies implies that most of the useful features that compact group actions have, are still available. For example, proper group actions admit local slices, the isotropy subgroups are always compact, and (the connected components of) the ***isotropy type*** submanifolds defined by $M_I := \{z \in M \mid G_z = I\}$, are embedded submanifolds of M for any isotropy subgroup $I \subset G$ of the action.

Proposition 3.7 (Law of conservation of the isotropy) *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale and let $S : TN \times M \to TM$ be a Stratonovich operator that is invariant with respect to a proper action of the Lie group G on the manifold M. Then, for any isotropy subgroup $I \subset G$, the isotropy type submanifolds M_I are invariant submanifolds of the stochastic differential equation associated to S and X.*

Proof. The properness of the action guarantees that for any isotropy subgroup $I \subset G$ and any $z \in M_I$,

$$T_z M_I = (T_z M)^I := \{v \in T_z M \mid T_z\Phi_g \cdot v = v, \text{ for any } g \in I\}. \tag{3.4}$$

Hence, for any $z \in M_I$ and $g \in I$, the G-invariance of the Stratonovich operator S implies that

$$T_z\Phi_g \circ S(x, z) = S(x, g \cdot z) = S(x, z),$$

which by (3.4) implies that $\operatorname{Im}(S(x, z)) \subset T_z M_I$. The invariance of the isotropy type manifolds follows then from Proposition 3.6 as well as from the equivariance of the stochastic flow associated to the stochastic differential equation determined by S and X. More explicitly, let $F : [0, \zeta) \times M \to M$ be the stochastic flow associated to the stochastic differential equation determined by S and X; by definition, $F_t(z)$ is the solution semimartingale Γ with initial condition $\Gamma_0(\omega) = z$ a.s., $z \in M$. The invariance of S implies that the flow F is such that $F_t(g \cdot z) = g \cdot F_t(z)$, for any $z \in M$ and $g \in G$, as it can be checked from the proof of Proposition 3.2. This equality guarantees that the isotropy subgroups of z and of $F_t(z)$ coincide, for any t. Consequently, if $z \in M_I$ then $F_t(z) \in M_I$, as required. ■

Remark 3.8 Some of the results that we just stated and others that will appear later on in the paper could be easily proved using their deterministic counterparts and the so called *Malliavin's Transfer Principle* [Ma78] which says, roughly speaking, that results from the theory of ordinary differential equations are valid for stochastic differential equations in Stratonovich form. The unavailability of a metatheorem that explicitly proves and shows the range of applicability of this principle makes advisable its use with care.

3.2 Reduction and reconstruction

This section is the core of the chapter. In the preceding paragraphs we explained how the symmetries of a stochastic differential equation imply the existence of certain conservation laws and degeneracies; reduction is a natural procedure to take advantage of the former and eliminate the latter via a combination of restriction and passage to the quotient operations. The end result of this strategy is the formulation of a stochastic differential equation with the same noise semimartingale but whose solutions take values in a manifold that is dimensionally smaller than the original one, which justifies the term *reduction* when we refer to this process. Smaller dimension and the absence of symmetry induced degeneracies usually make the reduced stochastic differential equation more tractable and easier to solve. The gain is therefore clear if once we have found the solutions of the reduced system, we know how to use them to find the solutions of the original system; that task is feasible and is the *reconstruction* process that will be explained in the second part of this section.

Theorem 3.9 (Reduction Theorem) *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale and let $S : TN \times M \to TM$ be a Stratonovich operator that is invariant with respect to a proper action of the Lie group G on the manifold M. Let $I \subset G$ be an isotropy subgroup of the G-action on M, M_I the corresponding isotropy type submanifold, and $L_I := N(I)/I$,* *with $N(I) := \{g \in G \mid gIg^{-1} = I\}$ the normalizer of I in G. L_I acts freely and properly on M_I and hence the orbit space M_I/L_I is a regular quotient manifold, that is, the projection $\pi_I : M_I \to M_I/L_I$ is a surjective submersion. Moreover, there is a well defined Stratonovich operator $S_{M_I/L_I} : TN \times M_I/L_I \to T(M_I/L_I)$ given by*

$$S_{M_I/L_I}(x, \pi_I(z)) = T_z\pi_I(S(x,z)), \quad \text{for any } x \in N \text{ and } z \in M_I \tag{3.5}$$

such that if Γ is a solution semimartingale of the stochastic differential equation associated to S and X, with initial condition $\Gamma_0 \subset M_I$, then so is $\Gamma_{M_I/L_I} := \pi_I(\Gamma)$ with respect to S_{M_I/L_I} and X, with initial condition $\pi_I(\Gamma_0)$. We will refer to S_{M_I/L_I} as the ***reduced Stratonovich operator*** *and to Γ_{M_I/L_I} as the* ***reduced solution****.*

Proof. The statement about M_I/L_I being a regular quotient manifold is a standard fact about proper group actions on manifolds (see for instance [DK99]). Now, observe that $S_{M_I/L_I} : TN \times M_I/L_I \to T(M_I/L_I)$ is well defined: if $z_1, z_2 \in M_I$ are such that $\pi_I(z_1) = \pi_I(z_2)$, then there exists some $g \in L_I$ satisfying $z_2 = \Phi_g(z_1)$ (we use the same symbol Φ to denote the G-action on M and the induced L_I-action on M_I). Hence,

$$\begin{aligned} S_{M_I/L_I}(x, \pi_I(z_2)) &= T_{z_2}\pi_I \circ S(x, z_2) = T_{z_2}\pi_I \circ T_{z_1}\Phi_g \circ S(x, z_1) \\ &= T_{z_1}\pi_I \circ S(x, z_1) = S_{M_I/L_I}(x, \pi_I(z_1)), \end{aligned}$$

where the G-invariance of S has been used. Let now Γ be a solution semimartingale of the stochastic differential equation associated to S and X with initial condition $\Gamma_0 \subset M_I$. The G-invariance of S implies via Proposition 3.7 that $\Gamma \subset M_I$ and hence $\Gamma_{M_I/L_I} := \pi_I(\Gamma)$ is well defined. In order to prove the statement, we have to check that for any one-form $\alpha \in \Omega(M_I/L_I)$

$$\int \langle \alpha, \delta\Gamma_{M_I/L_I}\rangle = \int \langle S^*_{M_I/L_I}(X, \Gamma_{M_I/L_I})\alpha, \delta X\rangle.$$

This equality follows in a straightforward manner from (3.5). Indeed,

$$\begin{aligned} \int \langle \alpha, \delta\Gamma_{M_I/L_I}\rangle &= \int \langle \alpha, \delta(\pi_I \circ \Gamma)\rangle = \int \langle \pi_I^*\alpha, \delta\Gamma\rangle \\ &= \int \langle S^*(X,\Gamma)(\pi_I^*\alpha), \delta X\rangle = \int \langle S^*_{M_I/L_I}(X, \Gamma_{M_I/L_I})\alpha, \delta X\rangle, \end{aligned}$$

as required. ■

We are now going to carry out the reverse procedure, that is, given an isotropy subgroup $I \subset G$ and a solution semimartingale Γ_{M_I/L_I} of the reduced stochastic differential equation with Stratonovich operator S_{M_I/L_I} we will *reconstruct* a solution Γ of the initial stochastic differential equation with Stratonovich operator S. In order to keep the notation not too heavy we will assume in the rest of this section that the G-action on M is not only proper but also free, so that the only isotropy subgroup is the identity element e and hence there is only one isotropy type submanifold, namely $M_e = M$. The general case can be obtained by replacing in the following paragraphs M by the isotropy type manifolds M_I, and G by the groups L_I.

We now make our goal more precise. The freeness of the action $\Phi : G \times M \to M$ guarantees that the canonical projection $\pi : M \to M/G$ is a principal bundle with structural group G. We saw in the previous theorem that for any solution Γ of a stochastic differential equation associated to a G-invariant Stratonovich operator S and a N-valued noise semimartingale X, we can build a solution $\Gamma_{M/G} = \pi(\Gamma)$ of the reduced stochastic differential equation associated to the projected Stratonovich operator $S_{M/G}$ introduced in (3.5) and to the stochastic component X. The main goal of the paragraphs that follow is to show how to reconstruct the dynamics of the initial system from solutions $\Gamma_{M/G}$ of the reduced system. As we will see in Theorem 3.10, any solution Γ of the original stochastic differential equation may be written as $\Gamma = \Phi_{g^{\Xi}}(d)$ where $d : \mathbb{R}_+ \times \Omega \to M$ is a semimartingale such that $\pi(d) = \Gamma_{M/G}$ and $g^{\Xi} : \mathbb{R}_+ \times \Omega \to G$ is a G-valued semimartingale which satisfies a suitable stochastic differential equation on the group G.

We start by picking $A \in \Omega^1(M; \mathfrak{g})$ ($\mathfrak{g}$ is the Lie algebra of G) an auxiliary principal connection on the left principal G-bundle $\pi : M \to M/G$ and let $TM = \mathrm{Hor} \oplus \mathrm{Ver}$ be the decomposition of the tangent bundle TM into the Whitney sum of the horizontal and vertical bundles associated to A. Analogously, the cotangent bundle T^*M admits a decomposition $T^*M = \mathrm{Hor}^* \oplus \mathrm{Ver}^*$ where, by definition, $\mathrm{Hor}^*_z := (\mathrm{Ver}_z)^\circ$ is the annihilator of the vertical subspace Ver_z at a point $z \in M$ and $\mathrm{Ver}^*_z := (\mathrm{Hor}_z)^\circ$ is the annihilator of the horizontal subspace. Hence, any one form $\alpha \in \Omega(M)$ may be uniquely written as $\alpha = \alpha^H + \alpha^V$ with $\alpha^H \in \mathrm{Hor}^*$ and $\alpha^V \in \mathrm{Ver}^*$. A section of the bundle $\pi_M : T^*M \to M$ taking values in Hor^* is called a horizontal one form. It is called vertical if $\alpha_z \in \mathrm{Ver}^*_z$ for any $z \in M$.

Let $\Gamma_{M/G} \subset M_{M/G}$ be a solution of the reduced stochastic differential equation associated to the Stratonovich operator $S_{M/G}$, and with stochastic component $X : \mathbb{R}_+ \times \Omega \to V$ as in Theorem 3.9 . As we claimed, we are going to find a solution Γ to the original G-invariant stochastic differential equation associated to S, such that $\pi(\Gamma) = \Gamma_{M/G}$ with a given initial condition Γ_0. We start by horizontally lifting $\Gamma_{M/G}$ to a M-valued semimartingale d. Indeed, by [S82, Theorem 2.1] (see also [C01]), there exists a M-valued semimartingale $d : \mathbb{R}_+ \times \Omega \to M$ such that $d_0 = \Gamma_0$, $\pi(d) = \Gamma_{M/G}$ and that satisfies

$$\int \langle A, \delta d \rangle = 0, \tag{3.6}$$

where (3.6) is a $\mathfrak{g}$-valued integral. More specifically, let $\{\xi_1, ..., \xi_m\}$ be a basis of the Lie algebra $\mathfrak{g}$ and let $A(z) = \sum_{i=1}^m A^i(z)\,\xi_i$ the expression of A in this basis. Then

$$\int \langle A, \delta d \rangle := \sum_{i=1}^{m} \int \langle A^i, \delta d \rangle\, \xi_i. \tag{3.7}$$

The condition (3.6) is equivalent to $\int \langle \alpha, \delta d \rangle = 0$ for any vertical one-form $\alpha \in \Omega(M)$ (see [C01, page 1641]) which, in turn, implies

$$\int \langle \theta, \delta d \rangle = 0 \tag{3.8}$$

for any T^*M-valued process $\theta : \mathbb{R}_+ \times \Omega \to \mathrm{Ver}^* \subset T^*M$ over d. We want to find a G-valued semimartingale $g^{\Xi} : \Omega \times \mathbb{R}_+ \to G$ such that $g^{\Xi}_0 = e$ a.s. and $\Gamma = g^{\Xi} \cdot d$ is a solution of the stochastic differential equation associated to the Stratonovich operator S and the N-valued noise semimartingale X.

Let $g \in G$, $z \in M$. It is easy to see that

$$\ker\left(T_g^*\Phi_z\right) = \left(T_{g\cdot z}\left(G\cdot z\right)\right)^\circ = \left(\mathrm{Ver}_{g\cdot z}\right)^\circ = \mathrm{Hor}^*_{g\cdot z}. \tag{3.9}$$

Where $G \cdot z$ denotes the G-orbit that contains the point $z \in M$. Therefore, the map

$$\widetilde{T_g^*\Phi_z} := T_g^*\Phi_z\big|_{\mathrm{Ver}^*_{g\cdot z}} : T^*_{g\cdot z}M \cap \mathrm{Ver}^*_{g\cdot z} \longrightarrow T_g^*G \tag{3.10}$$

is an isomorphism. Let

$$\begin{array}{rcl} \rho\left(g,z\right) : T_g^*G & \longrightarrow & T^*_{g\cdot z}M \cap \mathrm{Ver}^*_{g\cdot z} \subset T^*_{g\cdot z}M \\ \alpha_g & \longmapsto & \left(\widetilde{T_g^*\Phi_z}\right)^{-1}\left(\alpha_g\right) \end{array}$$

and define $\psi^*\left(x,z,g\right) : T_g^*G \to T_x^*N$ by

$$\psi^*\left(x,z,g\right) = S^*\left(x, g\cdot z\right) \circ \rho\left(g,z\right).$$

Finally, we define a dual Stratonovich operator between the manifolds G and $M \times N$ as

$$\begin{array}{rcl} K^*\left(\left(z,x\right),g\right) : T_g^*G & \longrightarrow & T_z^*M \times T_x^*N \\ \alpha_g & \longmapsto & \left(0, \psi^*\left(x,z,g\right)\left(\alpha_g\right)\right). \end{array} \tag{3.11}$$

Theorem 3.10 (Reconstruction Theorem) *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale and let $S : TN \times M \to TM$ be a Stratonovich operator that is invariant with respect to a free and proper action of the Lie group G on the manifold M. If we are given $\Gamma_{M/G}$ a solution semimartingale of the reduced stochastic differential equation then $\Gamma = g^\Xi \cdot d$ is a solution of the original stochastic differential equation such that $\pi(\Gamma) = \Gamma_{M/G}$.*

In this statement, $d : \mathbb{R}_+ \times \Omega \to M$ is the horizontal lift of $\Gamma_{M/G}$ using an auxiliary principal connection on $\pi : M \to M/G$ such that $\Gamma_0 = d_0$, and $g^\Xi : \mathbb{R}_+ \times \Omega \to G$ is the semimartingale solution of the stochastic differential equation

$$\delta g^\Xi = K\left(\Xi, g\right)\delta\Xi \tag{3.12}$$

*with initial condition $g_0^\Xi = e$, K the Stratonovich operator introduced in (3.11), and stochastic component $\Xi = (d, X)$ We will refer to d as the **horizontal lift** of $\Gamma_{M/G}$ and to $\Gamma = g^\Xi$ as the **stochastic phase** of the reconstructed solution.*

Remark 3.11 As we already pointed out, Theorem 3.10 is also valid when the group action is not free. In that situation, one is given a solution of the reduced stochastic differential equation on the quotient M_I/L_I, with I an isotropy subgroup of the G-action on M. The correct statement (and the proof that follows) of the reconstruction theorem in this case can be obtained from the one that we just gave by replacing M by the isotropy type manifold M_I and G by the group L_I.

Proof of Theorem 3.10. In order to check that $\Gamma = g^{\Xi} \cdot d$ is a solution of the original stochastic differential equation we have to verify that for any $\alpha \in \Omega(M)$,

$$\int \langle \alpha, \delta\Gamma \rangle = \int \langle S^*(X, \Gamma)\alpha, \delta X \rangle. \tag{3.13}$$

Since $\Gamma = g^{\Xi} \cdot d = \Phi(g^{\Xi}, d)$, the statement in [S82, Lemma 3.4] allows us to write

$$\int \langle \alpha, \delta\Gamma \rangle = \int \left\langle \Phi^*_{g^{\Xi}} \alpha, \delta d \right\rangle + \int \left\langle \Phi^*_d \alpha, \delta g^{\Xi} \right\rangle. \tag{3.14}$$

We split the verification of (3.13) into two cases:

(i) $\alpha \in \Omega(M)$ is horizontal or, equivalently, $\alpha = \pi^*(\eta)$ with $\eta \in \Omega(M/G)$. Since α is horizontal, then $\Phi^*_d \alpha = 0$ by (3.9). Then, using (3.14),

$$\begin{aligned} \int \langle \alpha, \delta\Gamma \rangle &= \int \left\langle \Phi^*_{g^{\Xi}} \alpha, \delta d \right\rangle = \int \left\langle \Phi^*_{g^{\Xi}} \left(\pi^*(\eta)\right), \delta d \right\rangle \\ &= \int \left\langle \left(\pi \circ \Phi_{g^{\Xi}}\right)^*(\eta), \delta d \right\rangle = \int \langle \pi^*(\eta), \delta d \rangle = \int \left\langle \eta, \delta\Gamma_{M/G} \right\rangle. \end{aligned}$$

We recall that $\Gamma_{M/G} = \pi(d)$ is a solution of the reduced system, that is,

$$\int \left\langle \eta, \delta\Gamma_{M/G} \right\rangle = \int \left\langle S^*_{M/G}\left(X, \Gamma_{M/G}\right)(\eta), \delta X \right\rangle$$

for any $\eta \in \Omega(M/G)$. This implies by (3.5) that

$$\int \left\langle \eta, \delta\Gamma_{M/G} \right\rangle = \int \left\langle S^*_{M/G}\left(X, \Gamma_{M/G}\right)(\eta), \delta X \right\rangle = \int \langle S^*(X, d)(\pi^*(\eta)), \delta X \rangle.$$

Now, due to the G-invariance of S, we know that $S^*(x, g \cdot z) = S^*(x, z) \circ T^*_z \Phi_g$, for any $g \in G$, $x \in N$, $z \in M$. Recall also that $T_z \Phi_g$ sends the horizontal space Hor_z to $\mathrm{Hor}_{g \cdot z}$ and the vertical space Ver_z to $\mathrm{Ver}_{g \cdot z}$. Moreover, since α is horizontal, $\Phi^*_g \alpha = \alpha$ for any $g \in G$. Therefore,

$$\begin{aligned} \int \left\langle \eta, \delta\Gamma_{M/G} \right\rangle &= \int \langle S^*(X, d)(\alpha), \delta X \rangle = \int \left\langle S^*(X, d)\left(\Phi^*_{g^{\Xi}} \alpha\right), \delta X \right\rangle \\ &= \int \left\langle S^*\left(X, g^{\Xi} \cdot d\right)(\alpha), \delta X \right\rangle = \int \langle S^*(X, \Gamma)(\alpha), \delta X \rangle \end{aligned}$$

and hence (3.13) holds.

(ii) $\alpha \in \Omega(M)$ is vertical. Since α is vertical, so is $\Phi^*_{g^{\Xi}} \alpha$ as a T^*M-valued process. Therefore, $\int \left\langle \Phi^*_{g^{\Xi}} \alpha, \delta d \right\rangle = 0$ by (3.8). Thus, using (3.14),

$$\int \langle \alpha, \delta\Gamma \rangle = \int \left\langle \Phi^*_d \alpha, \delta g^{\Xi} \right\rangle.$$

Now, as g^{Ξ} is a solution of the stochastic differential equation (3.12),

$$\begin{aligned}\int \left\langle \Phi_d^*\alpha, \delta g^{\Xi}\right\rangle &= \int \left\langle K^*\left(\Xi, g^{\Xi}\right)\left(\Phi_d^*\alpha\right), \delta\Xi\right\rangle = \int \left\langle \left(0, \psi^*\left(g^{\Xi}, X, d\right)\left(\Phi_d^*\alpha\right)\right), \delta\Xi\right\rangle \\ &= \int \left\langle \psi^*\left(g^{\Xi}, X, d\right)\left(\Phi_d^*\alpha\right), \delta X\right\rangle . \end{aligned} \tag{3.15}$$

Recall that $\psi^*(x,z,g) = S^*(x, g\cdot z)\circ\rho(g,z)$. Moreover $\rho(g,z)\left(\gamma_g\right) = \left(\widetilde{T_g^*\Phi_z}\right)^{-1}\left(\gamma_g\right)$ for any $\gamma_g \in T_g^*G$. Hence,

$$\rho(g,z)\circ T_g^*\Phi_z\left(\alpha_{g\cdot z}\right) = \left(\widetilde{T_g^*\Phi_z}\right)^{-1}\left(T_g^*\Phi_z\left(\alpha_{g\cdot z}\right)\right) = \alpha_{g\cdot z}$$

for any $\alpha_{g\cdot z} \in T_{g\cdot z}^*M \cap \mathrm{Ver}_{g\cdot z}^*$, since in that situation $T_g^*\Phi_z\left(\alpha_{g\cdot z}\right) = \widetilde{T_g^*\Phi_z}\left(\alpha_{g\cdot z}\right)$. Therefore, expression (3.15) equals

$$\int \left\langle \psi^*\left(X, d, g^{\Xi}\right)\left(\Phi_d^*\alpha\right), \delta X\right\rangle = \int \left\langle S^*\left(X, g^{\Xi}\cdot d\right)(\alpha), \delta X\right\rangle = \int \left\langle S^*(X,\Gamma)(\alpha), \delta X\right\rangle,$$

and hence (3.13) also holds whenever $\alpha \in \Omega(M)$ is vertical, as required. ■

The stochastic phase g^{Ξ} introduced in the Reconstruction Theorem admits another characterization that we present in the paragraphs that follow. Let $\{\xi_1, ..., \xi_m\}$ be a basis of $\mathfrak{g}$, the Lie algebra of G and write $A = \sum_{i=1}^m A^i\xi_i$, where $A^i \in \Omega(M)$ are the components of the auxiliary connection $A \in \Omega^1(M;\mathfrak{g})$ in this basis. Consider the $\mathfrak{g}$-valued semimartingale

$$Y = \sum_{i=1}^m \int \left\langle S^*(X,d)\left(A^i\right), \delta X\right\rangle \xi_i. \tag{3.16}$$

Proposition 3.12 *Let $Y : \mathbb{R}_+\times\Omega \to \mathfrak{g}$ be the $\mathfrak{g}$-valued semimartingale defined in (3.16). Then, the stochastic phase $g^{\Xi} : \mathbb{R}_+\times\Omega \to G$ introduced in (3.12) is the unique solution of the stochastic differential equation*

$$\delta g = L(Y,g)\,\delta Y \tag{3.17}$$

associated to the Stratonovich operator L given by

$$\begin{array}{rcl} L(\xi,g) : T_\xi\mathfrak{g} & \longrightarrow & T_gG \\ \eta & \longmapsto & T_eL_g(\eta), \end{array}$$

with initial condition $g_0 = e$. The symbol $L_g : G \to G$ denotes the left translation map by $g \in G$.

In the proof of this proposition, we will denote by $\xi_M(z) := \frac{d}{dt}\big|_{t=0}\exp t\xi\cdot z$ the infinitesimal vector field associated to $\xi \in \mathfrak{g}$ by the G-action on M evaluated at $z \in M$. Analogously, we will write ξ_G for the infinitesimal generators of the G-action on itself by left translations. We recall (see [OR04] for a proof) that for any $g \in G$, $\xi \in \mathfrak{g}$, and $z \in M$,

$$T_z\Phi_g\left(\xi_M(z)\right) = \left(\mathrm{Ad}_g\xi\right)_M(g\cdot z). \tag{3.18}$$

Moreover, $T_g\Phi_z\left(\xi_G\left(g\right)\right)=T_z\Phi_g\left(\xi_M\left(z\right)\right)$ or, in other words,

$$\xi_G\left(g\right)=\widetilde{T_g\Phi_z}^{-1}\circ T_z\Phi_g\left(\xi_M\left(z\right)\right), \tag{3.19}$$

where $\widetilde{T_g\Phi_z}^{-1}:T_{g\cdot z}M\cap\mathrm{Ver}_{g\cdot z}\rightarrow T_gG$ is the isomorphism introduced in (3.10).

Proof of Proposition 3.12. A result in [S82] shows that in order to prove the statement it suffices to check that $\int\left\langle\theta,\delta g^{\Xi}\right\rangle=Y$, where θ is the canonical $\mathfrak{g}$-valued one form on G defined by $\theta_g\left(\xi_G\left(g\right)\right)=\xi$, for any $g\in G$ and $\xi\in\mathfrak{g}$. Indeed, Lemmas 3.2 and 3.3 in [S82] show that a G-valued semimartingale g^G is such that $\int\left\langle\theta,\delta g^G\right\rangle=Y$ if and only if g^G is a solution of (3.17). Now, suppose that g^{Ξ} is a solution of (3.12),

$$\int\left\langle\theta,\delta g^{\Xi}\right\rangle=\int\left\langle\psi^*\left(X,d,g^{\Xi}\right)\left(\theta\right),\delta X\right\rangle=\int\left\langle S^*\left(X,g^{\Xi}\cdot d\right)\circ\rho\left(g^{\Xi},d\right)\left(\theta\right),\delta X\right\rangle.$$

We are now going to verify that for any $g\in G$ and $z\in M$,

$$\rho\left(g,z\right)\left(\theta\right)=\left(\Phi^*_{g^{-1}}A\right)\left(g\cdot z\right). \tag{3.20}$$

First of all notice that as $\rho\left(g,z\right)\left(\gamma_g\right)=\left(\widetilde{T_g^*\Phi_z}\right)^{-1}\left(\gamma_g\right)\in T^*_{g\cdot z}M\cap\mathrm{Ver}^*_{g\cdot z}$, for any $\gamma_g\in T_g^*G$ and since A vanishes when acting on horizontal vector fields, it suffices to verify (3.20) when acting on vector fields of the form ξ_M, for some $\xi\in\mathfrak{g}$. Using (3.18), the right hand side of (3.20) then reads

$$\left(\Phi^*_{g^{-1}}A\right)\left(g\cdot z\right)\left(\xi_M\left(g\cdot z\right)\right)=A\left(z\right)\left(T_z\Phi_{g^{-1}}\left(\xi_M\left(g\cdot z\right)\right)\right)=A\left(z\right)\left(\left(\mathrm{Ad}_{g^{-1}}\xi\right)_M\left(z\right)\right)=\mathrm{Ad}_{g^{-1}}\xi.$$

As to the left hand side, we can write using (3.18) and (3.19),

$$\begin{aligned}\rho\left(g,z\right)\left(\theta(g)\right)\left(\xi_M\left(g\cdot z\right)\right)&=\left[\left(\widetilde{T_g^*\Phi_z}\right)^{-1}\theta(g)\right]\left(\xi_M\left(g\cdot z\right)\right)=\theta\left(g\right)\left[\widetilde{T_g\Phi_z}^{-1}\left(\xi_M\left(g\cdot z\right)\right)\right]\\&=\theta\left(g\right)\left[\widetilde{T_g\Phi_z}^{-1}\circ T_z\Phi_g\circ T_{g\cdot z}\Phi_{g^{-1}}\left(\xi_M\left(g\cdot z\right)\right)\right]\\&=\theta\left(g\right)\left[\widetilde{T_g\Phi_z}^{-1}\circ T_z\Phi_g\circ\left(\mathrm{Ad}_{g^{-1}}\xi\right)_M\left(z\right)\right]\\&=\theta\left(g\right)\left[\left(\mathrm{Ad}_{g^{-1}}\xi\right)_G\left(g\right)\right]=\mathrm{Ad}_{g^{-1}}\xi.\end{aligned}$$

Thus,

$$\int\left\langle S^*\left(X,g^{\Xi}\cdot d\right)\circ\rho\left(g^{\Xi},d\right)\left(\theta\right),\delta X\right\rangle=\int\left\langle S^*\left(X,g^{\Xi}\cdot d\right)\left(\Phi^*_{\left(g^{\Xi}\right)^{-1}}A\right),\delta X\right\rangle.$$

Now, since the Stratonovich operator S is G-invariant, we have that $S^*\left(x,g\cdot z\right)=S^*\left(x,z\right)\circ T_z^*\Phi_g$, for any $x\in N$, $z\in M$, and $g\in G$, and hence

$$S^*\left(x,g\cdot z\right)\left(\left(\Phi^*_{g^{-1}}A\right)\left(g\cdot z\right)\right)=S^*\left(x,z\right)\circ T_z^*\Phi_g\circ T^*_{g\cdot z}\Phi_{g^{-1}}\left(A\left(z\right)\right)=S^*\left(x,z\right)\left(A\left(z\right)\right).$$

Therefore,

$$\int \left\langle \theta, \delta g^{\Xi} \right\rangle = \int \left\langle S^{*}\left(X, g^{\Xi} \cdot d\right) \left(\Phi^{*}_{(g^{\Xi})^{-1}} A\right), \delta X \right\rangle = \int \left\langle S^{*}\left(X, d\right)\left(A\right), \delta X \right\rangle = Y,$$

and consequently g^{Ξ} solves (3.17). The argument that we just gave can be easily reversed to prove that if g^{Ξ} is a solution of (3.17) then it is also a solution of (3.12). ■

The combination of the reduction and the reconstruction of the solution semimartingales of a symmetric stochastic differential equation can be seen as a method to split the problem of finding its solutions into three simpler tasks which we summarize as follows:

Step 1: Find a solution $\Gamma_{M/G}$ for the reduced stochastic differential equation associated to the reduced Stratonovich operator $S_{M/G}$ on the dimensionally smaller space M/G.

Step 2: Take an auxiliary principal connection $A \in \Omega^1 (M; \mathfrak{g})$ for the principal bundle $\pi : M \to M/G$ and a horizontally lifted semimartingale $d : \mathbb{R}_+ \times \Omega \to M$, that is $\int \langle A, \delta d \rangle = 0$, such that $d_0 = \Gamma_0$ and $\pi(d) = \Gamma_{M/G}$.

Step 3: Let $g^{\Xi} : \mathbb{R}_+ \times \Omega \to G$ be the solution semimartingale of the stochastic differential equation (3.17) on G

$$\delta g = L\left(Y, g\right) \delta Y$$

with initial condition $g_0 = e$ a.s. and with noise semimartingale $Y = \int \langle S^{*}(X, d)(A), \delta X \rangle$. The solution of the original stochastic differential equation associated to the Stratonovich operator S with initial condition Γ_0 is then $\Gamma = \Phi_{g^{\Xi}}(d)$.

Remark 3.13 Theorem 3.10 has as a consequence that the maximal existence times ζ and $\zeta_{M/G}$ of π-related solutions Γ and $\Gamma_{M/G}$ of the original symmetric and reduced systems, coincide. Indeed, if we write $\Gamma_t = g_t \cdot d_t$, with d_t a horizontal lift of $\Gamma_{M/G}$, then first, d_t is defined up to the same (maybe finite) explosion time $\zeta_{M/G}$ of $\Gamma_{M/G}$. Second, as the semimartingale g_t is the solution of the left-invariant stochastic differential equation (3.17) then it is in principle stochastically complete ([E82, Chapter VII §6, Example (i) page 131]) if its stochastic forcing is. Since in our case, the stochastic component Y (3.16) depends on d_t, we can conclude that g_t is defined again on the stochastic interval $[0, \zeta_{M/G})$. We consequently conclude that the maximal existence time of the solutions of the initial symmetric system (M, S, X, N) coincides with that of the corresponding solutions of the reduced system $(M/G, S_{M/G}, X, N)$. Notice that this in particular implies that if the reduced manifold M/G is compact then all the solutions of the original symmetric system are defined for all time, even if M is not compact.

3.3 Symmetries and skew-product decompositions

The skew-product decomposition of second order differential operators is a factorization technique that has been used in the stochastic processes literature in order to split the semielliptic and, in particular, the diffusion operators, associated to certain stochastic differential equations (see, for instance, [PR88, L89, T92], and references therein). This splitting has important consequences as to the properties of the solutions of these equations, like certain factorization properties of their probability laws and of the associated stochastic flows.

Symmetries are a natural way to obtain this kind of decompositions as it has already been exploited in [ELL04]. Our goal in the following pages consists of generalizing the existing results in two ways: first, we will generalize the notion of skew-product to arbitrary stochastic differential equations by working with the notion of skew-product decomposition of the Stratonovich operator; we will indicate below how our approach coincides with the traditional one in the case of diffusions. Second, we will show that the skew-product decompositions presented in [ELL04] for regular free action are also available (at least locally) for singular proper group actions.

Definition 3.14 *Let N, M_1, and M_2 be three smooth manifolds and $S(x,m) : T_xN \to T_m(M_1 \times M_2)$, $x \in N$, $m = (m_1, m_2) \in M_1 \times M_2$, a Stratonovich operator from N to the product manifold $M_1 \times M_2$. We will say that S admits a* ***skew-product decomposition*** *if there exists a Stratonovich operator $S_2(x, m_2) : T_xN \longrightarrow T_{m_2}M_2$ from N to M_2 and a M_2-dependent Stratonovich operator $S_1(x, m_1, m_2) : T_xN \to T_{m_1}M_1$ such that*

$$S(x,m) = (S_1(x, m_1, m_2), S_2(x, m_2)) \in \mathcal{L}(T_xN, T_{m_1}M_1 \times T_{m_2}M_2)$$

for any $m = (m_1, m_2) \in M_1 \times M_2$. The operators S_1 and S_2 will be called the ***factors*** *of S.*

In order to show the relation between this definition and the classical one used in the papers that we just quoted, we first have to briefly recall the relation between the global Stratonovich and Itô formulations for the stochastic differential equations (see [E89] for a detailed presentation of this subject). Given M and N two manifolds, a Schwartz operator is a family of Schwartz maps (see [E89, Definition 6.22]) $\mathcal{S}(x,z) : \tau_xN \to \tau_zM$ between the tangent bundles of second order τN and τM. In this context, the Itô stochastic differential equation defined by the Schwartz operator $\mathcal{S}$ with stochastic component a continuous semimartingale $X : \mathbb{R}_+ \times \Omega \to N$ is

$$d\Gamma = \mathcal{S}(X, \Gamma)\, dX. \tag{3.21}$$

Given a Stratonovich operator S, we saw in Subsection 1.4.4 that there is a unique Schwartz operator $\mathcal{S} : \tau N \times M \to \tau M$ that is an extension of S to the tangent bundles of second order and which makes the Itô and Stratonovich stochastic differential equations associated to S and $\mathcal{S}$ equivalent, in the sense that they have the same semimartingale solutions.

It is easy to show that if $S : TN \times (M_1 \times M_2) \to T(M_1 \times M_2)$ is a Stratonovich operator that admits a skew-product decomposition with factors S_1 and S_2 then the equivalent Schwartz operator $\mathcal{S} : \tau N \times (M_1 \times M_2) \to \tau(M_1 \times M_2)$ can be written as

$$\mathcal{S}(x, (m_1, m_2)) = \mathcal{S}_1(x, m_1, m_2) + \mathcal{S}_2(x, m_2), \tag{3.22}$$

for any $x \in N$ and any $m = (m_1, m_2) \in M_1 \times M_2$. In this expression, $\mathcal{S}_1(x, m_1, m_2) : \tau_xN \to \tau_m(M_1 \times M_2)$ and $\mathcal{S}_2(x, m_2) : \tau_xN \to \tau_m(M_1 \times M_2)$ are the equivalent Schwartz operators of the Stratonovich operators $\widetilde{S}_1, \widetilde{S}_2 : TN \times (M_1 \times M_2) \to T(M_1 \times M_2)$ defined by $\widetilde{S}_1(x, m) :=$

$T_{m_1} i_{m_2}\left(S_1(x, m_1, m_2)\right)$ and $\widetilde{S}_2(x, m) := T_{m_2} i_{m_1}\left(S_2(x, m_2)\right)$. The maps $i_{m_1} : M_2 \to M_1 \times M_2$ and $i_{m_2} : M_1 \to M_1 \times M_2$ are the natural inclusions obtained by fixing m_1 and m_2, respectively.

Now, the notion of skew-product decomposition of a second order differential operator $L \in \mathfrak{X}_2\left(M_1 \times M_2\right)$ on $M_1 \times M_2$ that one finds in the literature (see for instance [T92]) consists on the existence of two smooth maps $L_1 : M_2 \to \mathfrak{X}_2\left(M_1\right)$ and $L_2 \in \mathfrak{X}_2\left(M_2\right)$ such that for any $f \in C^{\infty}\left(M_1 \times M_2\right)$

$$L\left[f\right]\left(m_1, m_2\right) = \left(L_1\left(m_2\right)\left[f\left(\cdot, m_2\right)\right]\right)\left(m_1\right) + \left(L_2\left[f\left(m_1, \cdot\right)\right]\right)\left(m_2\right). \tag{3.23}$$

The relation between this notion and the one introduced in Definition 3.14 is very easy to establish for semielliptic diffusions. Indeed, suppose that a Stratonovich operator associated to a semielliptic diffusion admits a skew-product decomposition; we just saw that this implies in general the existence of a skew-product decomposition (3.22) of the corresponding Schwartz operator, which in turn implies the availability of a skew-product decomposition of the infinitesimal generator associated to (3.21) in the sense of (3.23). See [T92, page 15] for a sketch of the proof of this fact.

In conclusion, since in the cases that have already been studied, the skew-product decompositions of Stratonovich operators carry in their wake the skew-product decompositions as differential operators of the associated infinitesimal generators, we can focus in what follows on the more general situation that consists of adopting Definition 3.14.

3.3.1 Skew-products on principal fiber bundles. Free actions.

Let M, N be two manifolds, G a Lie group, and $\Phi : G \times M \to M$ a proper and free action. We already know that M/G is a smooth manifold under these hypotheses and that $\pi_{M/G} : M \to M/G$ is a principal fiber bundle with structural group G. The goal of the following paragraphs is to show that any G-invariant Stratonovich operator $S : TN \times M \to TM$ on M admits a local skew-product decomposition. This result is also true even if the action Φ is not free, as we will see in the next section. However, what makes this local decomposition possible in this simpler case is not the fact that the G-action is free and proper but that $\pi_{M/G} : M \to M/G$ is a principal fiber bundle. Consequently, in order to keep our exposition as general as possible, we will adopt as the setup for the rest of this subsection a G-invariant Stratonovich operator $S : TN \times P \to TP$ on an arbitrary (left) G-principal fiber bundle $\pi : P \to Q$. This setup has been studied in detail in [ELL04] for invariant diffusions. In the following proposition we generalize the vertical-horizontal splitting in that paper to arbitrary Stratonovich operators and we formulate it in terms of skew-products.

Proposition 3.15 *Let N be a manifold, $\pi : P \to Q$ a (left) principal bundle with structure group G, $S : TN \times P \to TP$ a G-invariant Stratonovich operator, $X : \mathbb{R}_+ \times \Omega \to N$ a N-valued semimartingale, and $\sigma : U \to \pi^{-1}\left(U\right) \subseteq P$ a local section of π defined on an open neighborhood $U \subseteq Q$. Then, S admits a skew-product decomposition on $\pi^{-1}\left(U\right)$. More explicitly, there exists a diffeomorphism $F : G \times U \to \pi^{-1}\left(U\right)$ and a skew-product split Stratonovich operator $S_{G\times U} : TN \times \left(G \times U\right) \to T\left(G \times U\right)$ such that F establishes a bijection between semimartingales Γ starting on $\pi^{-1}\left(U\right)$ which are solutions of the stochastic system $\left(P, S, X, N\right)$*

up to time $\tau = \inf \left\{ t > 0 \mid \Gamma_t \notin \pi^{-1}(U) \right\}$ *and the* $(G \times U)$*-valued semimartingales* $(\widetilde{g}_t, \Gamma_t^Q)$ *that solve* $(G \times U, S_{G\times U}, X, N)$,

$$\delta(\widetilde{g}_t, \Gamma_t^Q) = S_{G\times U}\left(X, (\widetilde{g}_t, \Gamma_t^Q)\right)\delta X_t. \tag{3.24}$$

Proof. Let $U \subseteq Q$ be an open neighborhood and $\sigma : U \to \pi^{-1}(U) \subseteq P$ a local section of $\pi : P \to Q$. Given that G acts freely on P, the map

$$\begin{array}{rcl} F : G \times U & \longrightarrow & \pi^{-1}(U) \\ (g, q) & \longmapsto & g \cdot \sigma(q) \end{array}$$

is a G-equivariant diffeomorphism, where $g \cdot \sigma(q) = \Phi_g(\sigma(q))$ denotes the (left) action of $g \in G$ on $\sigma(q) \in P$ via $\Phi : G \times P \to P$ and the product manifold $G \times U$ is considered as a left G-space with the action defined by $g \cdot (h, q) := (g \cdot h, q)$. Thus, we can use F to identify $\pi^{-1}(U) \subseteq P$ with the product manifold $G \times U$.

Now, given $p = g \cdot \sigma(q) \in \pi^{-1}(U)$, define $\mathrm{Hor}_p \subseteq T_pP$ as $\mathrm{Hor}_p := T_{\sigma(q)}\Phi_g \circ T_q\sigma(T_qQ)$. It is straightforward to see that the family of horizontal spaces $\{\mathrm{Hor}_p \mid p \in \pi^{-1}(U)\}$ is invariant by the G-action and hence defines a principal connection $A_\sigma \in \Omega^1\left(\pi^{-1}(U); \mathfrak{g}\right)$ on the open neighborhood $\pi^{-1}(U)$. Moreover, if $\Gamma^Q : \mathbb{R}_+ \times \Omega \to Q$ is a Q-valued semimartingale starting at q, then $\sigma\left(\Gamma^Q\right)$ is the unique horizontal lift on P of Γ^Q associated to the connection A_σ starting at $\sigma(q) \in \pi^{-1}(q)$ and defined up to time $\tau_U = \inf\{t > 0 \mid \Gamma_t^Q \notin U\}$.

Consider now the skew-product split Stratonovich operator $S_{G\times U}(x, (g, q)) : TN \times (G \times U) \to T(G \times U)$ such that, for any $x \in N$, $g \in G$, $q \in U$

$$S_{G\times U}(x, (g, q)) = \left(K\left((\sigma(q), x), g\right), S_{P/G}(x, q)\right) \in \mathcal{L}\left(T_xN, T_gG \times T_qU\right),$$

where K is the Stratonovich operator introduced in (3.11) and $S_{P/G}$ the reduced Stratonovich operator constructed out of S as in (3.5). Let $(\widetilde{g}_t, \Gamma_t^Q)$ be a $(G \times U)$-valued semimartingale solution of the stochastic system (3.24), i.e.

$$\delta(\widetilde{g}_t, \Gamma_t^Q) = S_{G\times U}\left(X, (\widetilde{g}_t, \Gamma_t^Q)\right)\delta X,$$

with initial condition $(g, q) \in G \times U$. We claim that $\Gamma_t = F(\widetilde{g}_t, \Gamma_t^Q) = \widetilde{g}_t \cdot \sigma(\Gamma_t^Q)$ is a solution of the stochastic system (P, S, X, N) with initial condition $g \cdot \sigma(q)$ up to the first exit time $\tau_U = \inf\{t > 0 \mid \Gamma_t^Q \notin U\}$. This is a consequence of the Reconstruction Theorem 3.10 and the fact that $\sigma(\Gamma_t^Q)$ is the horizontal lift of a solution of the reduced system $\left(Q, S_{P/G}, X, N\right)$. Conversely, let Γ be a solution of the stochastic system (P, S, X, N) with initial condition $p = g \cdot \sigma(q) \in \pi^{-1}(U)$. By the Reconstruction Theorem 3.10, Γ can be written as $\Gamma_t = \widetilde{g}_t \cdot d_t$. We recall that d_t the horizontal lift with respect to an arbitrary connection $A \in \Omega^1(Q; \mathfrak{g})$ of the solution $\Gamma_t^Q = \pi(\Gamma_t)$ of the reduced system $\left(Q, S_{P/G}, X, N\right)$ (see Theorem 3.9) with initial condition $\sigma(q)$. On the other hand, $\widetilde{g}_t$ is the solution of the stochastic system (3.12) with initial condition $g \in G$. If we take in this procedure $A_\sigma \in \Omega^1\left(\pi^{-1}(U); \mathfrak{g}\right)$ as the auxiliary connection, that is, the one given by the local section $\sigma : U \to \pi^{-1}(U)$, then $d_t = \sigma(\Gamma_t^Q)$ and it is straightforward to check that $(\widetilde{g}_t, \Gamma_t^Q)$ is a solution of (3.24) with initial condition $(g, q) \in G \times U$. ■

Example 3.16 Let G be a Lie group, $H \subseteq G$ a closed subgroup, and R a smooth manifold. In [PR88], Pauwels and Rogers show several examples of skew-product decompositions of Brownian motions on manifolds of the type $R \times G/H$ which share a common feature, namely, they are obtained from skew-product split Brownian motions on $R \times G$ via the reduction $\pi : R \times G \to R \times G/H$. The H-action on $R \times G$ is $h \cdot (r, g) := (r, gh)$, for any $h \in H$, $r \in R$, and $g \in G$. An important result in this paper is Theorem 2 which reads as follows: suppose that $R \times G/H$ is a Riemannian manifold with Riemannian metric η and that the tensor $\pi^*\eta$ is G-invariant. Furthermore, suppose that the decomposition $T_{(r,g)}(R \times G) = T_rR \oplus T_gG$ is orthogonal with respect to $\pi^*\eta$, for any $r \in R$, $g \in G$, and that the Lie algebra $\mathfrak{g}$ of G admits an Ad_H-invariant inner product. Under these hypotheses, $R \times G$ admits a G-invariant Riemannian metric $\hat{\eta}$ such that if Γ is a Brownian motion on $R \times G$ with respect to $\hat{\eta}$ then Γ has a skew-product decomposition and moreover, $\pi(\Gamma)$ is a Brownian motion on $(R \times G/H, \eta)$. This result is repeatedly used in [PR88] to obtain skew-product decompositions of Brownian motions on various manifolds of matrices.

Example 3.17 (Brownian motion on symmetric spaces) Let (M, η) be a Riemannian symmetric space with Riemannian metric η. We want to define Brownian motions on (M, η) by reducing a suitable process defined on the connected component containing the identity of its group of isometries. The notation and most of the results in this example, in addition to a comprehensive exposition on symmetric spaces, can be found in [H78] and [KN69]. The reader is encouraged to check with [ELL98] to learn more about stochastics in the context of homogeneous spaces.

We start by recalling that a M-valued process Γ is a Brownian motion whenever

$$f(\Gamma) - f(\Gamma_0) - \frac{1}{2} \int \Delta(f)(\Gamma_s)\, ds$$

is a real valued local semimartingale for any $f \in C^\infty(M)$, where Δ denotes the Laplacian. The Laplacian is defined as the trace of the Hessian associated to the Levi-Civita connection ∇ of η, that is,

$$\Delta(f)(m) = \sum_{i=1}^{r} \left(\mathcal{L}_{Y_i} \circ \mathcal{L}_{Y_i} - \nabla_{Y_i} Y_i\right)(f)(m)$$

where $\{Y_1, ..., Y_r\} \subset \mathfrak{X}(M)$ is family or vector fields such that $\{Y_1(m), ..., Y_r(m)\}$ is an orthonormal basis of T_mM, $m \in M$.

Let G be the connected component containing the identity of the isometries group $I(M) \subseteq \mathrm{Diff}(M)$ of M. Take $o \in M$ a fixed point and let s be a geodesic symmetry at o. The Lie group G acts on M transitively and, if K denotes the isotropy group of o, M is diffeomorphic to G/K ([H78, Chapter IV, Theorem 3.3]). Denote by $\pi : G \to G/K$ the canonical projection and suppose that $\dim(G) < \infty$. Let $\sigma : G \to G$ be the involutive automorphism of G defined by $\sigma(g) = s \circ \Phi_g \circ s$ for any $g \in G$, where $\Phi : G \times M \to G$ denotes as usual the left action of G on M. $T_e\sigma : \mathfrak{g} \to \mathfrak{g}$ induces an involutive automorphism of $\mathfrak{g}$. That is, $T_e\sigma \circ T_e\sigma = \mathrm{Id}$ but $T_e\sigma \neq \mathrm{Id}$. Let $\mathfrak{k}$ and $\mathfrak{m}$ be the the eigenspaces in $\mathfrak{g}$ associated to the eigenvalues 1 and -1 of $T_e\sigma$, respectively, such that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$. It can be checked that $\mathfrak{k}$ is a Lie subalgebra of $\mathfrak{g}$ and that (see [KN69, Chapter XI Proposition 2.1]).

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{m}] \subseteq \mathfrak{m}, \text{ and } [\mathfrak{m}, \mathfrak{m}] \subseteq \mathfrak{k}.$$

Since the infinitesimal generators $\xi^M \in \mathfrak{X}(M)$ of the G-action Φ on M, with $\xi \in \mathfrak{m}$, span the tangent space at any point $gK \in G/K$, any affine connection is fully characterized by its value on the left-invariant vector fields ξ^M with $\xi \in \mathfrak{m}$. In the particular case of the Levi-Civita connection ∇ associated to the metric η, its G-invariance implies via [KN69, Chapter XI, Theorem 3.3] that

$$\nabla_{\xi^M} \zeta^M (gK) = 0 \tag{3.25}$$

for any pair of left-invariant vector fields ξ^M and ζ^M. A consequence of (3.25) is that the Laplacian Δ takes the expression $\Delta(f)(gK) = \sum_{i=1}^r \mathcal{L}_{\xi_i^M} \circ \mathcal{L}_{\xi_i^M}(f)(gK)$, $gK \in G/K$, where $\left\{\xi_1^M(gK), ..., \xi_r^M(gK)\right\}$ is an orthonormal basis of $T_{gK}(G/K)$.

Let $\{\xi_1, ..., \xi_r\}$ be a basis of $\mathfrak{m}$ such that $\{T_e\pi(\xi_1) ..., T_e\pi(\xi_r)\}$ is an orthonormal basis of $T_K(G/K) \simeq T_oM$ with respect to η_o and let $\left\{\xi_1^G, ..., \xi_r^G\right\} \subset \mathfrak{X}(G)$ be the corresponding family of *right-invariant* vector fields built from $\{\xi_1, ..., \xi_r\}$. Observe that $\left\{\xi_1^M, ..., \xi_r^M\right\}$ is an orthonormal basis of the tangent space at any point $gK \in G/K$ due to the transitivity of the G-action on M and to the G-invariance of the metric η. Consider now the Stratonovich stochastic differential equation on G

$$\delta g_t = \sum_{i=1}^r \xi_i^G(g_t)\delta B_t^i, \tag{3.26}$$

where $\left(B_t^1, ..., B_t^r\right)$ is a $\mathbb{R}^r$-valued Brownian motion. The equation (3.26) is by construction K-invariant with respect to the *right action* $R : K \times G \to G$, $R_k(g) = gk$. In addition, it is straightforward to check that the projection $\pi : G \to G/K$ sends any right-invariant vector field $\xi^G \in \mathfrak{X}(G)$, $\xi \in \mathfrak{g}$, to the infinitesimal generator $\xi^M \in \mathfrak{X}(M)$ of the G-action $\Phi : G \times M \to M$. Indeed, for any $\xi \in \mathfrak{g}$, $g \in G$, and $k \in K$

$$T_g\pi\left(\xi^G(g)\right) = T_g\pi \circ T_eR_g(\xi) = \left.\frac{d}{dt}\right|_{t=0} \pi(\exp(t\xi)\, g) = \left.\frac{d}{dt}\right|_{t=0} \Phi(\exp(t\xi), \pi(g)) = \xi^M(gK),$$

and hence (3.26) projects to the stochastic differential equation

$$\delta\Gamma_t = \sum_{i=1}^r \xi_i^M(\Gamma_t)\delta B_t^i \tag{3.27}$$

on M by the Reduction Theorem 3.9. A straightforward computation shows that that the solution semimartingales of (3.27) have as infinitesimal generator the Laplacian $\Delta = \sum_{i=1}^r \mathcal{L}_{\xi_i^M} \circ \mathcal{L}_{\xi_i^M}$ and hence by the Itô formula

$$f(\Gamma) - f(\Gamma_0) - \frac{1}{2}\int \Delta(f)(\Gamma_s)\, ds = \sum_{i=1}^r \int \xi_i^M[f](\Gamma) dB^i$$

which allows us to conclude that they are Brownian motions. It is worth noticing that since right-invariant systems such that (3.26) are stochastically complete (see [E82, Chapter VII §6]) and by the Reduction and Reconstruction Theorems 3.9 and 3.10 any solution of (3.27) may be written as $\Gamma_t = \pi(g_t)$ for a suitable solution g_t of (3.26), the Brownian motion on a symmetric space is stochastically complete.

3.3.2 Skew-products induced by non-free actions. The tangent-normal decomposition

In this section we will show how the results that we just presented for free actions can be generalized to the non-free case by using the notion of ***slice*** [Ko53, P61] and a generalization to the context of Stratonovich operators of the so-called ***tangent-normal decomposition*** of G-equivariant vector fields with respect to proper group actions [K90, F91].

Let $\Phi : G \times M \to M$ be a proper action of the Lie group G on the manifold M and let M/G be the associated orbit space, M/G. Observe that as the group action is not necessarily free, the orbit space M/G needs not be a smooth manifold.

In order to introduce the notion of slice we start by considering a subgroup $H \subset G$ of G. Suppose that H acts on the left on a certain manifold A. The ***twisted action*** of H on the product $G \times A$ is defined by

$$h \cdot (g, a) = (gh, h^{-1} \cdot a), \quad h \in H, \ g \in G, \ \text{and } a \in A.$$

Note that this action is free and proper by the freeness and properness of the action on the G-factor. The ***twisted product*** $G \times_H A$ is defined as the orbit space $(G \times A)/H$ corresponding to the twisted action. The elements of $G \times_H A$ will be denoted by $[g, a]$, $g \in G$, $a \in A$. The twisted product $G \times_H A$ is a G-space relative to the left action defined by $g' \cdot [g, a] := [g'g, a]$. Also, it can be shown that the action of H on A is proper if and only if the G-action on $G \times_H A$ just defined is proper (see [OR04, Proposition 2.3.17]).

Let now $m \in M$ and denote $H := G_m$. A ***tube*** around the orbit $G \cdot m$ is a G–equivariant diffeomorphism

$$\varphi : G \times_H A \longrightarrow U,$$

where U is a G–invariant neighborhood of the orbit $G \cdot m$ and A is some manifold on which H acts. Note that the G–action on the twisted product $G \times_H A$ is proper since by the properness of the G-action on M, the isotropy subgroup H is compact and, consequently, its action on A is proper.

Definition 3.18 *Let M be a manifold and G a Lie group acting properly on M. Let $m \in M$ and denote $H := G_m$. Let W be a submanifold of M such that $m \in W$ and $H \cdot W = W$. We say that W is a **slice** at m if the G–equivariant map*

$$\begin{array}{rccc} \varphi : & G \times_H W & \longrightarrow & U \\ & [g, s] & \longmapsto & g \cdot s \end{array}$$

is a tube about $G \cdot m$ for some G–invariant open neighborhood U of $G \cdot m$. Notice that if W is a slice at m then $\Phi_g(W)$ is a slice at the point $\Phi_g(m)$.

The Slice Theorem of Palais [P61] proves that there exists a slice at any point of a proper G-manifold. The following theorem, whose proof can be found in [OR04] provides several equivalent characterizations of the concept of slice that are available in the literature.

Theorem 3.19 *Let M be a manifold and G a Lie group acting properly on M. Let $m \in M$, denote $H := G_m$, $\mathfrak{h}$ the Lie algebra of H, and let W be a submanifold of M containing m. Then the following statements are equivalent:*

(i) *There is a tube $\varphi : G \times_H A \longrightarrow U$ about $G \cdot m$ such that $\varphi[e, A] = W$.*

(ii) *W is a slice at m.*

(iii) *The submanifold W satisfies the following properties:*

- ***(a)*** *The set $G \cdot W$ is an open neighborhood of the orbit $G \cdot m$ and W is closed in $G \cdot W$.*
- ***(b)*** *For any $z \in W$ we have that $T_z M = \mathfrak{g} \cdot z + T_z W$. Moreover, $\mathfrak{g} \cdot z \cap T_z W = \mathfrak{h} \cdot z$. In particular, for $z = m$ the sum $\mathfrak{g} \cdot z + T_z W$ is direct.*
- ***(c)*** *W is H–invariant. Moreover, if $z \in W$ and $g \in G$ are such that $g \cdot z \in W$, then $g \in H$.*
- ***(d)*** *Let $\sigma : V \subset G/H \to G$ be a local section of the submersion $G \to G/H$. Then, the map $F : V \times W \to M$ given by $F(gH, z) := \sigma(gH) \cdot z$ is a diffeomorphism onto an open subset of M.*

(iv) *$G \cdot W$ is an open neighborhood of $G \cdot m$ and there is an equivariant smooth retraction*

$$r : G \cdot W \longrightarrow G \cdot m$$

of the injection $G \cdot m \hookrightarrow G \cdot W$ such that $r^{-1}(m) = W$.

Let now $S : TN \times M \to TM$ be a G-invariant Stratonovich operator. The existence of slices for the G-action allow us to carry out two decompositions of S. The first one, that we will call ***tangent-normal decomposition*** is semi-global in the sense that it shares the properties that the Slice Theorem has in this respect, which is global in the orbit directions and local in the directions transversal to the orbits; this decomposition consists of writing S as the sum of two Stratonovich operators such that, roughly speaking, one is tangent to the orbits of the G-action and the other one is transversal to them. The second one is purely local and yields a ***skew-product decomposition*** of S in the sense of Definition 3.14, provided that an additional hypothesis on the isotropies in the slice is present. This hypothesis, whose impact will be explained in detail later on, is generically satisfied and hence the following theorem shows that S admits a skew product decomposition in a neighborhood of most points in M (those points form an open and dense subset of M). These statements are rigorously proved in the following theorem.

Theorem 3.20 *Let $X : \mathbb{R}_+ \times \Omega \to N$ be a N-valued semimartingale, $\Phi : G \times M \to M$ a proper Lie group action, and $S : TN \times M \to TM$ a G-invariant Stratonovich operator. Let $m \in M$ and W a slice at m. Then, there exist two Stratonovich operators $S_N : TN \times W \to TW$ and $S_T : TN \times G \cdot W \to T(G \cdot W)$ such that the following statements hold:*

(i) *Let* $\mathrm{Lie}\,(N(G_z))$ *denote the Lie algebra of the normalizer* $N(G_z)$ *in* G *of the isotropy group* G_z, $z \in G \cdot W$. *The Stratonovich operator* S_T *is* G*-invariant and* $S_T(x,z) \in \mathcal{L}\left(T_xN, \mathrm{Lie}\,(N(G_z)) \cdot z\right)$ *for any* $x \in N$ *and any* $z \in G \cdot W$. *Moreover, there exists an adjoint* G*-equivariant map* $\xi : TN \times G \cdot W \to \mathfrak{g}$, *(that is,* $\xi\,(x, g \cdot z) = \mathrm{Ad}_g \circ \xi\,(x,z)$, *for any* $g \in G$*) such that* $S_T(x,z) = T_e\Phi_z \circ \xi\,(x,z)$.

(ii) *The Stratonovich operator* $S_N : TN \times W \to TW$ *is* G_m*-invariant.*

(iii) *If* $z = g \cdot w \in G \cdot W$, *with* $g \in G$ *and* $w \in W$, *then*

$$S\,(x,z) = S_T\,(x,z) + T_w\Phi_g \circ S_N\,(x,w) = T_w\Phi_g \circ \left(S_T\,(x,w) + S_N\,(x,w)\right). \tag{3.28}$$

This sum of Stratonovich operators will be referred to as the ***tangent-normal decomposition*** *of* S.

(iv) *Let* φ *be the flow of the stochastic system* (W, S_N, X, N) *so that* $\varphi\,(w)$ *denotes the solution of*

$$\delta\Gamma = S_N(X,\Gamma)\delta X \tag{3.29}$$

with initial condition $\Gamma_{t=0} = w$ *a.s.. Let* $S_{\mathfrak{g}\times W} : TN \times (\mathfrak{g} \times W) \to T\,(\mathfrak{g} \times W)$ *be the Stratonovich operator defined as* $S_{\mathfrak{g}\times W}\,(x,(\eta,w)) = \xi\,(x,w) \times S_N\,(x,w) \in \mathcal{L}\,(T_xN, \mathfrak{g} \times T_wW)$ *and let* (η^w, Γ^w) *be the solution semimartingale of the stochastic system* $(\mathfrak{g} \times W, S_{\mathfrak{g}\times W}, X, N)$ *with initial condition* $(0,w) \in \mathfrak{g} \times W$. *Finally, let* $\widetilde{g} : \{0 \leq t < \tau_\varphi\} \to G$ *be the solution of the stochastic system* $(G, L, \eta^w, \mathfrak{g})$ *with initial condition* $g \in G$ *and where* $L : T\mathfrak{g} \times G \to TG$ *is such that* $L\,(\eta, g)\,(\nu) = T_eL_g(\nu)$. *Then, the semimartingale*

$$\Gamma_t = \widetilde{g}_t \cdot \varphi_t\,(w)$$

is a solution up to time τ_φ *of the stochastic system* (M, S, X, N) *with initial condition* $z = g \cdot w \in G \cdot W$.

(v) *Suppose now that* $G_w = G_z$, *for any* $w \in W$. *Then* S *admits a local skew-product decomposition. More specifically, for any point* $m \in M$, *there exists an open neighborhood* $V \subseteq G/G_m$ *of* G_m, *a diffeomorphism* $F : V \times W \to U \subseteq M$, *and a skew-product split Stratonovich operator* $S_{V\times W} : TN \times (V \times W) \to T\,(V \times W)$ *such that* F *establishes a bijection between semimartingales* Γ *starting on* U *which are solution of the stochastic system* (U, S, X, N) *and semimartingales on* $V \times W$ *solution of the stochastic system* $(V \times W, S_{V\times W}, X, N)$. *Moreover,*

$$S_{V\times W}\,(x,(gG_m, w)) = T_g\pi_{G_m} \circ T_eL_g\,(\xi\,(x,w)) \times S_N\,(x,w)$$

for any $x \in N$, $gG_m \in V \subset G/G_m$, *and any* $g \in G$ *such that* $\pi_{G_m}\,(g) = gG_m$.

Remark 3.21 The last point in this theorem shows that proper symmetries of Stratonovich operators imply the availability of skew-products decompositions around most points in the manifold where the solutions take place. Indeed, the Principal Orbit Type Theorem (see for instance [DK99]) shows that there exists an isotropy subgroup H whose associated isotropy type manifold $M_{(H)} := \{z \in M \mid G_z = kHk^{-1}, k \in G\}$ is open and dense in M. Hence, for any point $m \in M_{(H)}$ there exist slice coordinates around the orbit $G \cdot m$ in which the manifold M looks locally like $G \times_H W = G \times_H W_H \simeq G/H \times W_H$. This local trivialization of the manifold M into two factors and the results in part **(v)** of the theorem can be used to split the Stratonovich operator S, in order to obtain a locally defined skew-product around all the points in the open dense subset $M_{(H)}$ of M.

Proof. As we already said, this construction is much inspired by a similar one available in the context of equivariant vector fields [K90, F91]. In this proof we will mimic the strategy for that result followed in [OR04, Theorem 3.3.5].

We start by noting that the properness of the action guarantees that the isotropy subgroup G_m is compact and hence there exists an open G_m-invariant neighborhood $V \subseteq G/G_m$ of G_m and a local section $\sigma : V \subseteq G/G_m \to G$ with the following equivariance property [F91]: $\sigma(h \cdot gG_m) = h\sigma(gG_m)h^{-1}$, for any $h \in G_m$ and $gG_m \in V$. If we now construct with this section the map $F : V \times W \to U \subseteq M$ introduced in Theorem 3.19, that is

$$F(gG_m, w) := \sigma(gG_m) \cdot w, \tag{3.30}$$

we obtain a G_m-equivariant map by considering the diagonal G_m-action in $V \times W$. Since for any $w \in W$ we have that $F^{-1}(w) = (G_m, \sigma(G_m)^{-1} \cdot w)$,

$$T_w F^{-1} \circ S(x, w) =: S_V(x, w) \times S_W(x, w) \in \mathcal{L}\left(T_x N, T_{G_m} V \times T_{\sigma(G_m)^{-1} \cdot w} W\right). \tag{3.31}$$

Define

$$S_N(x, w) := T_{\sigma(G_m)^{-1} \cdot w} \Phi_{\sigma(G_m)} \circ S_W(x, w) \in T_w W \tag{3.32a}$$

$$\begin{aligned} S_T(x, g \cdot w) &:= T_w \Phi_g \circ T_e \Phi_w \circ T_{\sigma(G_m)} R_{\sigma(G_m)^{-1}} \circ T_{G_m} \sigma \circ S_V(x, w) \\ &= T_e \Phi_{g \cdot w} \circ \mathrm{Ad}_g \circ T_{\sigma(G_m)} R_{\sigma(G_m)^{-1}} \circ T_{G_m} \sigma \circ S_V(x, w). \end{aligned} \tag{3.32b}$$

(i) Let $z = g \cdot w \in G \cdot W$, $g \in G$, $w \in W$, $x \in N$, and define $\xi : TN \times G \cdot W \to \mathfrak{g}$ by

$$\xi(x, z) = \mathrm{Ad}_g \circ T_{\sigma(G_m)} R_{\sigma(G_m)^{-1}} \circ T_{G_m} \sigma \circ S_V(x, w). \tag{3.33}$$

It can be seen that $\xi(x, z)$ is well defined by reproducing the steps taken in [OR04, Theorem 3.3.5 (i)]. More specifically, it can be shown that if z is written as $z = g' \cdot w'$ for some other $g' \in G$ and $w' \in W$ then

$$\mathrm{Ad}_g \circ T_{\sigma(G_m)} R_{\sigma(G_m)^{-1}} \circ T_{G_m} \sigma \circ S_V(x, w) = \mathrm{Ad}_{g'} \circ T_{\sigma(G_m)} R_{\sigma(G_m)^{-1}} \circ T_{G_m} \sigma \circ S_V(x, w').$$

Using (3.32b) and (3.33) we have that

$$S_T(x, g \cdot w) = T_w \Phi_{g \cdot w} \circ \xi(x, g \cdot w)$$

It is an exercise to check that $\xi(x, g \cdot w) = \mathrm{Ad}_g \circ \xi(x, w)$, for any $g \in G$, and hence the Stratonovich operator S_T is G-invariant. This G-invariance implies by Proposition 3.7 that the image of $S_T(x,z)$ is such that $\mathrm{Im}(S_T(x,z)) \subseteq T_z M_{G_z}$. On the other hand, $\mathrm{Im}(S_T(x,z)) = \mathrm{Im}(T_e\Phi_z \circ \xi(x,z)) \subseteq \mathfrak{g}\cdot z$, therefore

$$\mathrm{Im}\left(S_T(x,z)\right) \subseteq T_z M_{G_z} \cap \mathfrak{g}\cdot z = T_z\left(N(G_z)\cdot z\right)$$

by [OR04, Proposition 2.4.5] and hence $\mathrm{Im}(\xi(x,z)) \subset \mathrm{Lie}\left(N(G_z)\right)$.

(ii) and **(iii)** It is immediate to see that the Stratonovich operator $S_N : TN \times W \to TW$ defined in (3.32a) is G_m-invariant. Let $w \in W$; using (3.31) and (3.30)

$$\begin{aligned} S\left(x,w\right) &= T_{\left(G_m, \sigma(G_m)^{-1}\cdot w\right)} F \circ \left(S_V(x,w) \times S_W(x,w)\right) \\ &= T_e\Phi_w \circ T_{\sigma(G_m)} R_{\sigma(G_m)^{-1}} \circ T_{G_m}\sigma \circ S_V(x,w) + T_{\sigma(G_m)^{-1}\cdot w}\Phi_{\sigma(G_m)} \circ S_W(x,w) \\ &= S_T\left(x,w\right) + S_N\left(x,w\right), \end{aligned}$$

where (3.32a) and (3.32b) have been used. The equality (3.28) then follows from the G-invariance of S and S_T.

(iv) First of all observe that if (η^w, Γ^w) is the $\mathfrak{g}\times W$-valued semimartingale solution of the stochastic system $(\mathfrak{g}\times W, S_{\mathfrak{g}\times W}, X, N)$ with constant initial condition $(0,w) \in \mathfrak{g}\times W$, then

$$\langle \mu, \eta^w \rangle = \int \langle \xi\left(X, \varphi_t(w)\right)^* (\mu), \delta X\rangle$$

for any $\mu \in \mathfrak{g}^*$. In other words, η^w may be regarded as the solution of the stochastic differential equation

$$\delta\eta^w = \xi\left(X, \varphi_t(w)\right)\delta X \tag{3.34}$$

with initial condition $\eta^w_{t=0} = 0$ a.s.. Notice that η^w is defined up to time $\tau_{\varphi(w)}$, that is, the time of existence of the solution $\varphi(\omega)$. Let now $\Gamma_t = \widetilde{g}_t \cdot \varphi_t\left(w\right)$ be the M-valued semimartingale in the statement. Applying the rules of Stratonovich differential calculus and the Leibniz rule we obtain

$$\delta\Gamma_t = T_{\widetilde{g}_t}\Phi_{\varphi_t(w)}(\delta\widetilde{g}_t) + T_{\varphi_t(w)}\Phi_{\widetilde{g}_t}(\delta\varphi_t(w)) \tag{3.35}$$

We rewrite the first summand in this expression as

$$\begin{aligned} T_{\widetilde{g}_t}\Phi_{\varphi_t(w)}(\delta\widetilde{g}_t) &= T_{\varphi_t(w)}\Phi_{\widetilde{g}_t} \circ T_e\Phi_{\varphi_t(w)} \circ T_{\widetilde{g}_t} L_{\widetilde{g}_t^{-1}}(\delta\widetilde{g}_t) \\ &= T_{\varphi_t(w)}\Phi_{\widetilde{g}_t} \circ T_e\Phi_{\varphi_t(w)}(\delta\eta^w_t) \\ &= T_{\varphi_t(w)}\Phi_{\widetilde{g}_t} \circ T_e\Phi_{\varphi_t(w)} \circ \xi(X_t, \varphi_t(w))\delta X_t \\ &= T_{\varphi_t(w)}\Phi_{\widetilde{g}_t} \circ S_T(X, \varphi_t(w))\delta X_t, \end{aligned}$$

where in the second and third line we have used that $\widetilde{g}_t$ is a solution of $(G, L, \eta^w, \mathfrak{g})$ and equation (3.34), respectively. The second summand of (3.35) can be written as

$$T_{\varphi_t(w)}\Phi_{\widetilde{g}_t}(\delta\varphi_t(w)) = T_{\varphi_t(w)}\Phi_{\widetilde{g}_t} \circ S_N(X, \varphi_t(w))\delta X_t$$

because $\varphi_t(w)$ is a solution of (3.29). Therefore, using (3.28) we can conclude that

$$\begin{aligned} \delta\Gamma_t &= T_{\varphi_t(w)}\Phi_{\widetilde{g}_t} \circ \left(S_N(X, \varphi_t(w)) + S_T(X, \varphi_t(w))\right)\delta X_t \\ &= S\left(X, \widetilde{g}_t \cdot \varphi_t(w)\right)\delta X_t = S\left(X, \Gamma_t\right)\delta X_t \end{aligned}$$

which shows that Γ_t is a solution up to time τ_φ of the stochastic system (M, S, X, N) with initial condition $z = g \cdot w \in G \cdot W$

(v) Let $w \in W$ and $h \in G_m = G_w$. Let Ψ be the twisted action of G_m on W, that is, $\Psi : G_m \times (G \times W) \to (G \times W)$ defined as $\Psi_h(g, w) := \left(gh, h^{-1} \cdot w\right)$, and whose orbit space is the twisted product $G \times_{G_m} W$. The hypothesis $G_m = G_w$, for any $w \in W$, implies that $G \times_{G_m} W$ can be easily identified with $G/G_m \times W$ using the diffeomorphism

$$\begin{array}{rcl} G \times_{G_m} W & \longrightarrow & G/G_m \times W \\ {[g, w]} & \longmapsto & (gG_m, w)\,. \end{array}$$

Consider now the Stratonovich operator defined by

$$S_{G\times W}\left(x, (g, w)\right) = T_e L_g \circ \xi\left(x, w\right) \times S_N\left(x, w\right).$$

We are going to show that $S_{G\times W}$ is G_m-invariant under the action defined by Ψ. Indeed, given that $G_w = G_m$, $\Psi_h(g, w) = (gh, w)$ for any $h \in G_m$, $g \in G$, and $w \in W$, we have

$$\begin{aligned} S_{G\times W}\left(x, \Psi_h(g, w)\right) &= S_{G\times W}\left(x, (gh, w)\right) = T_e L_{gh} \circ \xi\left(x, w\right) \times S_N\left(x, w\right) \\ &= T_h L_g \circ T_e L_h \circ \xi\left(x, w\right) \times S_N\left(x, w\right) \\ &= T_h L_g \circ T_e R_h \circ \operatorname{Ad}_h \circ \xi\left(x, w\right) \times S_N\left(x, w\right) \\ &= T_g R_h \circ T_e L_g \circ \operatorname{Ad}_h \circ \xi\left(x, w\right) \times S_N\left(x, w\right). \end{aligned} \tag{3.36}$$

But due to the G-equivariance of ξ we have $\xi(x, w) = \xi(x, h \cdot w) = \operatorname{Ad}_h \circ \xi(x, w)$, for any $h \in G_m$. In addition, $T_{(g,w)}\Psi_h = T_g R_h \times \mathrm{Id}$, so (3.36) equals

$$T_{(g,w)}\Psi_h \circ \left(T_e L_g \circ \xi\left(x, w\right) \times S_N\left(x, w\right)\right) = T_{(g,w)}\Psi_h \circ S_{G\times W}\left(x, (g, w)\right),$$

which shows that $S_{G\times W}$ is G_m-invariant.

We can therefore apply the Reduction Theorem 3.9 to conclude that $S_{G\times W}$ projects onto a stochastic system $\left(G/G_m \times W, S_{G/G_m\times W}, X, N\right)$ on $G \times_{G_m} W \simeq G/G_m \times W$ with Stratonovich operator

$$\begin{aligned} S_{G/G_m\times W}\left(x, (gG_m, w)\right) \quad : &= T_g \pi_{G_m} \circ S_{G\times W}\left(x, (g, w)\right) \\ &= T_g \pi_{G_m} \circ T_e L_h \circ \xi\left(x, w\right) \times S_N\left(x, w\right), \end{aligned} \tag{3.37}$$

where $x \in N$, $w \in W$, and $g \in G$ is any element such that $\pi_{G_m}(g) = gG_m$. Notice that by (3.28), expression (3.37) proves that the Stratonovich operator $S_{G/G_m\times W}$ is a local skew-product decomposition of S on $G/G_m \times W$.

Concerning the solutions, by **(iv)** any solution of the stochastic system (M, S, X, N) starting at some point $z = g \cdot w \in U \subseteq G \cdot W$ can be written as the image by the action Φ of the solution $(g_t, \varphi_t(w))$ of the stochastic system $(G \times W, S_{G\times W}, X, N)$ starting at $(g, w) \in G \times W$ and defined up to time $\tau_{\varphi(w)}$. Then, the Reduction Theorem 3.9 guarantees that this solution can be projected to a solution of $\left(G/G_m \times W, S_{G/G_m\times W}, X, N\right)$ starting at $(gG_m, w) \in G/G_m \times W$, also defined up to time $\tau_{\varphi(w)}$. Conversely, in order to recover a solution of the original system from a solution $\left((gG_m)_t, w_t\right)$ of $\left(G/G_m \times W, S_{G/G_m\times W}, X, N\right)$ we need to invoke the

Reconstruction Theorem 3.10 by choosing an auxiliary connection $A \in \Omega^1(G; \mathfrak{g}_m)$. This will yield a solution (g_t, w_t) of $(G \times W, S_{G\times W}, X, N)$ with g_t a G-valued semimartingale that can be written as

$$g_t = d_t h_t,$$

where $d_t : \mathbb{R}_+ \times \Omega \to G$ is the horizontal lift of $(gG_m)_t$ with respect to A and $h_t : \mathbb{R}_+ \times \Omega \to G_m$ is a suitable semimartingale on G_m. The key point is that the image by the action Φ of the solution (g_t, w_t) of $(G \times W, S_{G\times W}, X, N)$, that is,

$$\Phi(g_t, w_t) = g_t \cdot w_t = d_t h_t \cdot w_t = d_t \cdot w_t$$

yields a solution of (M, S, X, N). Notice that the semimartingale h_t plays no role. Indeed, let $\sigma : V \subseteq G/G_m \to G$ be the local G_m-equivariant section introduced in the beginning of the proof. We already saw in Proposition 3.15 that if $(gG_m)_t : \mathbb{R}_+ \times \Omega \to G/G_m$ is a G/G_m-valued semimartingale then $\sigma((gG_m)_t)$ is the horizontal lift with respect to the connection $A_\sigma \in \Omega^1\left(\pi_{G_m}^{-1}(V); \mathfrak{g}_m\right)$ induced by the local section σ. Consequently, any solution Γ_t of the initial stochastic system (M, S, X, N) with initial condition $\Gamma_{t=0} = g \cdot w \in U \subset G \cdot W$ can be locally expressed as $\sigma((gG_m)_t) \cdot w_t$ where $((gG_m)_t, w_t)$ is a solution of the stochastic system $\left(G/G_m \times W, S_{G/G_m\times W}, X, N\right)$ with initial condition $(\pi_{G_m}(g), w) \in G/G_m \times W$. ■

Example 3.22 (Liao decomposition of Markov processes) The possibility of decomposing stochastic processes using a group invariance property has been used beyond the context of stochastic differential equations. For example, Liao [L07] has used what he calls the ***transversal submanifolds*** of a compact group action to carry out an ***angular-radial decomposition*** of the Markov processes that are equivariant with respect to those actions. To be more specific, let M be a manifold acted upon by a Lie group G and let $\Gamma : \mathbb{R}_+ \times \Omega \to M$ be a M-valued Markov process with transition semigroup P_t; that is, Γ is a process with càdlàg paths that satisfies the simple Markov property

$$E[f(\Gamma_{t+s}) | \mathcal{F}_t] = P_s f(\Gamma_t)$$

a.s. for $s < t$ and $f \in C_b^\infty(M)$, where $C_b^\infty(M)$ is the space of bounded smooth functions on M, and $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ is the natural filtration induced by Γ. Furthermore, suppose that the Markov process Γ or, equivalently, its transition semigroup P_t is G-equivariant in the sense that

$$P_t(f \circ \Phi_g) = (P_t f) \circ \Phi_g$$

for any $g \in G$. Additionally, in [L07] it is assumed the existence of a submanifold $W \subseteq M$ which is ***globally transversal*** to the G-action. This means that W intersects each G-orbit at exactly one point, that is, for any $w \in W$, $G \cdot w \cap W = \{w\}$ and $M = \bigcup_{w\in W} G \cdot w$. The existence of such global transversal section is a strong hypothesis that only a limited number of actions satisfy. A larger range of applicability of the results in [L07] can be obtained if one is willing to work locally using the slices introduced in this section. Indeed, suppose now that the group G is not compact but just that the group action is proper; let $m \in M$ and $\varphi : G \times_{G_m} W \to U \subseteq M$ a tube around the orbit $G \cdot m$ where, additionally, we assume that $G_w = G_m$ for any $w \in W$. With this hypothesis which, incidentally is the same one that in part **(v)** of Theorem 3.20

allowed us to obtain a skew-product decomposition of the invariant Stratonovich operator, the slice W *is a local transversal manifold* in the sense of [L07].

Let now $J : U \subseteq M \to W$ be the projection that associates to each point, the unique element in its orbit that intersects W. Liao proves [L07, Theorem 1] that the ***radial part*** $y := J(\Gamma)$ of the Markov process Γ is also a Markov process with transition semigroup $Q_t := J^* P_t$. Moreover, if the group G is compact and Γ is Feller then so is y and its generator is fully determined by that of Γ.

Let now $\pi_{G_m} : G \to G/G_m$ be the canonical projection, $V \subseteq G/G_m$ as in Theorem 3.19 **(iii)-(d)**, and let $\phi : V \times W \to U$ be the diffeomorphism associated to the local section $\sigma : V \to \pi_{G_m}^{-1}(V) \subseteq G$ such that $\phi(gG_m, w) = \sigma(gG_m) \cdot w$. Let Γ be U-valued Markov process starting at m and $y = J(\Gamma)$ its radial part. Let $\bar{\Gamma} : \{0 \leq t < \tau_U\} \to V \subseteq G/G_m$ be the process such that $\Gamma_t = \sigma\left(\bar{\Gamma}_t\right) \cdot y_t$, where $\tau_U = \inf\{t > 0 \mid \Gamma_t \notin U\}$. $\bar{\Gamma}$ is called the ***angular part*** of Γ. Liao shows (see [L07, Theorem 3]) that the angular process $\bar{\Gamma}_t$ is a ***nonhomogeneous Lévy process*** under the conditional probability built by conditioning with respect to the σ-algebra generated by the radial process. The reader is encouraged to check with [L07] (see also Section 4.3) for precise definitions and statements (see also [L04]).

3.4 Projectable stochastic differential equations on associated bundles

In the previous section we saw how the availability of the slices associated to a proper group action allows the local splitting of the invariant Stratonovich operators using what we called the tangent-normal decomposition. Additionally, this decomposition yields generically a local skew-product splitting of the invariant Stratonovich operator in question. The key idea behind these splittings was the possibility of locally modeling the manifold where the solutions of the stochastic differential equation take place as a twisted product. A natural setup that we could consider are the manifolds M where this product structure is *global*, that is $M = P \times_G W$, with P and W two G-manifolds. The most standard situation where such manifolds are encountered is when M is the ***associated bundle*** to the G-principal bundle $\pi : P \to Q$: let W be an effective left G-space and $\bar{\pi} : P \times_G W \to Q$, $\bar{\pi}([p, w]) = \pi(p)$. A classical theorem in bundle theory shows that such construction is a fiber bundle with typical fiber W and it is usually referred to as the bundle associated to $\pi : P \to Q$ with fiber W. To be more specific, consider the commutative diagram that defines $\bar{\pi}$:

$$\begin{array}{ccc} P \times W & \xrightarrow{\kappa} & P \times_G W \\ {\scriptstyle pr_1} \downarrow & & \downarrow {\scriptstyle \bar{\pi}} \\ P & \xrightarrow{\pi} & Q. \end{array} \tag{3.38}$$

In this diagram, $\kappa_p : \{p\} \times W \to \bar{\pi}^{-1}(\pi(p)) =: (P \times_H W)_{\pi(p)}$ is a diffeomorphism (see for instance [KMS93, 10.7]). Hence, the correspondence $p \to \kappa_p$, $p \in P$, allows us to consider the elements of P as diffeomorphisms from the typical fiber W of $P \times_G W$ to $\bar{\pi}^{-1}(q)$, with $q = \pi(p)$.

Stochastic processes and diffusions on associated bundles have deserved certain attention in the literature (see [L89] for example) because, as we will see in the following paragraphs, the available geometric structure makes possible a Reduction-Reconstruction procedure that in some cases implies the existence of a global skew-product decomposition. In this context, the notion of invariance is replaced by what we will call $\bar{\pi}$-projectability: if N is a manifold and $S : TN \times M \to TM$ a Stratonovich operator from N to M, we say that S is **$\bar{\pi}$-projectable** if the Stratonovich operator S_Q from N to Q

$$S_Q(x,q) := T_{[p,w]}\bar{\pi} \circ S(x,[p,w]) \in \mathcal{L}\left(T_xN, T_{[p,w]}M\right)$$

is well defined, where $[p,w] \in M$ is any point such that $\bar{\pi}([p,w]) = q \in Q$.

Theorem 3.23 *Let $\bar{\pi} : M = P \times_G W \to Q$ be the associated bundle introduced in the previous discussion. Let N be a manifold, $S : TN \times M \to TM$ a $\bar{\pi}$-projectable Stratonovich operator onto Q, and $X : \mathbb{R}_+ \times \Omega \to N$ a N-valued semimartingale. Then there exist a Stratonovich operator $S_{P\times W} : TN \times (P \times W) \to TP \times TW$ with the property that if (p_t, w_t) is any solution of the stochastic system $(P \times W, S_{P\times W}, X, N)$ with initial condition $(p,w) \in P \times W$, then $\Gamma_t := \kappa(p_t, w_t)$ is the solution of (M, S, X, N) starting at $[p,m]$. Furthermore, p_t can be written as the horizontal lift of $\bar{\pi}(\Gamma_t)$ with respect to an auxiliary connection $A \in \Omega^1(P;\mathfrak{g})$. Conversely, if Γ_t is a solution of (M,S,X,N) and p_t the horizontal lift of $\bar{\pi}(\Gamma_t)$ with respect to A, then $\left(p_t, \kappa_{p_t}^{-1}(\Gamma_t)\right)$ is a solution of $(P \times W, S_{P\times W}, X, N)$.*

Proof. Let $A \in \Omega^1(P;\mathfrak{g})$ be an auxiliary principal connection for $\pi : P \to Q$ and let $\widehat{A}_p : T_{\pi(p)}Q \to \operatorname{Hor}_p P \subseteq T_pP$ be the inclusion of the tangent space T_qQ at $q = \pi(p)$ into the horizontal space $\operatorname{Hor}_p P$ at $p \in P$ defined by A. Consider the family of linear maps $\widehat{\mathbf{A}}_{[p,w]} : T_{\bar{\pi}([p,w])}Q \to T_{[p,w]}M$ for any $[p,w] \in P \times_G W$ as

$$\widehat{\mathbf{A}}_{[p,w]} = T_p\kappa_w \circ \widehat{A}_p, \tag{3.39}$$

where $\kappa_w(p) := \kappa(p,w)$ for any $w \in W$. The family of maps $\{\widehat{\mathbf{A}}_{[p,w]} \mid [p,w] \in M\}$ define what is called the ***induced connection*** $\mathbf{A}$ ([KMS93, 11.8]) on $P \times_G W$ by $A \in \Omega^1(P;\mathfrak{g})$. It can be easily checked that $\mathbf{A}$ is well-defined, that is, the expression (3.39) does not depend on the particular choice of $p \in P$ and $w \in W$ in the class $[p,w] \in P \times_G W$ used to define it. Indeed, if $[p,w] = [p',w']$ then there exists some $g \in G$ such that $p' = g \cdot p$ and $w' = g^{-1} \cdot w$. Since the connection A is principal, $\widehat{A}_{p'} = T_pR_g \circ \widehat{A}_p$, where $R : G \times P \to P$ denotes the G-right action on P. On the other hand, since $\kappa(p' = p \cdot g, w') = \kappa(p, g \cdot w')$, we have

$$T_{p'}\kappa_{w'} \circ T_pR_g = T_p\kappa_{g\cdot w'} \quad \text{or, equivalently,} \quad T_{p'}\kappa_{w'} = T_p\kappa_{g\cdot w'} \circ T_{p'}R_{g^{-1}}. \tag{3.40}$$

Therefore, $\widehat{\mathbf{A}}_{[p',w']} = T_{p'}\kappa_{w'} \circ \widehat{A}_{p'} = T_p\kappa_{g\cdot w'} \circ T_{p'}R_{g^{-1}} \circ T_pR_g \circ \widehat{A}_p = T_p\kappa_w \circ \widehat{A}_p = \widehat{\mathbf{A}}_{[p,w]}$.

Let $S_Q : TN \times Q \to TQ$ be the Stratonovich operator defined as

$$S_Q(x,q) := T_{[p,w]}\bar{\pi} \circ S(x,[p,w]), \tag{3.41}$$

where $[p,w] \in P \times_G W$ is any point such that $\bar{\pi}([p,w]) = q$, $x \in N$, and $w \in W$. This Stratonovich operator is well-defined because S is by hypothesis $\bar{\pi}$-projectable. Let $\widehat{H}_{[p,w]} : T_{[p,w]}M \to \operatorname{Hor}_{[p,w]} M \subseteq T_{[p,w]}M$ and $\widehat{V}_{[p,w]} : T_{[p,w]}M \to \operatorname{Ver}_{[p,w]} M \subseteq T_{[p,w]}M$ be the projections onto the horizontal and vertical spaces associated to $\mathbf{A}$, respectively, at $[p,w] \in P \times_G W$. Define the Stratonovich operator $S_{P\times W} : TN \times (P \times W) \to TP \times TW$ as

$$S_{P\times W}\left(x,(p,w)\right) = \widehat{A}_p \circ S_Q\left(x, \pi\left(p\right)\right) \times \left(T_w \kappa_p\right)^{-1} \circ \widehat{V}_{[p,w]} \circ S\left(x,[p,w]\right) \in \mathcal{L}\left(T_x N, T_{(p,w)}(P\times W)\right) \tag{3.42}$$

for any $x \in N$, $w \in W$, and $p \in P$. Recall from (3.38) that $\kappa_p : W \to M_{\pi(p)}$ is a diffeomorphism for any $p \in P$ and hence $\left(T_w \kappa_p\right)^{-1}$ exists as a map. Now, we claim that if (p_t, w_t) is a $(P\times W)$-valued semimartingale solution of the stochastic system $(P\times W, S_{P\times W}, X, N)$ then $\Gamma_t := \kappa_{p_t}(w_t)$ is a solution of (M,S,X,N). Indeed, applying the Stratonovich rules for differential calculus,

$$\begin{aligned}
\delta\Gamma_t &= T_{w_t}\kappa_{p_t}\left(\delta w_t\right) + T_{p_t}\kappa_{w_t}\left(\delta p_t\right)\\
&= T_{w_t}\kappa_{p_t} \circ \left(T_{w_t}\kappa_{p_t}\right)^{-1} \circ \widehat{V}_{[p_t,w_t]} \circ S\left(X_t,[p_t,w_t]\right)\delta X_t + T_{p_t}\kappa_{w_t} \circ \widehat{A}_{p_t} \circ S_Q\left(X_t, \pi\left(p_t\right)\right)\delta X_t\\
&= \widehat{V}_{[p_t,w_t]} \circ S\left(X_t,[p_t,w_t]\right)\delta X_t + \widehat{\mathbf{A}}_{[p_t,w_t]} \circ S_Q\left(X_t, \pi\left(p_t\right)\right)\delta X_t\\
&= \widehat{V}_{[p_t,w_t]} \circ S\left(X_t,[p_t,w_t]\right)\delta X_t + \widehat{\mathbf{A}}_{[p_t,w_t]} \circ T_{[p_t,w_t]}\bar{\pi} \circ S\left(X_t,[p_t,w_t]\right)\delta X_t\\
&= \widehat{V}_{[p_t,w_t]} \circ S\left(X_t,[p_t,w_t]\right)\delta X_t + \widehat{H}_{[p_t,w_t]} \circ S\left(X_t,[p_t,w_t]\right)\delta X_t\\
&= S\left(X_t,[p_t,w_t]\right)\delta X_t = S\left(X_t,\Gamma_t\right)\delta X_t,
\end{aligned}$$

and hence Γ_t is a solution of (M,S,X,N).

Conversely, let Γ_t be a solution of (M,S,X,N) such that $\Gamma_{t=0} = [p,m]$ a.s. and let p_t be the horizontal lift of $\bar{\pi}\left(\Gamma_t\right)$ with respect to the auxiliary connection $A \in \Omega^1\left(P;\mathfrak{g}\right)$ starting at some $p_0 \in \pi^{-1}\left(\bar{\pi}\left([p,w]\right)\right)$. Define $\widetilde{w}_t := \kappa_{p_t}^{-1}\left(\Gamma_t\right)$. Observe that $\widetilde{w}_{t=0} = \widetilde{w}_0$ is such that $[p_0,\widetilde{w}_0] = [p,m]$. Since $\kappa_p : W \to M_{\pi(p)}$ is a diffeomorphism, $\widetilde{w}_t$ is uniquely determined a.s. by Γ_t once p_t is fixed. Indeed, $\widetilde{w}_t$ is the unique semimartingale such that $\kappa_{p_t}\left(\widetilde{w}_t\right) = \Gamma_t$. But we have already seen that the solution of $(P\times W, S_{P\times W}, X, N)$ starting at $(p_0,\widetilde{w}_0) \in P\times W$ may be expressed as (p_t,w_t), with p_t the fixed horizontal lift of $\bar{\pi}\left(\Gamma_t\right)$ that we have been using all along. Therefore $w_t = \widetilde{w}_t$ a.s. necessarily and $w_t = \kappa_{p_t}^{-1} \circ \kappa_{p_t}\left(w_t\right) = \kappa_{p_t}^{-1}\left(\Gamma_t\right)$. ■

Corollary 3.24 *Using the same notation as in the proof of Theorem 3.23, suppose that* $\left(T_w\kappa_p\right)^{-1} \circ \widehat{V}_{[p,w]} \circ S\left(x,[p,w]\right)$ *in (3.42) does not depend on* $p \in P$. *In such case there exists a unique* G*-invariant Stratonovich operator* $S_W : TN \times W \to TW$ *from* N *to* W *determined by the relation*

$$T_w\kappa_p \circ S_W\left(x,w\right) = \widehat{V}_{[p,w]} \circ S\left(x,[p,w]\right) \tag{3.43}$$

for any $x \in N$, $w \in W$, *and* $p \in P$. *Moreover,* $S_{P\times W}$ *in (3.42) admits the skew-product decomposition*

$$S_{P\times W}\left(x,(p,w)\right) = \widehat{A}_p \circ S_Q\left(x,\pi\left(p\right)\right) \times S_W\left(x,w\right).$$

Proof. First of all notice that as $\left(T_w\kappa_p\right)^{-1} \circ \widehat{V}_{[p,w]} \circ S\left(x,[p,w]\right)$ does not depend on $p \in P$, the expression (3.43) is a good definition that uniquely determines S_W. The only non-trivial point in the statement that needs proof is the G-invariance of S_W: let $g \in G$ and (p',w'), $(p,w) \in P\times W$ such that $p' = p\cdot g$ and $w' = g^{-1}\cdot w$. Since $\widehat{V}_{[p',w']} \circ S\left(x,[p',w']\right) = \widehat{V}_{[p,w]} \circ S\left(x,[p,w]\right)$, we necessarily have

$$T_w\kappa_p \circ S_W\left(x,w\right) = T_{w'}\kappa_{p'} \circ S_W\left(x,w'\right).$$

As $\kappa(p \cdot g, w) = \kappa(p, g \cdot w)$, we have that $T_{g \cdot w}\kappa_p \circ T_w l_g = T_w \kappa_{p \cdot g}$, where $l : G \times W \to W$ is the G-action on W. Thus,

$$T_{w'}\kappa_{p'} \circ S_W\left(x, w'\right) = T_w \kappa_p \circ T_{g^{-1} \cdot w} l_g \circ S_W\left(x, g^{-1} \cdot w\right).$$

Since $T_w \kappa_p : T_w W \to T_{[p,w]}(P \times_G W)$ is an isomorphism, we conclude comparing the two previous relations that

$$S_W(x, w) = T_{g^{-1} \cdot w} l_g \circ S_W\left(x, g^{-1} \cdot w\right),$$

necessarily, which amounts to S_W being G-invariant. ■

Remark 3.25 It is worth noticing that, under the hypotheses of Corollary 3.24 and unlike Theorem 3.20, the skew-product decomposition of $S_{P \times W} : TN \times (P \times W) \to T(P \times W)$ is now global.

Remark 3.26 If the hypotheses of Corollary 3.24 hold, we can solve a stochastic system (M, S, X, N) on the associated bundle $\bar{\pi} : M = P \times_G W \to Q$ with $\bar{\pi}$-projectable Stratonovich operator S using the following reduction-reconstruction scheme. On one hand, we find the solution starting at $\bar{\pi}([p, w])$ on the base space system (Q, S_Q, X, N), where S_Q was given in (3.41). We lift then this solution to the principal bundle P using an auxiliary connection $A \in \Omega^1(P; \mathfrak{g})$. We choose the lift p_t starting at some $p_0 \in \pi^{-1}(\bar{\pi}([p, w]))$. On the other hand, we find the solution w_t of the independent stochastic system (W, S_W, X, N) with initial condition w_0 such that $\kappa(p_0, w_0) = [p, w]$. Then $\kappa_{p_t}(w_t)$ is the solution of (M, S, X, N) starting at $[p, w]$.

Example 3.27 Projectable SDEs and the horizontal-vertical factorization of diffusion operators. In this example we show how some of the results in [L89] on the factorization of certain semielliptic differential operators on associated bundles can be rethought in the light of the results in Theorem 3.23 and Corollary 3.24. We recall that a second order differential operator $L_Q \in \mathfrak{X}_2(Q)$ on a manifold Q is called semielliptic if any point $q \in Q$ has an open neighborhood U where L_Q can be locally written as

$$L_Q|_U = \sum_{i=1}^{s} \mathcal{L}_{Y_i}\mathcal{L}_{Y_i} + \mathcal{L}_{Y_0} \tag{3.44}$$

for some Y_0, $Y_i \in \mathfrak{X}(U)$, $i = 1, ..., s$. Such a semielliptic operator can be seen as the infinitesimal generator for the laws of the solution semimartingales of the following stochastic system $(Q, S_Q, X, \mathbb{R} \times \mathbb{R}^s)$ (see for instance [IW89, Theorem 1.2, page 238]): let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R} \times \mathbb{R}^s$ be the semimartingale

$$X_t(\omega) = \left(t, B_t^1(\omega), ..., B_t^s(\omega)\right),$$

where $\left(B^1, ..., B^s\right)$ is a s-dimensional Brownian motion and consider the Stratonovich operator

$$\begin{array}{rcl} S_Q(x, q) : T_x(\mathbb{R} \times \mathbb{R}^s) & \longrightarrow & T_q U \subseteq T_q Q \\ \left(u, v^1, ..., v^s\right) & \longmapsto & uY_0 + \sum_{i=1}^{s} v^i Y_i. \end{array}$$

Let now G be a Lie group, $\pi : P \to Q$ a principal G-bundle, and consider a manifold W acted upon by G via the map $l : G \times W \to W$. Let $L_W \in \mathfrak{X}_2(W)$ be the semielliptic differential operator on W given by

$$L_W = \sum_{i=1}^{n} \mathcal{L}_{Z_i}\mathcal{L}_{Z_i} + \mathcal{L}_{Z_0}$$

where $Z_0, Z_1, ..., Z_n \in \mathfrak{X}(V)$ on some $V \subseteq W$. As we just did, we will consider L_W as the generator for the laws of the solutions of the stochastic system $\left(W, S_W, X', \mathbb{R}^{n+1}\right)$, where $X' : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{n+1}$ is a noise semimartingale constructed using the time process t and n independent Brownian motions, and S_W is the Stratonovich operator given by

$$\begin{array}{rcl} S_W(x, w) : T_x(\mathbb{R} \times \mathbb{R}^n) & \longrightarrow & T_w V \subseteq T_w W \\ \left(u, v^1, ..., v^n\right) & \longmapsto & uZ_0 + \sum_{i=1}^{n} v^i Z_i. \end{array}$$

In addition, we will assume that both L_W and S_W are G-invariant. Let $\widehat{\mathbf{A}}$ be a connection on the associated bundle $M = P \times_G Q$ and define the Stratonovich operator $S : T\mathbb{R}^{n+s+1} \times M \to TM$ as

$$S(x, [p, w]) = T_w\kappa_p \circ S_W(x, w) + \widehat{\mathbf{A}}_{[p,w]} \circ S_Q(x, \pi(p))$$

consistently with the notation introduced so far. Taking $\left(B_t^1, ..., B_t^{n+s}\right)$ a $(n+s)$-dimensional Brownian motion, the stochastic system $(M, S, \widetilde{X}, \mathbb{R}^{n+s+1})$ with stochastic component $\widetilde{X} : \mathbb{R}_+ \times \Omega \to \mathbb{R}^{n+s+1}$ given by $\widetilde{X}_t(\omega) = \left(t, B_t^1(\omega), ..., B_t^{n+s}(\omega)\right)$ satisfies by construction the hypotheses of Theorem 3.23 and Corollary 3.24. The projected stochastic system of $(M, S, \widetilde{X}, \mathbb{R}^{n+s+1})$ onto Q is obviously $\left(Q, S_Q, X, \mathbb{R}^{s+1}\right)$ and the one induced in the typical fiber W is $\left(W, S_W, X', \mathbb{R}^{n+1}\right)$. It is straightforward to check that the probability laws of the solutions of $(M, S, \widetilde{X}, \mathbb{R}^{n+s+1})$ have as infinitesimal generator

$$L_M = \widetilde{L}_Q + L_W^*, \tag{3.45}$$

where $\widetilde{L}_Q$ is what Liao [L89] calls the **horizontal lift** of L_Q and L_W^* the **vertical operator** induced by L_W.

Many of the results presented in [L89] about the factorization (3.45) of semielliptic operators on associated bundles and their related diffusions can be understood from the perspective of stochastic systems and stochastic differential equations that we have adopted here using Theorem 3.23 and Corollary 3.24. In order to illustrate this point consider the following result in Liao's article about Riemannian submersions (see also [EK85]): let (M, η) be a complete Riemannian space with Riemann metric tensor η and let $\bar{\pi} : M \to Q$ be a Riemannian submersion with totally geodesic fibers. In this setup, $\bar{\pi} : M \to Q$ is an associated bundle whose structure group G is the group of isometries of the standard fiber $W := \bar{\pi}^{-1}(q_0)$ for some $q_0 \in Q$ [H60]. Indeed, it can be checked that all the fibers of $\bar{\pi} : M \to Q$ are isometric, so we can take any of them as a standard fiber, and that G has finite dimension [BB82, Remark 1.10, page 185]. Let $\pi : P \to Q$ be the corresponding principal bundle. Additionally, since $\kappa_p : W \to \bar{\pi}^{-1}(q)$ is an isometry for any $p \in P$, the restriction $\eta_{\bar{\pi}^{-1}(q)}$ of the metric η to $\bar{\pi}^{-1}(q)$ may be considered as induced from the metric $\eta_{\bar{\pi}^{-1}(q_0)}$ of W by κ_p which, in addition, is invariant by G. Then,

$$\Delta_M = \widetilde{\Delta}_Q + \Delta_W^* \tag{3.46}$$

where Δ_Q is the Laplacian on Q and Δ_W the Laplacian on W ([L89, Proposition 3]). As a consequence of (3.46), if Γ_t is a M-valued Brownian motion associated to the Laplacian Δ_M

on M then $\bar{\pi}(\Gamma_t)$ is a Brownian motion on Q with generator Δ_Q (see also [E82, Theorem 10E]). Let now $A \in \Omega^1(P, \mathfrak{g})$ be the principal connection on P whose associated connection $\mathbf{A}$ on $\bar{\pi} : M \to Q$ is such that $\mathrm{Hor}_m^{\perp} = \mathrm{Ver}_m$ for any $m \in M$, that is, the horizontal subspace $\mathrm{Hor}_m \subset T_m M$ of $\mathbf{A}$ is the orthogonal complement of Ver_m, $m \in M$. Then, if p_t denotes the horizontal lift of $\bar{\pi}(\Gamma_t)$ to P with respect to A then $\kappa_{p_t}^{-1}(\Gamma_t)$ is a Brownian motion on W with generator Δ_W^* [L89, Propositions 6].

3.5 The Hamiltonian case

Hamiltonian dynamical systems are a class of differential equations in the non-stochastic deterministic context in which reduction techniques have been much developed. This is mainly due to their central role in mechanics and applications to physics and also to the added value that symmetries usually have in this category. As we saw in Proposition 3.7 the symmetries of a stochastic differential equation bring in their wake certain invariance properties of its flow that have to do with the preservation of the isotropy type submanifolds. Symmetric Hamiltonian deterministic systems also preserve isotropy type submanifolds but they usually exhibit additional invariance features caused by the presence of symmetry induced ***first integrals*** or ***constants of motion***, usually encoded as components of a ***momentum map***.

The goal in this section is to show that the reduction and reconstruction techniques that have been developed for deterministic Hamiltonian dynamical systems can be extended to the stochastic Hamiltonian systems that have been introduced in Chapter 2 as a generalization of those in [B81] and that we now briefly review. In the following paragraphs we will assume certain familiarity with standard deterministic Hamiltonian systems and reduction theory (see for instance [AM78, OR04] and references therein).

Let $(M, \{\cdot,\cdot\})$ be a finite dimensional Poisson manifold, $X : \mathbb{R}_+ \times \Omega \to V$ a continuous semimartingale that takes values on the vector space V with $X_0 = 0$, and let $h : M \to V^*$ be a smooth function with values in V^*, the dual of V. Let $\{\epsilon^1, \ldots, \epsilon^r\}$ be a basis of V^* and let $h_1, \ldots, h_r \in C^\infty(M)$ be such that $h = \sum_{i=1}^r h_i \epsilon^i$. According to Definition 2.2, the ***stochastic Hamiltonian system*** associated to h with ***stochastic component*** X is the stochastic differential equation

$$\delta\Gamma^h = H(X, \Gamma)\delta X \tag{2.5}$$

defined by the Stratonovich operator $H(v, z) : T_v V \to T_z M$ defined by

$$H(v, z)(u) := \sum_{i=1}^r \left\langle \epsilon^i, u \right\rangle X_{h_i}(z), \tag{2.6}$$

where X_{h_i} is the Hamiltonian vector field associated to $h_i \in C^\infty(M)$. In this case, the dual Stratonovich operator $H^*(v, z) : T_z^* M \to T_v^* V$ of $H(v, z)$ is given by $H^*(v, z)(\alpha_z) = -\mathbf{d}h(z) \cdot B^\sharp(z)(\alpha_z)$, where $B^\sharp : T^*M \to TM$ is the vector bundle map naturally associated to the Poisson tensor $B \in \Lambda^2(M)$ of $\{\cdot,\cdot\}$ and $\mathbf{d}h = \sum_{i=1}^r \mathbf{d}h_i \otimes \epsilon^i$. We will usually summarize this construction by saying that $(M, \{\cdot,\cdot\}, h, X)$ is a ***stochastic Hamiltonian system***. We will dedicate particular attention to the symplectic case (M, ω) in which the bracket $\{\cdot,\cdot\}$ is obtained from the symplectic form ω via the expression $\{f, h\} = \omega(X_f, X_h)$, $f, h \in C^\infty(M)$.

3.5.1 Invariant manifolds and conserved quantities of a stochastic Hamiltonian system

As we already said, the presence of symmetries in a Hamiltonian system forces the appearance of invariance properties that did not use to occur for arbitrary symmetric dynamical systems. Before we proceed with the study of those conservation laws in the stochastic Hamiltonian case, we extract some conclusions on invariant manifolds that can be obtained from Proposition 3.6 in that situation, some of them already stated in Chapter 2.

Proposition 3.28 *Let $(M, \{\cdot,\cdot\}, h : M \to V^*, X)$ be a stochastic Hamiltonian system. Let $\{\epsilon^1, \ldots, \epsilon^r\}$ be a basis of V^* and write $h = \sum_{i=1}^r h_i \epsilon^i$. Consider the following situations:*

- ***(i)*** *Suppose that M is symplectic (respectively, Poisson) and let $z \in M$ be such that $\mathbf{d}h(z) = 0$ (respectively, $X_{h_i}(z) = 0$, for all $i \in \{1, \ldots, r\}$). Then, the Hamiltonian semimartingale Γ^h with constant initial condition $\Gamma_0(\omega) = z$, for all $\omega \in \Omega$, is an equilibrium, that is $\Gamma^h = \Gamma_0$.*
- ***(ii)*** *Let $S_1, \ldots, S_r$ be submanifolds of M with transverse intersection $S := S_1 \cap \ldots \cap S_r$, such that $X_{h_i}(z_i) \in T_{z_i}S_i$, for all $z_i \in S_i$ and $i \in \{1, \ldots, r\}$. Then S is a local invariant submanifold of the stochastic Hamiltonian system $(M, \{\cdot,\cdot\}, h : M \to V^*, X)$.*
- ***(iii)*** *The symplectic leaves of $(M, \{\cdot,\cdot\})$ are local invariant submanifolds of the stochastic Hamiltonian system $(M, \{\cdot,\cdot\}, h : M \to V^*, X)$.*

Proof. It is a direct consequence of Proposition 3.6 and of the fact that the Stratonovich operator is given by $H(v, z)(u) := \sum_{i=1}^r \langle \epsilon^j, u\rangle X_{h_j}(z)$. In **(i)** the hypothesis $\mathbf{d}h(z) = 0$ implies in the symplectic case that $X_{h_i}(z) = 0$, for all $i \in \{1, \ldots, r\}$. Hence, both in the symplectic and in the Poisson cases $H(v, z) = 0$ and hence by Proposition 3.6, the point z is an invariant submanifold and consequently an equilibrium. For **(ii)** it suffices to recall that the transversality hypothesis implies that $T_zS = T_zS_1 \cap \ldots \cap T_zS_r$, for any $z \in S$. **(iii)** is a restatement of Proposition 2.9. ■

In the Hamiltonian case, most of the invariant manifolds of a system come as the level sets of a conserved quantity (also called first integral) of the motion. Recall that, according to Definition 2.12, a function $f \in C^\infty(M)$ is said to be a ***(strongly) conserved quantity*** of the stochastic differential equation associated to X and S when for any solution semimartingale Γ we have that $f(\Gamma) = f(\Gamma_0)$. We now concentrate on the conserved quantities that one can associate to the invariance of a Hamiltonian system with respect to a group action. We recall that given a Lie group G acting on the Poisson manifold $(M, \{\cdot,\cdot\})$ (respectively, symplectic (M, ω)) via the map $\Phi : G \times M \to M$, we will say that the action is ***canonical*** when for any $g \in G$ and $f, h \in C^\infty(M)$, $\{f, h\} \circ \Phi_g = \{\Phi_g^* f, \Phi_g^* h\}$ (respectively, $\Phi_g^*\omega = \omega$). In this context, we will say that the Hamiltonian system $(M, \{\cdot,\cdot\}, h : M \to V^*, X)$ is ***G-invariant*** whenever the G-action on M is canonical and the Hamiltonian function $h : M \to V^*$ is G-invariant. Notice that the invariance of h and the canonical character of the action imply that the associated Stratonovich operator H is also G-invariant. Indeed, Let $\{\epsilon^1, \ldots, \epsilon^r\}$ be a basis of V^* and write $h = \sum_{i=1}^r h_i \epsilon^i$; if h is G-invariant, then so are the components h_i, $i \in \{1, \ldots, r\}$, that is $h_i \in C^\infty(M)^G$, and hence, for any $g \in G$ we have that $T\Phi_g \circ X_{h_i} = X_{h_i} \circ \Phi_g$, which implies that $H(v, z)(u) := \sum_{i=1}^r \left\langle \epsilon^i, u\right\rangle X_{h_i}(z)$is G-invariant.

Now suppose that M is a Poisson manifold $(M, \{\cdot, \cdot\})$ acted properly and canonically upon by a Lie group G. We also recall that the ***optimal momentum map*** [OR02] $\mathcal{J} : M \to M/D_G$ of the G-action on $(M, \{\cdot, \cdot\})$ is the projection onto the leaf space of the integrable distribution $D_G \subset TM$ (in the generalized sense of Stefan-Sussmann) given by $D_G := \{X_f \mid f \in C^\infty(M)^G\}$.

Proposition 3.29 *Let (M, h, X, V) be a standard Hamiltonian system acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Suppose that $h : M \to V^*$ is a G-invariant function.*

(i) ***Law of conservation of the isotropy****: The isotropy type submanifolds M_I are invariant submanifolds of the stochastic Hamiltonian system associated to h and X, for any isotropy subgroup $I \subset G$.*

(ii) ***Noether's Theorem****: If the G-action on $(M, \{\cdot, \cdot\})$ has a momentum map associated $\mathbf{J} : M \to \mathfrak{g}^*$ then its level sets are left invariant by the stochastic Hamiltonian system associated to h and X. Moreover, its components are conserved quantities.*

(iii) ***Optimal Noether's Theorem****: The level sets of the optimal momentum map $\mathcal{J} : M \to M/D_G$ are local invariant subsets of the stochastic Hamiltonian system associated to h and X.*

Proof. **(i)** As we already saw, the G-invariance of h implies that

$$H(v, z)(u) := \sum_{i=1}^{r} \langle \epsilon^j, u \rangle X_{h_j}(z)$$

is G-invariant. The statement follows from Proposition 3.7. **(ii)** Let $\xi \in \mathfrak{g}$ be arbitrary and let $\mathbf{J}^\xi := \langle \mathbf{J}, \xi \rangle \in C^\infty(M)$ be the corresponding component. The G-invariance of the components h_i of the Hamiltonian implies that $\{\mathbf{J}^\xi, h_i\} = -\mathbf{d}h_i \cdot \xi_M = 0$, where $\xi_M \in \mathfrak{X}(M)$ is the infinitesimal generator associated to the element ξ. By formula (2.8) we have that

$$\mathbf{J}^\xi(\Gamma^h) - \mathbf{J}^\xi(\Gamma_0) = \sum_{j=1}^{r} \int \{\mathbf{J}^\xi, h_j\} \, \delta X^j = 0,$$

where X_j, $j \in \{1, \ldots, r\}$, are the components of X in the basis $\{e_1, \ldots, e_r\}$ of V dual to the basis $\{\epsilon^1, \ldots, \epsilon^r\}$ of V^*. Since this equality holds for any $\xi \in \mathfrak{g}$, we have that $\mathbf{J}(\Gamma^h) = \mathbf{J}(\Gamma_0)$ and the result follows. **(iii)** It is a straightforward consequence of the construction of the optimal momentum map and Proposition 3.6. ■

Remark 3.30 When the manifold M is symplectic and the group action has a standard momentum map $\mathbf{J} : M \to \mathfrak{g}^*$ associated, part **(iii)** in the previous proposition implies the first two since it can be shown that in that situation (see [OR02]) the level sets of the optimal momentum map coincide with the connected components of the intersections $\mathbf{J}^{-1}(\mu) \cap M_I$, with $\mu \in \mathfrak{g}^*$ and I an isotropy subgroup of the G-action on M.

Remark 3.31 The level sets of the momentum map $\mathbf{J}$ may not be submanifolds of M unless the G-action is, in addition to proper and canonical, also free ([OR04, Corollary 4.6.2]). If this is the case, the relation $\{\mathbf{J}^\xi, h_i\} = -\mathbf{d}h_i \cdot \xi_M = 0$, which stems from the G-invariance of h, implies then that $\operatorname{Im}(H(v,z)) \subset T_z\mathbf{J}^{-1}(\mu)$, for any $z \in \mathbf{J}^{-1}(\mu)$ and any $v \in V$. Then Proposition 3.6 may be invoked to prove the invariance of the fibers $\mathbf{J}^{-1}(\mu)$ under the stochastic Hamiltonian system associated to H.

3.5.2 Stochastic Hamiltonian reduction and reconstruction

The goal of this section is showing that stochastic Hamiltonian systems share with their deterministic counterpart a good behavior with respect to symmetry reduction. The main idea that our following theorem tries to convey to the reader is that *the symmetry reduction of a stochastic Hamiltonian system yields a stochastic Hamiltonian system*, that is, the stochastic Hamiltonian category is stable under reduction.

The following theorem spells out, in the simplest possible case, how to reduce symmetric Hamiltonian stochastic systems. In a remark below we give the necessary prescriptions to carry this procedure out in more general situations. The main simplifying hypothesis is the freeness of the action. We recall that in this situation, the orbit space M/G inherits from M a Poisson structure $\{\cdot,\cdot\}_{M/G}$ naturally obtained by projection of that in M, that is, $\{f,g\}_{M/G} \circ \pi := \{f \circ \pi, g \circ \pi\}$, for any $f, g \in C^\infty(M/G)$, with $\pi : M \to M/G$ the orbit projection. Moreover, if M is actually symplectic with symplectic form ω, and the action has a coadjoint equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$, then the symplectic leaves of this Poisson structure are naturally symplectomorphic to the (connected components) of the ***Marsden-Weinstein*** [MW74] ***symplectic quotients*** $(M_\mu := \mathbf{J}^{-1}(\mu)/G_\mu, \omega_\mu)$, with $\mu \in \mathfrak{g}^*$ and G_μ the coadjoint isotropy of μ. The symplectic structure ω_μ on M_μ is uniquely determined by the expression $\pi_\mu^*\omega_\mu = i_\mu^*\omega$, with $i_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow M$ the injection and $\pi_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mu)/G_\mu$ the projection. See [AM78, OR04] and references therein for a general presentation of reduction theory.

Theorem 3.32 *Let $(M, \{\cdot,\cdot\}, h : M \to V^*, X)$ be a stochastic Hamiltonian system that is invariant with respect to the canonical, free , and proper action $\Phi : G \times M \to M$ of the Lie group G on M.*

(i) Poisson reduction*: The projection $h_{M/G}$ of the Hamiltonian function h onto M/G, uniquely determined by $h_{M/G} \circ \pi = h$, with $\pi : M \to M/G$ the orbit projection, induces a stochastic Hamiltonian system on the Poisson manifold $(M/G, \{\cdot,\cdot\}_{M/G})$ with stochastic component X and whose Stratonovich operator $H_{M/G} : TV \times M/G \to T(M/G)$ is given by*

$$H_{M/G}(v, \pi(z))(u) = T_z\pi\left(H(v,z)(u)\right) = \sum_{i=1}^{r} \langle \epsilon^i, u\rangle X_{h_i^{M/G}}(\pi(z)), \qquad u, v \in V \text{ and } z \in M. \tag{3.47}$$

In the previous expression $\{\epsilon^1, \ldots, \epsilon^r\}$ is a basis of V^*, $h_{M/G} = \sum_{i=1}^r h_i^{M/G}\epsilon^i$, and $h = \sum_{i=1}^r h_i\epsilon^i$; notice that the functions $h_i^{M/G} \in C^\infty(M/G)$ are the projections of the *components $h_i \in C^\infty(M)^G$, that is $h_i^{M/G} \circ \pi = h_i$. Moreover, if Γ is a solution semimartingale of the Hamiltonian system associated to H with initial condition Γ_0, then so is $\Gamma_{M/G} := \pi(\Gamma)$ with respect to $H_{M/G}$, with initial condition $\pi(\Gamma_0)$.*

(ii) ***Symplectic reduction**: Suppose that M is now symplectic and that the group action has a coadjoint equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$ associated. Then for any $\mu \in \mathfrak{g}^*$, the function $h_\mu : M_\mu := \mathbf{J}^{-1}(\mu)/G_\mu \to V^*$ uniquely determined by the equality $h_\mu \circ \pi_\mu = h \circ i_\mu$, induces a stochastic Hamiltonian system on the symplectic reduced space $(M_\mu := \mathbf{J}^{-1}(\mu)/G_\mu, \omega_\mu)$ with stochastic component X and whose Stratonovich operator $H_\mu : TV \times M_\mu \to TM_\mu$ is given by*

$$H_\mu(v, \pi_\mu(z))(u) = T_z\pi_\mu\left(H(v, i_\mu(z))(u)\right) = \sum_{i=1}^r \langle \epsilon^i, u\rangle X_{h_i^\mu}(\pi_\mu(z)), \tag{3.48}$$

$u, v \in V$ and $z \in \mathbf{J}^{-1}(\mu)$, where Remark 3.31 has been implicitly used. In the previous expression, the functions $h_i^\mu \in C^\infty(\mathbf{J}^{-1}(\mu)/G_\mu)$ are the coefficient functions in the linear combination $h_\mu = \sum_{i=1}^r h_i^\mu \epsilon^i$ and are related to the components $h_i \in C^\infty(M)^G$ of h via the relation $h_i^\mu \circ \pi_\mu = h_i \circ i_\mu$. Moreover, if Γ is a solution semimartingale of the Hamiltonian system associated to H with initial condition $\Gamma_0 \subset \mathbf{J}^{-1}(\mu)$, then so is $\Gamma_\mu := \pi_\mu(\Gamma)$ with respect to H_μ, with initial condition $\pi_\mu(\Gamma_0)$.

Remark 3.33 In the absence of freeness of the action the orbit spaces M/G and $\mathbf{J}^{-1}(\mu)/G_\mu$ cease to be regular quotient manifolds. Moreover, it could be that (even for free actions) there is no standard momentum map available (this is generically the case for Poisson manifolds). This situation can be handled by using the so called optimal momentum map [OR02] and its associated reduction procedure [O02]. Given that by part **(iii)** of Proposition 3.29 the fibers of the optimal momentum map are preserved by the Hamiltonian semimartingales associated to invariant Hamiltonians one can formulate, for any proper group action on a Poisson manifold, a theorem identical to part **(ii)** of Theorem 3.32 with the standard momentum map replaced by the optimal momentum map. In the particular case of a (non-necessarily free) symplectic proper action that has a standard momentum map associated, such result guarantees the good behavior of the symmetric stochastic Hamiltonian systems with respect to the singular reduced spaces in [SL91]; see also [OR06, OR06a] for the symplectic case without a standard momentum map.

Proof of Theorem 3.32. **(i)** can be proved by mimicking the proof of Theorem 3.9 by simply taking into account the fact that the G-invariance of h implies that of H and that for any $i \in \{1, \ldots, r\}$, one has that $T\pi \circ X_{h_i} = X_{h_i^{M/G}} \circ \pi$.

(ii) Expression (3.48) is guaranteed by the fact that $X_{h_i^\mu} \circ \pi_\mu = T\pi_\mu \circ X_{h_i} \circ i_\mu$, for any $i \in \{1, \ldots, r\}$ (see for instance [OR04, Theorem 6.1.1]). Let now Γ be a solution semimartingale of the Hamiltonian system associated to H with initial condition $\Gamma_0 \subset \mathbf{J}^{-1}(\mu)$. Notice first that by part **(ii)** in Proposition 3.29, $\Gamma \subset \mathbf{J}^{-1}(\mu)$ and hence the expression $\Gamma_\mu := \pi_\mu(\Gamma)$ is well defined. In order to prove the statement, we have to check that for any one-form $\alpha_\mu \in \Omega(M_\mu)$

$$\int \langle \alpha_\mu, \delta\Gamma_\mu\rangle = \int \langle H_\mu^*(X, \Gamma_\mu)\alpha_\mu, \delta X\rangle.$$

This equality follows in a straightforward manner from (3.48). Indeed,

$$\begin{aligned}\int\langle\alpha_\mu,\delta\Gamma_\mu\rangle &= \int\langle\alpha_\mu,\delta\left(\pi_\mu\circ\Gamma\right)\rangle=\int\langle\pi_\mu^*\alpha_\mu,\delta\Gamma\rangle\\ &= \int\langle H^*(X,\Gamma)\left(\pi_\mu^*\alpha_\mu\right),\delta X\rangle=\int\langle H_\mu^*(X,\Gamma_\mu)\alpha_\mu,\delta X\rangle,\end{aligned}$$

as required. ■

As to the reconstruction problem of solutions of a symmetric stochastic differential equation starting from a solution of the Poisson or symplectic reduced stochastic differential equation, Theorem 3.10 can be trivially modified to handle this situation. In the Poisson reduction case the theorem works without modification and when working with a solution of the symplectic reduced space it suffices to change the principal fiber bundle $\pi : M \to M/G$ by $\pi_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mu)/G_\mu$ all over.

3.6 Examples

3.6.1 Stochastic collective Hamiltonian motion

Our first example shows a situation in which the symplectic reduction of a symmetric stochastic Hamiltonian system offers, not only the advantage of cutting its dimension, but also of *making it into a deterministic system*. From the point of view of obtaining the solutions of the system, the procedures introduced in the previous section allow in this case the splitting of the problem into two parts: first, the solution of a standard ordinary differential equation for the reduced system and second, the solution of a stochastic differential equation in the group at the time of the reconstruction.

Let (M,ω) be a symplectic manifold, G a Lie group and $\Phi : G\times M\to M$ a free, proper, and canonical action. Additionally, suppose that this action has a coadjoint equivariant momentum map $\mathbf{J} : M\to\mathfrak{g}^*$ associated. Let $h_0\in C^\infty(M)^G$ be a G-invariant function and consider the deterministic Hamiltonian system with Hamiltonian function h_0.

A function of the form $f\circ\mathbf{J}\in C^\infty(M)$, for some $f\in C^\infty(\mathfrak{g}^*)$, is called ***collective***. We recall that by the Collective Hamiltonian Theorem (see for instance [MR99])

$$X_{f\circ\mathbf{J}}(z)=\left(\frac{\delta f}{\delta\mu}\right)_M(z),\quad z\in M,\ \mu=\mathbf{J}(z), \tag{3.49}$$

where the functional derivative $\frac{\delta f}{\delta\mu}\in\mathfrak{g}$ is the unique element such that for any $\nu\in\mathfrak{g}^*$, $Df(\mu)\cdot\nu=\langle\nu,\frac{\delta f}{\delta\mu}\rangle$. A straightforward consequence of (3.49) is that the G-invariant functions, in particular h_0, commute with the collective functions. Indeed, if $h\in C^\infty(M)^G$, then for any $z\in M$,

$$\{h,f\circ\mathbf{J}\}(z)=\mathbf{d}h(z)\cdot X_{f\circ\mathbf{J}}(z)=\mathbf{d}h(z)\cdot\left(\frac{\delta f}{\delta\mu}\right)_M(z)=0.$$

Collective functions play an important role to prove the complete integrability of certain dynamical systems (see [GS83]). Moreover, some relevant physical systems may be described using collective Hamiltonian functions. In that case, the (deterministic) equations of motion exhibit special features and, in some favorable cases, may be partially integrated using geometrical arguments (see [GS80]). The aim of this example is to study stochastic perturbations of deterministic symmetric mechanical systems introduced by means of collective Hamiltonians.

Let $Y : \mathbb{R}_+ \times \Omega \to \mathbb{R}^r$ be a $\mathbb{R}^r$-valued continuous semimartingale and $\{f_1, ..., f_r\} \subset C^\infty(\mathfrak{g}^*)$ a finite family of Ad^*_G-invariant functions on $\mathfrak{g}^*$. The coadjoint equivariance of the momentum map and the Ad^*_G-invariance of the functions allows us to construct the following G-invariant Hamiltonian function

$$\begin{array}{rcl} h : M & \longrightarrow & \mathbb{R} \times \mathbb{R}^r \\ m & \longmapsto & (h_0(m), (f_1(\mathbf{J}(m)), ..., f_r(\mathbf{J}(m)))) . \end{array}$$

Let X be the continuous semimartingale

$$\begin{array}{rcl} X : \mathbb{R}_+ \times \Omega & \longrightarrow & \mathbb{R}_+ \times \mathbb{R}^r \\ (t, \omega) & \longmapsto & (t, Y_t(\omega)) . \end{array}$$

Consider the stochastic Hamiltonian system (M, ω, h, X) which is, by construction, G-invariant. Noether's theorem (Proposition 3.29 **(ii)**) guarantees that the level sets of $\mathbf{J}$ are left invariant by the solution semimartingales of (M, ω, h, X). As to the reduction of this system, its main feature is that if we apply to it the reduction scheme introduced in Theorem 3.32 **(ii)**, for any $\mu \in \mathfrak{g}^*$, the reduced stochastic Hamiltonian system $(M_\mu, \omega_\mu, h_\mu, X)$ is such that

$$h_\mu \circ \pi_\mu = h_0 \circ i_\mu,$$

since $\mathbf{J}$, and hence the functions $f_i \circ \mathbf{J}$, are constant on the level sets $\mathbf{J}^{-1}(\mu)$, for any $i = 1, ..., r$. Consequently, the reduced system $(M_\mu, \omega_\mu, h_\mu, X)$ is equivalent to the deterministic Hamiltonian system $(M_\mu, \omega_\mu, h_\mu)$. In other words, the reduced system obtained from (M, ω, h, X) coincides with the one obtained in deterministic mechanics by symplectic reduction of $(M, h_0, t, \mathbb{R}_+)$. Thus, *we have perturbed stochastically a symmetric mechanical system preserving its symmetries and without changing the deterministic behavior of its corresponding reduced system.*

Remark 3.34 If we want to perturb the deterministic Hamiltonian system associated to h_0 with the only prescription that the level set $\mathbf{J}^{-1}(\mu)$ is left invariant, for a given value $\mu \in \mathfrak{g}^*$, we can weaken the requirement on the Ad^*_G-invariance of the functions $f_i \in C^\infty(\mathfrak{g}^*)$, $i = 1, ..., r$. Indeed, if we just ask that $\delta f_i / \delta \mu \in \mathfrak{g}_\mu$, we then have that $X_{h_0}(z), X_{f_1 \circ \mathbf{J}}(z), \ldots, X_{f_r \circ \mathbf{J}}(z) \in T_z \mathbf{J}^{-1}(\mu)$, for any $z \in \mathbf{J}^{-1}(\mu)$. The required invariance property follows then from (3.49) and Proposition 3.6.

Remark 3.35 In this example, the reduction-reconstruction scheme provides a global decomposition of the system (M, ω, h, X) into its deterministic and stochastic parts. If one is willing to work only locally, this splitting could be carried out without reduction in the neighborhood of any point in phase space, given that as $\{h_0, f_i \circ \mathbf{J}\} = 0$, for any $i \in \{1, \ldots, r\}$, then $[X_{h_0}, X_{f_i \circ \mathbf{J}}] = 0$.

3.6.2 Stochastic mechanics on Lie groups

The presence of mechanical systems whose phase space is the cotangent bundle of a Lie group is widespread. Besides the importance that this general case has in specific applications it is also very useful at the time of illustrating some of the theoretical developments in this chapter since most of the constructions that we presented admit very explicit characterizations. We start by recalling the main features of (deterministic) Hamiltonian systems over Lie groups. The reader interested in further details is encouraged to check with [AM78, MR99] and references therein.

Let G be a Lie group. The tangent bundle TG of G is trivial since it is isomorphic to the product $G \times \mathfrak{g}$, where $\mathfrak{g} = T_eG$ is the Lie algebra of G and $e \in G$ is the identity element. The identification $TG = G \times \mathfrak{g}$ is usually carried out by means of two isomorphisms, denoted by λ and ρ and induced by left and right translations on G, respectively. More specifically, let $\lambda : TG \to G \times \mathfrak{g}$ be the map given by $\lambda(v) = \left(g, T_gL_{g^{-1}}(v)\right)$, where $g = \tau_G(v)$ with $\tau_G : TG \to G$ the natural projection. On the other hand, $\rho : TG \to G \times \mathfrak{g}$ is defined by $\rho(v) = \left(g, T_gR_{g^{-1}}(v)\right)$. We refer to the image of λ as ***body coordinates*** and to the image of ρ as ***space coordinates***. The cotangent bundle T^*G is also trivial and isomorphic to $G \times \mathfrak{g}^*$. We can introduce ***body coordinates*** and ***space coordinates*** on T^*G by $\bar{\lambda}(\alpha) = (g, T_e^*L_g(\alpha)) \in G \times \mathfrak{g}^*$ and $\bar{\rho}(\alpha) = (g, T_e^*R_g(\alpha))$ respectively, where $g = \pi_G(\alpha)$ and $\pi_G : T^*G \to G$ is the canonical projection. The transition from body to space coordinates is as follows:

$$\begin{aligned}\left(\rho \circ \lambda^{-1}\right)(g,\xi) &= \rho\left(g, T_eL_g(\xi)\right) = \left(g, T_gR_{g^{-1}} \circ T_eL_g(\xi)\right) = \left(g, \mathrm{Ad}_g(\xi)\right)\\ \left(\bar{\rho} \circ \bar{\lambda}^{-1}\right)(g,\mu) &= \rho\left(g, T_g^*L_{g^{-1}}(\mu)\right) = \left(g, T_e^*R_g \circ T_g^*L_{g^{-1}}(\mu)\right) = \left(g, \mathrm{Ad}^*_{g^{-1}}(\mu)\right),\end{aligned}$$

for any $(g,\xi) \in G \times \mathfrak{g}$ and any $(g,\mu) \in G \times \mathfrak{g}^*$. The group action of G by left or right translations can be lifted to both TG and T^*G. We will denote by $\Phi_L : G \times TG \to TG$ and $\bar{\Phi}_L : G \times T^*G \to T^*G$ the lifted action of left translations on the tangent and cotangent bundle respectively, and by $\Phi_R : G \times TG \to TG$ and $\bar{\Phi}_R : G \times T^*G \to T^*G$ the lifted actions of right translations. The lifted actions have particularly simple expressions in suitable body or space coordinates. Indeed, it is more convenient to express Φ_L and $\bar{\Phi}_L$ in body coordinates, where for any $g, h \in G$, $\xi \in \mathfrak{g}$, and $\mu \in \mathfrak{g}^*$,

$$\begin{aligned}\left(\Phi_L\right)_g(h,\xi) &= \left(\lambda \circ TL_g \circ \lambda^{-1}\right)(h,\xi) = (gh,\xi),\\ \left(\bar{\Phi}_L\right)_g(h,\mu) &= \left(\bar{\lambda} \circ T^*L_{g^{-1}} \circ \bar{\lambda}^{-1}\right)(h,\mu) = \left(g^{-1}h,\mu\right).\end{aligned}$$

As to Φ_R and $\bar{\Phi}_R$, space coordinates are particularly convenient; for any $g, h \in G$, $\zeta \in \mathfrak{g}$, and $\alpha \in \mathfrak{g}^*$,

$$\begin{aligned}\left(\Phi_R\right)_g(h,\zeta) &= \left(\rho \circ TR_g \circ \rho^{-1}\right)(h,\zeta) = (hg,\zeta)\\ \left(\bar{\Phi}_R\right)_g(h,\alpha) &= \left(\bar{\rho} \circ T^*R_{g^{-1}} \circ \bar{\rho}^{-1}\right)(h,\alpha) = \left(hg^{-1},\alpha\right).\end{aligned}$$

The actions $\bar{\Phi}_L$ and $\bar{\Phi}_R$, being the cotangent lifted actions to T^*G of an action on G, have canonical momentum maps $\mathbf{J}_L : T^*G \to \mathfrak{g}^*$ and $\mathbf{J}_R : T^*G \to \mathfrak{g}^*$, respectively, when we endow T^*G with its canonical symplectic form. Let $\theta \in \Omega^1(T^*G)$ be the Liouville canonical one-form on T^*G. Then, $\mathbf{J}_L$ and $\mathbf{J}_R$ are given by

$$\langle \mathbf{J}_L(z_g), \xi \rangle = \langle z_g, (\xi)^L_G(g) \rangle, \qquad \langle \mathbf{J}_R(z_g), \xi \rangle = \langle z_g, (\xi)^R_G(g) \rangle,$$

for any $z_g \in T_g^*G$ and any $\xi \in \mathfrak{g}$. Here $(\xi)_G^L \in \mathfrak{X}(G)$ (respectively $(\xi)_G^R \in \mathfrak{X}(G)$) denotes the infinitesimal generator associated to $\xi \in \mathfrak{g}$ by the left (respectively right)action of G on itself. This expression clearly shows that $\mathbf{J}_L$ is right-invariant and $\mathbf{J}_R$ left-invariant. Observe that $\mathbf{J}_L = \mathrm{Ad}_{g^{-1}}^* \circ \mathbf{J}_R$. For example, in body coordinates, these momentum maps have the following expressions ([AM78, Theorem 4.4.3])

$$(\mathbf{J}_L)_B\left((g,\mu)\right) = \mathrm{Ad}_{g^{-1}}^*(\mu) \quad \text{and} \quad (\mathbf{J}_R)_B\left((g,\mu)\right) = \mu. \tag{3.50}$$

In this context, the classical results on symplectic and Poisson reduction that we have described in the previous section admit a particularly explicit formulation. In all that follows we will suppose that the action with respect to which we are reducing is lifted left translations. Using body coordinates, it is easy to see that in this case the Poisson reduced space T^*G/G coincides with the dual of the Lie algebra $\mathfrak{g}^*$ endowed with the ***Lie-Poisson*** structure given by

$$\{f_1, f_2\}_-^{\mathfrak{g}^*}(\mu) = -\left\langle \mu, \left[\frac{\delta f_1}{\delta \mu}, \frac{\delta f_2}{\delta \mu}\right]\right\rangle,$$

for any $\mu \in \mathfrak{g}^*$ and $f_1, f_2 \in C^\infty(\mathfrak{g}^*)$. The symplectic reduced spaces $\mathbf{J}_L^{-1}(\mu)/G_\mu$ are naturally symplectomorphic to the symplectic leaves of the Lie-Poisson structure on $\mathfrak{g}^*$, that is, the coadjoint orbits endowed with the so-called ***Kostant-Kirillov-Souriau*** symplectic form ω_μ^-:

$$\omega_\mu^-(\mu)(\xi_{\mathfrak{g}^*}(\mu), \eta_{\mathfrak{g}^*}(\mu)) = \omega_\mu^-(\mu)(-\mathrm{ad}_\xi^*\mu, -\mathrm{ad}_\eta^*\mu) = -\langle \mu, [\xi, \eta]\rangle.$$

Let now V be a vector space, $X : \mathbb{R}_+ \times \Omega \to V$ a continuous semimartingale, and $h : T^*G \to V^*$ a smooth map invariant under the lifted left translations of G on T^*G. If we use body coordinates and we visualize T^*G as the product $G \times \mathfrak{g}^*$, the invariance of $h : G \times \mathfrak{g}^* \to V^*$ allows us to write it as $h = \sum_{i=1}^r h_i \epsilon^i$, where $\{\epsilon^1, \ldots, \epsilon^r\}$ is a basis of V^* and $h_1, \ldots, h_r \in C^\infty(\mathfrak{g}^*)$. Let $\{e_1, \ldots, e_r\}$ be the dual basis of V and write $X = \sum_{i=1}^r X^i e_i$. Using the left trivialized expression of the Hamiltonian vector fields in the deterministic case (see [OR04, Theorem 6.2.5]) it is easy to see that the stochastic Hamiltonian equations in this setup are

$$\delta\Gamma^h = \sum_{i=1}^r \left(T_e L_{\Gamma^G}\left(\frac{\delta h_i}{\delta \Gamma^{\mathfrak{g}^*}}\right), \mathrm{ad}^*_{\frac{\delta h_i}{\delta \Gamma^{\mathfrak{g}^*}}} \Gamma^{\mathfrak{g}^*}\right) \delta X^i \tag{3.51}$$

where Γ^G and $\Gamma^{\mathfrak{g}^*}$ are the G and $\mathfrak{g}^*$ components of Γ^h, respectively, that is, $\Gamma^h := \left(\Gamma^G, \Gamma^{\mathfrak{g}^*}\right)$. In the left trivialized representation, the reduced Poisson and symplectic Hamiltonians are simply the restrictions $h^{\mathfrak{g}^*}$ and $h^{\mathcal{O}_\mu}$ of h to $\mathfrak{g}^*$ and to the coadjoint orbits $\mathcal{O}_\mu \subset \mathfrak{g}^*$, respectively. Additionally, the reduced stochastic Hamilton equations on $\mathfrak{g}^*$ and $\mathcal{O}_\mu$ are given by

$$\delta\Gamma^{\mathfrak{g}^*} = \sum_{i=1}^r \mathrm{ad}^*_{\frac{\delta h_i^{\mathfrak{g}^*}}{\delta \Gamma^{\mathfrak{g}^*}}} \Gamma^{\mathfrak{g}^*} \delta X^i \quad \text{and} \quad \delta\Gamma^{\mathcal{O}_\mu} = \sum_{i=1}^r \mathrm{ad}^*_{\frac{\delta h_i^{\mathcal{O}_\mu}}{\delta \Gamma^{\mathcal{O}_\mu}}} \Gamma^{\mathcal{O}_\mu} \delta X^i \tag{3.52}$$

where $h^{\mathfrak{g}^*} = \sum_{i=1}^r h_i^{\mathfrak{g}^*} \epsilon^i$ and $h^{\mathcal{O}_\mu} = \sum_{i=1}^r h_i^{\mathcal{O}_\mu} \epsilon^i$.

The combination of expressions (3.51) and (3.52) shows that in this setup, the dynamical reconstruction of reduced solutions is particularly simple to write down. Indeed, suppose that we are given a solution $\Gamma^{\mathfrak{g}^*}$ of, say, the Poisson reduced system. In order to obtain the solution Γ^h of the original system such that $\Gamma_0^h = (\Gamma_0^G, \Gamma_0^{\mathfrak{g}^*})$ and $\pi(\Gamma^h) = \Gamma^{\mathfrak{g}^*}$, with $\pi : T^*G \simeq G \times \mathfrak{g}^* \to T^*G/G \simeq \mathfrak{g}^*$ the Poisson reduction projection, it suffices to solve the stochastic differential equation in G

$$\delta\Gamma^G = \sum_{i=1}^r T_e L_{\Gamma^G}\left(\frac{\delta h_i}{\delta\Gamma^{\mathfrak{g}^*}}\right)\delta X^i, \tag{3.53}$$

with the initial condition Γ_0^G. The reconstructed solution that we are looking for is then $\Gamma^h = \left(\Gamma^G, \Gamma^{\mathfrak{g}^*}\right)$.

3.6.3 Stochastic perturbations of the free rigid body

The free rigid body, also referred to as Euler top, is a particular case of systems introduced in the previous section where the group G is $SO(3,\mathbb{R})$. We recall that in the context of mechanical systems on groups, a Hamiltonian system is called *free* when the energy of the system is purely kinetic and there is no potential term. Let $(\cdot,\cdot)$ be a left invariant Riemannian metric on G; the kinetic energy E associated to $(\cdot,\cdot)$ is $E(v) = \frac{1}{2}(v,v)$, $v \in TG$. Then, using the left invariance of the metric, we can write in body coordinates

$$E(g,\xi) = \frac{1}{2}(\xi,\xi)_e = \frac{1}{2}\langle I(\xi),\xi\rangle,$$

for any $(g,\xi) \in G \times \mathfrak{g}$, where $\langle\cdot,\cdot\rangle$ is the natural pairing between elements of $\mathfrak{g}^*$ and $\mathfrak{g}$, and $I : \mathfrak{g} \to \mathfrak{g}^*$ is the map given by $\xi \longmapsto (\xi,\cdot)_e$ and usually known as the ***inertia tensor*** associated to the metric $(\cdot,\cdot)$. The Legendre transformation associated to E can be used to define a Hamiltonian function $h : T^*G \to \mathbb{R}$ that, in body coordinates, can be written as

$$h(g,\mu) = \frac{1}{2}\langle\mu,\Lambda(\mu)\rangle, \tag{3.54}$$

where $\Lambda = I^{-1}$. Notice that as the kinetic energy is left invariant (invariant with respect to the lifted G-action to T^*G of the action of G on itself by left translations), then the components of $\mathbf{J}_L$ are conserved quantities of the corresponding Hamiltonian system. In order to connect with example in Section 3.6.1, let $f \in C^\infty(\mathfrak{g}^*)$ be the function $f : \mathfrak{g}^* \to \mathbb{R}$ given by $\mu \mapsto \frac{1}{2}\langle\mu,\Lambda(\mu)\rangle$. By (3.50), the Hamiltonian function h may be expressed as $h = f \circ \mathbf{J}_R$. Therefore h is collective with respect to $\mathbf{J}_R$.

We now go back to the free rigid body case, that is, $G = SO(3,\mathbb{R})$. We recall that the Lie algebra $\mathfrak{so}(3,\mathbb{R})$ is the vector space of three dimensional skew-symmetric real matrices whose bracket is just the commutator of two matrices. As a Lie algebra, $(\mathfrak{so}(3),[\cdot,\cdot])$ is naturally isomorphic to $\left(\mathbb{R}^3,\times\right)$, where $\times$ denotes the cross product of vectors in $\mathbb{R}^3$. Under this isomorphism, the adjoint representation of $SO(3,\mathbb{R})$ on its Lie algebra is simply the action of matrices on vectors of $\mathbb{R}^3$ and the Lie-Poisson structure on $\mathfrak{so}(3)^* \simeq \mathbb{R}^3$ is given by $\{f,g\}(v) = -v\cdot(\nabla f \times \nabla g)$, for any $f,g \in C^\infty(\mathbb{R}^3)$, where ∇ is the usual Euclidean gradient and $\cdot$ denotes the Euclidean inner product.

Given a free rigid body with inertia tensor $I : \mathbb{R}^3 \to \mathbb{R}^3$, since $\delta h_B/\delta\mu = \Lambda(\mu)$, for any $\mu \in \mathbb{R}^3$, the left-trivialized equations of motion of the system are

$$(\dot{A},\dot{\mu}) = \left(A\cdot\widehat{\Lambda(\mu)}, \mu \times \Lambda(\mu)\right), \tag{3.55}$$

where the dot in the right hand side of (3.55) stands for matrix multiplication and $\widehat{\Lambda(\mu)}$ is the skew-symmetric matrix associated to $\Lambda(\mu) \in \mathbb{R}^3$ via the mapping that implements the Lie algebra isomorphism between $(\mathfrak{so}(3), [\cdot, \cdot])$ and $(\mathbb{R}^3, \times)$. In the context of the free rigid body motion the momentum map $\mathbf{J}_L$ (respectively, $\mathbf{J}_R$) is called ***spatial angular momentum*** (respectively, ***body angular momentum***). The second component of (3.55), that is,

$$\dot{\mu} = \mu \times \Lambda(\mu) \tag{3.56}$$

are the well-known ***Euler equations*** for the free rigid body.

Random perturbations of the body angular momentum. We now introduce stochastic perturbations of the free rigid body by using some of the geometrical tools that we have introduced above. Later on we will compare this example with the model of the randomly perturbed rigid body studied in [L97] and [LW05], whose physical justification, as we will briefly discuss, involves the same ideas as ours.

Let $V = \mathbb{R} \times \mathfrak{so}(3) \simeq \mathbb{R}^+ \times \mathbb{R}^3$ and let h be the Hamiltonian function $h : T^*SO(3) \to V^* = \mathbb{R} \times \mathfrak{so}(3)^*$ defined as $h = (h_0, \mathbf{J}_R)$, where h_0 is the Hamiltonian function of the free (deterministic) rigid body. Observe that h is a left-invariant function because so is $\mathbf{J}_R$. Let $Y : \mathbb{R}_+ \times \Omega \to \mathfrak{g}$ be a continuous semimartingale which we may suppose, for the sake of simplicity, is a $\mathfrak{g}$-valued Brownian motion and let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^* \times \mathfrak{g}$ be the semimartingale defined as $X_t(\omega) = (t, Y_t(\omega))$ for any $(t, \omega) \in \mathbb{R} \times \Omega$. Consider the stochastic Hamiltonian system on T^*G associated to h and X. Since h is left invariant, the momentum map $\mathbf{J}_L$ is preserved by the solution semimartingales of this system and moreover, we can apply the reduction scheme introduced in the previous sections. For example, if we carry out Poisson reduction we have a reduced Hamiltonian function $h^{\mathfrak{g}^*} : \mathfrak{g}^* \to V^*$ given by $h^{\mathfrak{g}^*}(\mu) = \left(\frac{1}{2}\langle \mu, \Lambda(\mu)\rangle, \mu\right)$. Let $\{\xi_1, \xi_2, \xi_3\}$ a basis of the Lie algebra $\mathfrak{g}$ and $\{\epsilon^1, \epsilon^2, \epsilon^3\} \subset \mathfrak{g}^*$ its dual basis. Observe that if we write $\mathbf{J}_R(\mu) = \sum_{i=1}^3 \langle \mu, \xi_i\rangle \epsilon^i$ and $Y = \sum_{i=1}^3 Y^i \xi_i$, then the reduced stochastic Lie-Poisson equations can be expressed as

$$\delta\mu_t = \mu_t \times \Lambda(\mu_t)\,\delta t + \sum_{i=1}^3 (\mu_t \times \xi_i)\,\delta Y_t^i. \tag{3.57}$$

Regarding the reconstruction of the reduced dynamics, one has to solve the stochastic differential equation on the rotations group $SO(3)$ given by (3.53) that, in this particular case, is given by

$$\delta A_t = A_t \cdot \widehat{\Lambda(\mu_t)}\delta t + \sum_{i=1}^3 A_t \cdot \widehat{\xi}_i \delta Y_t^i. \tag{3.58}$$

A physical model whose description fits well in a stochastic Hamiltonian differential equation like the one associated to h and X is that of a free rigid body subjected to small random impacts. Each impact causes a small and instantaneous change in the *body* angular momenta μ_t at time t that justifies the extra term in (3.57), when compared to the Euler equations (3.56).

Our model is very similar to the one proposed in [L97] where, instead of introducing the random perturbation by means of a Hamiltonian function, a stochastic differential equation on the group G is introduced. This equation, also studied in detail in [LW05], is

$$\delta A_t = A_t \cdot \Lambda \cdot \mathrm{Ad}^*_{A_t}(\alpha)\,\delta t + \sum_{i=1}^{3}\left(A_t \cdot \Lambda \cdot \mathrm{Ad}^*_{A_t}\left(\epsilon^i\right)\right)\delta Y^i, \tag{3.59}$$

where $\alpha \in \mathfrak{g}^*$ is a constant vector. It important to note that the drift terms of equations (3.58) and (3.59) coincide. Indeed, for any $(g,\mu) \in G \times \mathfrak{g}^*$ we can write

$$\mu = \mathrm{Ad}^*_g \circ \mathrm{Ad}^*_{g^{-1}}(\mu) = \mathrm{Ad}^*_g\left(\mathbf{J}_L(g,\mu)\right).$$

Since in our model the spatial angular momentum is conserved, $\Lambda(\mu_t) = \Lambda\left(\mathrm{Ad}^*_{A_t}\left(\mathrm{Ad}^*_{A_t^{-1}}\mu_t\right)\right) = \Lambda\left(\mathrm{Ad}^*_{A_t}\left(\mathbf{J}_L(A_t,\mu_t)\right)\right) = \Lambda\left(\mathrm{Ad}^*_{A_t}(\alpha)\right)$, where $\alpha = \mathbf{J}_L(A_t,\mu_t)$ is the preserved value of the spatial angular momentum of a solution (A_t,μ_t) of (3.57) and (3.58). The difference between (3.58) and (3.59) lies in the stochastic terms. The justification given by the author in [L97] for the equation (3.59) is the following: since in the (deterministic) rigid body the spatial angular momentum $\mathbf{J}_L$ is conserved, once we have fixed the value of this conserved quantity, we can simply study the dynamics of the free rigid body by looking at the first component of the ordinary differential equation (3.55), now rewritten as

$$\dot{A} = A\left(\Lambda\left(\mathrm{Ad}^*_A(\alpha)\right)\right) \tag{3.60}$$

where $\alpha \in \mathfrak{g}^*$ is the $\mathbf{J}_L$-value of the solution. Under random impacts, the spatial angular momentum α, which was preserved in the deterministic case, is now randomly modified. The idea is then to replace αdt in (3.60) by $\alpha\delta t + \sum_{i=1}^{3}\epsilon^i\delta Y^i$. Unlike our model, where the random perturbation is introduced in the cotangent bundle respecting the underlying symmetries of the deterministic system, there is no preservation of $\mathbf{J}_L$ in the stochastic model of [L97].

One advantage of working on T^*G is that, even in the stochastic context, classical quantities such as the angular momentum, are still well defined. These objects do not have a clear counterpart if one follows the configuration space based approach in [L97] (see for instance [LW05] for a non-trivial definition of angular velocity in the stochastic context).

Not so rigid rigid bodies. Random perturbation of the inertia tensor. In this example we want to write the equations that describe a rigid body some of whose parts are slightly loose, that is, the body is not a true rigid body and hence its mass distribution is constantly changing in a random way. This will be modelled by stochastically perturbing the tensor of inertia.

For the sake of simplicity, we will write $G = SO(3,\mathbb{R})$ and $\mathfrak{g} = \mathfrak{so}(3)$. Let $\mathcal{L}(\mathfrak{g}^*,\mathfrak{g})$ be the vector space of linear maps from $\mathfrak{g}^*$ to $\mathfrak{g}$. As we know $(\mathfrak{so}(3),[\cdot,\cdot]) \simeq \left(\mathbb{R}^3,\times\right)$. Furthermore, we can establish an isomorphism $\mathbb{R}^3 \simeq \left(\mathbb{R}^3\right)^*$ using the Euclidean inner product and hence we can write $\mathfrak{g} \simeq \mathfrak{g}^*$. Let $V = \mathcal{L}_S(\mathfrak{g}^*,\mathfrak{g}) = \{A \in \mathcal{L}(\mathfrak{g}^*,\mathfrak{g}) \mid A^* = A\}$ be the vector space of self-adjoint linear maps from $\mathfrak{g}^*$ to $\mathfrak{g}$. Define the Hamiltonian $h : T^*G \to V^*$ in *body coordinates* as

$$\begin{array}{rcl} h : T^*G \simeq G \times \mathfrak{g}^* & \longrightarrow & V^* \\ (g,\mu) & \longmapsto & \bar{\mu}, \end{array}$$

where $\bar{\mu}$ is such that

$$\begin{array}{rcl} \bar{\mu} : \mathcal{L}_S\left(\mathfrak{g}^*, \mathfrak{g}\right) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & \frac{1}{2}\left\langle \mu, A\left(\mu\right)\right\rangle . \end{array}$$

Observe that in body coordinates the Hamiltonian h does not depend on G, so the Hamiltonian is G-invariant by the action $\bar{\Phi}_L$ on T^*G. On the other hand, consider some filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}}, P\right)$ and introduce a stochastic component $X : \mathbb{R}_+ \times \Omega \to V$ in the following way:

$$\begin{array}{rcl} X : \mathbb{R}_+ \times \Omega & \longrightarrow & \mathcal{L}_S\left(\mathfrak{g}^*, \mathfrak{g}\right) \\ (t, \omega) & \longmapsto & \Lambda t + \varepsilon A_t\left(\omega\right), \end{array}$$

where $\Lambda \in \mathcal{L}_S\left(\mathfrak{g}^*, \mathfrak{g}\right)$ plays the role of the inverse of the tensor of inertia given by the deterministic (rigid) description of the body, ε is a small parameter, and A is an arbitrary $\mathcal{L}_S\left(\mathfrak{g}^*, \mathfrak{g}\right)$-valued semimartingale. In order to show how the stochastic Hamiltonian system on T^*G associated to h and X models a free rigid body whose inertia tensor undergoes random perturbations, we write down the associated stochastic reduced Lie-Poisson equations in Stratonovich form

$$\delta\mu_t = \mu_t \times \Lambda\left(\mu_t\right)\delta t + \varepsilon\mu_t \times \delta A_t\left(\mu_t\right).$$

Thus we see that these Lie-Poisson equations consist in changing $\Lambda\left(\mu_t\right)dt$ in the Euler equations (3.56) by $\Lambda\left(\mu_t\right)\delta t + \varepsilon\delta A_t\left(\mu_t\right)$, which accounts for the stochastic perturbation of the inertia tensor.

4

Superposition rules and stochastic Lie-Scheffers systems

A differential equation is said to have a ***superposition rule*** (a more explicit definition is provided in the next section) whenever any of its solutions can be written as a given (in general nonlinear) function of the initial condition and of a fixed set of particular solutions. The first characterization of the existence of superposition rules was given by the Norwegian mathematician Sophus Lie in a remarkable piece of work [Lie93] where he established a link between the existence of superposition rules and what we nowadays call the Lie algebraic properties of the vector fields that define a time-dependent differential equation. This result is referred to as the ***Lie-Scheffers Theorem*** and systems that satisfy its hypotheses as ***Lie-Scheffers systems***.

Lie-Scheffers systems have been the subject of much attention due to their widespread occurrence in physics and mathematics. The reader is encouraged to check with [CGM00, CGM07], and references therein, for various presentations of the classical Lie-Scheffers Theorem, an excellent collection of examples of applications of this theorem, and for historical remarks.

The main goal of this chapter is the extension of the Lie-Scheffers Theorem to stochastic differential equations. This generalization is stated in Theorem 4.7. It is worth emphasizing that the main result of the chapter, Theorem 4.7, cannot be seen just as a mere transcription of the deterministic Lie-Scheffers Theorem into the context of Stratonovich stochastic integration by using the so called Malliavin's Transfer Principle [Ma78]. As we will see later on, there are purely stochastic conditions that appear in the statement of the theorem.

Additionally, in proving Theorem 4.7 we have carefully spelled out the regularity conditions needed for the result to be valid; those conditions are only vaguely evoked in the classical references or in the cited papers that study the deterministic case. More importantly, a careful construction of the proof has lead us to realize that the hypotheses under which we can guarantee the existence of superposition rules can be weakened: the Lie algebra condition in the classical theorem can be replaced by an involutivity hypothesis that is, in general, less restrictive.

The contents of the chapter are structured as follows. Section 4.1 explains in detail the notion of superposition rule and includes a proposition that translates this concept into geometric terms. Section 4.2 contains the main theorem that we have already described.

Section 4.3 is dedicated to the study of Lie-Scheffers systems on Lie groups and homogeneous spaces; this case is particularly relevant since, as we show in the first result of that section (Proposition 4.12), classical Lie-Scheffers systems (roughly speaking, those generated by vector fields that close a Lie algebra) can be locally reduced to this case via a theorem due to Palais. In that section we also show, as an example, how Lévy stochastic processes can be seen as Lie group valued Lie-Scheffers systems. The section concludes with a brief presentation of the classical Wei-Norman method for solving Lie-Scheffers systems, adapted to the stochastic context.

Section 4.4 contains a discussion on how the existence of a superposition rule for a stochastic differential equation makes available a remarkable feature that has deserved certain attention in the context of standard stochastic differential equations, namely, the fact that the stochastic flow can be written as a fixed deterministic function of the Brownian forcing of the equation in question. Indeed, a well know theorem by Ben Arous [B89], that we state in this thesis and whose proof is based on the use of stochastic Taylor expansions, shows that this property of the flow is available under exactly the same hypotheses as the classical Lie-Scheffers Theorem. Our main theorem allows, admittedly only to a certain extent, the generalization of this statement to any stochastic differential equation that satisfies its hypotheses; more specifically, any SDE generated by vector fields that span an involutive distribution has a superposition rule and hence its flow can be written as a fixed deterministic function of the initial conditions and of a set of solutions that contain the stochastic behavior of the resulting map.

This chapter, a transcription of the paper [LO08a] witten by the author of this thesis in collaboration with J.-P. Ortega, concludes with a section that contains a number of examples that illustrate our developments.

4.1 Superposition rules for stochastic differential equations

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We start by considering the stochastic differential equation

$$\delta\Gamma = S\left(X,\Gamma\right)\delta X, \tag{4.1}$$

where $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^l$ is a given $\mathbb{R}^l$-valued semimartingale and $S\left(x,z\right) : T_x\mathbb{R}^l \longrightarrow T_z\mathbb{R}^n$ is a Stratonovich operator from $\mathbb{R}^l$ to $\mathbb{R}^n$. Sometimes we will choose a basis in $T^*\mathbb{R}^l$ and will write down the Stratonovich operator $S(x,z)$ in terms of its components $(S_1\left(x,z\right), \ldots, S_l(x,z))$ with respect to that basis.

Definition 4.1 *A **superposition rule** of the stochastic differential equation (4.1) is a pair $(\Phi, \{\Gamma_1, \ldots, \Gamma_m\})$, where $\Phi : \mathbb{R}^{n(m+1)} \longrightarrow \mathbb{R}^n$ is a (not necessarily smooth) function and $\{\Gamma_i : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n \mid i = 1, \ldots, m\}$ is a set of particular solutions of (4.1) such that any solution Γ of (4.1) can be written, at least up to a sufficiently small stopping time τ, as*

$$\Gamma = \Phi\left(z^1, \ldots, z^n; \Gamma_1, \ldots, \Gamma_m\right) =: \Phi\left(z; \Gamma_1, \ldots, \Gamma_m\right),$$

where $z = \left(z^1, \ldots, z^n\right)$ is a set of n arbitrary constants associated with the initial condition of the solution Γ, that is, $\Gamma(0,\omega) = (z^1, \ldots, z^n)$, for all $\omega \in \Omega$. We extend to the stochastic context the terminology used for standard differential equations and we will call ***Lie-Scheffers*** *systems the stochastic differential equations that admit a superposition rule.*

Remark 4.2 As we will see in examples later on in the paper, superposition rules exist only locally. That is why we can, without loss of generality, restrict our attention to stochastic differential equation on Euclidean spaces. Observe also that we are requiring that Φ does not depend on time, the probability space, or the noise X. This prevents us from using certain regularization techniques at the time of testing the existence of superposition rules. For example, when dealing with a deterministic differential equation, the standard transformation of a time-dependent system $\dot{\gamma} = f\left(t, \gamma\right)$ on $\mathbb{R}^n$, $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ into the autonomous one

$$\dot{\gamma} = f\left(t, \gamma\right) \quad \text{and} \quad \dot{t} = 1$$

on $\mathbb{R}^{n+1}$ obtained by adding an extra trivial differential equation for the time, is not allowed; indeed, if we find a superposition rule for the transformed autonomous system, that rule does not yield a superposition rule for the original system that satisfies the requirements of our definition, precisely due to the explicit dependence on time that appears in the superposition function.

In order to study the implications of the presence of a superposition rules we take a more geometric approach. Let Ψ be the function defined by

$$\begin{array}{rcl} \Psi : \mathbb{R}^{n(m+2)} & \longrightarrow & \mathbb{R}^n \\ (z, q_0, q_1, \ldots, q_m) & \longmapsto & q_0 - \Phi\left(z; q_1, \ldots, q_m\right). \end{array} \tag{4.2}$$

Notice that for any $z \in \mathbb{R}^n$, the function $\Psi_z := \Psi\left(z, \cdot\right) : \mathbb{R}^{n(m+1)} \to \mathbb{R}^n$ is constant on a $(m+1)$-tuple $(\Gamma, \Gamma_1 \ldots, \Gamma_m)$ of solutions of the system (4.1), at least up to a given stopping time τ, provided that $\Gamma_{t=0} = z \in \mathbb{R}^n$ a.s.. From now on we assume that all the solutions Γ that we are dealing with are constant a.s. at $t = 0$. Additionally, if the function Φ is smooth then the map $\Psi_z : \mathbb{R}^{n(m+1)} \to \mathbb{R}^n$ is a submersion for any fixed $z \in \mathbb{R}^n$, because

$$\operatorname{rank}\left(\frac{\partial \Psi_z^j}{\partial q_0^i}\right)_{j,i=1,\ldots,n} = \operatorname{rank}\left(I_n\right) = n \tag{4.3}$$

where I_n is the identity matrix of dimension n. Consequently, for any $z \in \mathbb{R}^n$, the level set $\Psi_z^{-1}\left(0\right) \subset \mathbb{R}^{n(m+1)}$ is a closed embedded submanifold of $\mathbb{R}^{n(m+1)}$ of dimension nm . That is, the function Ψ defines a family $\mathcal{G}$ of regular nm-dimensional submanifolds $\mathcal{G}_z$ via the zero level sets $\Psi_z^{-1}\left(0\right) = \{p \in \mathbb{R}^{n(m+1)} \mid \Psi\left(z, p\right) = 0\} =: \mathcal{G}_z$ of Ψ_z, for any $z \in \mathbb{R}^n$. The submanifolds $\mathcal{G}_z$ are globally diffeomorphic to $\mathbb{R}^{nm}$ via the restriction $\pi_m|_{\mathcal{G}_z}$ to $\mathcal{G}_z$ of the projection $\pi_m : \mathbb{R}^{n(m+1)} = \mathbb{R}^n \times \overset{m+1}{\cdots} \times \mathbb{R}^n \longrightarrow \mathbb{R}^{nm} = \mathbb{R}^n \times \overset{m}{\cdots} \times \mathbb{R}^n$ onto the last m $\mathbb{R}^n$ factors. This is easy to see by verifying that the inverse $\Xi_z : \mathbb{R}^{mn} \to \mathcal{G}_z$ of $\pi_m|_{\mathcal{G}_z}$ is given by $\Xi_z(q_1, \ldots, q_m) = (\Phi(z; q_1, \ldots, q_m), q_1, \ldots, q_m)$, which is obviously a diffeomorphism. In order to study the significance of the family of submanifolds $\mathcal{G}$ we start by introducing the following definition.

Definition 4.3 *Let $Y : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field. The vector field*

$$\begin{array}{rcl} \widetilde{Y} : \mathbb{R}^{n(m+1)} & \longrightarrow & \mathbb{R}^{n(m+1)} \\ (q_0, \ldots, q_m) & \longmapsto & (Y(q_0), \ldots, Y(q_m)) \end{array}$$

is called the ***diagonal extension*** *of Y.*

It can be easily checked that the set of diagonal extensions of vector fields in $\mathfrak{X}(\mathbb{R}^n)$ are a subalgebra of $\mathfrak{X}(\mathbb{R}^{n(m+1)})$; more explicitly, for any $Y_1, Y_2, Y_3 \in \mathfrak{X}(\mathbb{R}^n)$ and $\lambda \in \mathbb{R}$,

$$[\widetilde{Y}_1, \widetilde{Y}_2 + \lambda \widetilde{Y}_3] = \widetilde{[Y_1, Y_2 + \lambda Y_3]}. \tag{4.4}$$

The following proposition states that, roughly speaking, the family of submanifolds $\mathcal{G}$ completely characterizes the superposition rule.

Proposition 4.4 *Suppose that the stochastic differential equation (4.1) admits a smooth superposition rule $(\Phi, \{\Gamma_1, \ldots, \Gamma_m\})$. Suppose that $(\Gamma_1, \ldots, \Gamma_m)_{t=0} = (p_1, \ldots, p_m) \in \mathbb{R}^{mn}$ a.s.. Then, there exists a family $\mathcal{G}$ of closed embedded nm-dimensional submanifolds of $\mathbb{R}^{n(m+1)}$ such that for any $z \in \mathbb{R}^n$ there exists $\mathcal{G}_z \in \mathcal{G}$ such that $(\Gamma^z, \Gamma_1, \ldots, \Gamma_m) \subset \mathcal{G}_z$, with Γ^z the solution of (4.1) such that $(\Gamma^z)_{t=0} = z$. Moreover, for any $\mathcal{G}_z \in \mathcal{G}$ the map $\pi_m|_{\mathcal{G}_z} : \mathcal{G}_z \to \mathbb{R}^{nm}$ is a diffeomorphism.*

Conversely, let $\mathcal{G}$ be a family of (not necessarily embedded) submanifolds of $\mathbb{R}^{n(m+1)}$ diffeomorphic to $\mathbb{R}^{nm}$ via π_m and $\{\Gamma_1, \ldots, \Gamma_m\}$ a set of distinct solutions of (4.1) such that $(\Gamma_1, \ldots, \Gamma_m)_{t=0} = (p_1, \ldots, p_m) \in \mathbb{R}^{mn}$ a.s.. Then, if for any point $z \in \mathbb{R}^n$ there is an element $\mathcal{G}_z$ that contains the point $(z, p_1, \ldots, p_m)$ and the diagonal extensions $(\widetilde{S}_1(X, \cdot), \ldots, \widetilde{S}_l(X, \cdot))$ of the vector fields $(S_1(X, \cdot), \ldots, S_l(X, \cdot))$ that define (4.1) are tangent to $\mathcal{G}_z$ when evaluated at $(\Gamma^z, \Gamma_1, \ldots, \Gamma_m)$, then (4.1) admits a (possibly nonsmooth) superposition rule.

Proof. In view of the remarks preceding Definition 4.3 we just need to prove that having a family $\mathcal{G}$ that satisfies the hypotheses in the statement allows us to recover the superposition rule.

Let $\{\Gamma_1, \ldots, \Gamma_m\}$ be the set of fixed distinct solutions of (4.1). Denote $p_i = (\Gamma_i)_{t=0}$ the (necessarily different) constant initial conditions of Γ_i, $i = 1, \ldots, m$. Let $z = (z^1, \ldots, z^n) \in \mathbb{R}^n$ be a point and let $\mathcal{G}_z$ be the submanifold in $\mathcal{G}$ such that $(z, p_1, \ldots, p_m) \in \mathcal{G}_z$; by hypothesis, this manifold is diffeomorphic to $\mathbb{R}^{nm}$ via the map $\varphi_z = \pi_m|_{\mathcal{G}_z}$, where $\pi_m : \mathbb{R}^{n(m+1)} \longrightarrow \mathbb{R}^{nm}$ is the projection onto the last nm factors. In other words, the last nm coordinates of a point in $\mathbb{R}^{n(m+1)}$ serve as global coordinates of $\mathcal{G}_z$. Introduce the projection

$$\begin{array}{rcl} \pi^0_{\mathbb{R}^n} : \mathbb{R}^{n(m+1)} & \longrightarrow & \mathbb{R}^n \\ (q_0, \ldots, q_m) & \longmapsto & q_0. \end{array} \tag{4.5}$$

We now define

$$(\Gamma_0)_t(\omega) := \pi^0_{\mathbb{R}^n} \circ \varphi_z^{-1}\left((\Gamma_1)_t(\omega), \ldots, (\Gamma_m)_t(\omega)\right). \tag{4.6}$$

It is immediate to see that $(\Gamma_0)_{t=0} = z$ and that Γ_0 is a semimartingale because, by construction, it is a composition of smooth functions with semimartingales. Let now Γ^z be the unique solution of (4.1) with a.s. initial condition $z \in \mathbb{R}^n$. We will proceed by proving that Γ_0 defined in (4.6) equals Γ^z and we will therefore have a superposition rule Φ given by the map $\Phi(z;\Gamma_1,\ldots,\Gamma_m) := \pi^0_{\mathbb{R}^n} \circ \varphi_z^{-1}(\Gamma_1,\ldots,\Gamma_m)$. Notice that unless additional hypotheses are assumed on the family $\mathcal{G}$, there is no guarantee on the smoothness of Φ on the z variable.

In order to prove that Γ_0 equals Γ^z, denote by $\left(q^k; k=1,\ldots,n\right)$ the coordinates on $\mathbb{R}^n$ and by $\left(q^k_a; k=1,\ldots,n; a=0,\ldots,m\right)$ the coordinates on $\mathbb{R}^{n(m+1)}$. Let $F^a_k : \mathbb{R}^{nm} \to \mathbb{R}^n$ and $X^a_k : \mathbb{R}^{nm} \to \mathbb{R}^{n(m+1)}$ be the maps defined as

$$F^a_k\left(q_1,\ldots,q_m\right) = T_{(q_1,\ldots,q_m)}(\pi^0_{\mathbb{R}^n} \circ \varphi_z^{-1} \circ \pi_m)\left(\frac{\partial}{\partial q^k_a}\right)$$

$$\begin{aligned} X^a_k\left(\varphi_z^{-1}(q_1,\ldots,q_m)\right) &= T_{(q_1,\ldots,q_m)}(\varphi_z^{-1} \circ \pi_m)\left(\frac{\partial}{\partial q^k_a}\right) \\ &= (F^a_k\left(q_1,\ldots,q_m\right), 0, \overset{a-1}{\ldots}, \overbrace{(0, \overset{k-1}{\ldots}, 1, \ldots, 0)}^{n \text{ entries}}, \overset{m-a}{\ldots}, 0), \end{aligned}$$

where $a=1,\ldots,m$, $k=1,\ldots,n$. Observe that, by construction, the nm vector fields X^a_k are linearly independent and span $T_q\mathcal{G}_z$ at any $q \in \mathcal{G}_z$, since φ_z^{-1} is a diffeomorphism form $\mathbb{R}^{nm}$ to $\mathcal{G}_z$.

Now, we notice that for any $j=1,\ldots,l$, the vectors

$$\widetilde{S}_j\left(X;\Gamma^z,\Gamma_1,\ldots,\Gamma_m\right) = \left(S_j\left(X,\Gamma^z\right), S_j\left(X,\Gamma_1\right),\ldots,S_j\left(X,\Gamma_m\right)\right) \tag{4.7}$$

are by hypothesis tangent to $\mathcal{G}_z$. Additionally, due to (4.6) and the Stratonovich differentiation rules we can write

$$\delta\Gamma_0 = \sum_{a=1}^m\sum_{k=1}^n F^a_k\left(\Gamma_1,\ldots,\Gamma_m\right)\delta\Gamma^k_a = \sum_{a=1}^m\sum_{k=1}^n\sum_{j=1}^l F^a_k\left(\Gamma_1,\ldots,\Gamma_m\right)S^k_j\left(X,\Gamma_a\right)\delta X^j. \tag{4.8}$$

Moreover,

$$\left(\sum_{a=1}^m\sum_{k=1}^n F^a_k\left(\Gamma_1,\ldots,\Gamma_m\right)S^k_j\left(X,\Gamma_a\right), S_j\left(X,\Gamma_1\right),\ldots,S_j\left(X,\Gamma_m\right)\right) \in \mathbb{R}^{n(m+1)} \tag{4.9}$$

belongs also to $T\mathcal{G}_z$ for any $j=1,\ldots,l$, since (4.9) can be written as a linear combination of the nm linearly independent vector fields X^a_k. Indeed,

$$\begin{aligned} &\left(\sum_{a=1}^m\sum_{k=1}^n F^a_k\left(\Gamma_1,\ldots,\Gamma_m\right)S^k_j\left(X,\Gamma_a\right), S_j\left(X,\Gamma_1\right),\ldots,S_j\left(X,\Gamma_m\right)\right) \\ &\qquad = \sum_{a=1}^m\sum_{k=1}^n S^k_j\left(X,\Gamma_a\right)X^a_k\left(\Gamma_1,\ldots,\Gamma_m\right). \end{aligned}$$

Subtracting (4.9) from (4.7), we see that for any $j = 1, \ldots, l$,

$$W_j := \left(S_j\left(X, \Gamma^z\right) - \sum_{a=1}^{m} \sum_{k=1}^{n} F_k^a\left(\Gamma_1, \ldots, \Gamma_m\right) S_j^k\left(X, \Gamma_a\right), 0, \ldots, 0 \right) \in T\mathcal{G}_z.$$

Any of these vectors fields, if different from zero, is obviously linearly independent from all the X_k^a, $a = 1, \ldots, m$, $k = 1, \ldots, n$. If that is the case we could therefore conclude that $\dim(\mathcal{G}_z)$ is strictly bigger than nm, which is obviously a contradiction. Therefore, $W_j = 0$ necessarily, and hence

$$S_j\left(X, \Gamma^z\right) = \sum_{a=1}^{m} \sum_{k=1}^{n} F_k^a\left(\Gamma_1, \ldots, \Gamma_m\right) S_j^k\left(X, \Gamma_a\right),$$

which guarantees that Γ_0 is a solution of (4.1) because by (4.8)

$$\delta\Gamma_0 = \sum_{j=1}^{l} S_j\left(X, \Gamma^z\right) \delta X^j = \delta\Gamma^z. \quad \blacksquare$$

Remark 4.5 In the previous proposition we saw how the tangency of the diagonal extensions of the vector fields that define the SDE to the submanifolds in $\mathcal{G}$ is a sufficient condition to ensure the existence of a superposition rule. Is it necessary? Suppose that we have a smooth superposition rule $(\Phi, \Gamma_1, \ldots, \Gamma_m)$ and let Ψ be the associated map introduced in (4.2). As we have that $\Psi_z\left(\Gamma^z, \Gamma_1, \ldots, \Gamma_m\right) = 0$, the Stratonovich differentiation rules yield

$$0 = \sum_{i=1}^{n} \sum_{a=0}^{m} \frac{\partial \Psi_z}{\partial q_a^i}\left(\Gamma^z, \Gamma_1, \ldots, \Gamma_m\right) \delta\Gamma_a^i = \sum_{j=1}^{l} \sum_{i=1}^{n} \sum_{a=0}^{m} \frac{\partial \Psi_z}{\partial q_a^i}\left(\Gamma^z, \Gamma_1, \ldots, \Gamma_m\right) S_j^i\left(X, \Gamma_a\right) \delta X^j. \tag{4.10}$$

A sufficient condition for this identity to hold is that, for any $j \in \{1, .., l\}$,

$$\sum_{i=1}^{n} \sum_{a=0}^{m} \frac{\partial \Psi_z}{\partial q_a^i}\left(\Gamma^z, \Gamma_1, \ldots, \Gamma_m\right) S_j^i\left(X, \Gamma_a\right) = 0 \tag{4.11}$$

or, equivalently, that the diagonal extensions $\widetilde{S}_j\left(X, \Gamma^z, \Gamma_1, \ldots, \Gamma_m\right)$ are tangent to the elements of the family of submanifolds $\mathcal{G}$ given by the zero fibers of the maps Ψ_z. Additionally, one can find situations in which (4.10) implies (4.11): for instance if $j = 1$ and (like in the case of the Brownian motion) the quadratic variation $[X, X]$ is a strictly increasing process, a straightforward application of the Doob-Meyer decomposition and the Itô isometry make in this case (4.10) and (4.11) equivalent.

Remark 4.6 If we add to the hypotheses of Proposition 4.4 that for any $z \in \mathbb{R}^n$ and for any $(p_1, \ldots, p_m) \in \mathbb{R}^{nm}$ there exist a submanifold $\mathcal{G}_z$ in $\mathcal{G}$ such that $(z, p_1, \ldots, p_m) \in \mathcal{G}_z$ (for instance when $\mathcal{G}$ is a foliation of $\mathbb{R}^{n(m+1)}$ whose leaves are diffeomorphic to $\mathbb{R}^{nm}$ via π_m) then the superposition function that we constructed in the proof of that result has the following extremely convenient property: the superposition function is the same for any fundamental sets of solutions $\{\Gamma_1, \ldots, \Gamma_m\}$ that we may want to choose. In other words, once Φ is know, we can take m arbitrary independent solutions of (4.1) to write down any solution. This situation frequently occurs in mechanics; see for instance, the study of the classical Riccati equation in ([CMN98]).

4.2 The stochastic Lie-Scheffers Theorem

The main goal of this section is proving a theorem that characterizes the existence of a superposition rule for a stochastic differential equation in terms of the integrability properties of the distribution spanned by the vector fields that define it. This can be translated into a Lie algebraic requirement, which allows us to recover the classical Lie-Scheffers Theorem in the stochastic context (Corollary 4.11).

In order to have at hand the necessary concepts to state the main theorem, we start by briefly recalling some standard results on generalized distributions due to Stefan [St74a, St74b] and Sussman [Su73]. Let M be a smooth manifold, $\mathcal{D} \subset \mathfrak{X}(M)$ be a family of smooth vector fields, and D the smooth generalized distribution spanned by $\mathcal{D}$. Let $G_{\mathcal{D}}$ be the pseudogroup of transformations generated by the flows of the vector fields in $\mathcal{D}$ and constructed as follows: let $k \in \mathbb{N}^*$ be a positive natural number, $\mathcal{X}$ an ordered family $\mathcal{X} = (X_1, \ldots, X_k)$ of k elements of $\mathcal{D}$, and T a k–tuple $T = (t_1, \ldots, t_k) \in \mathbb{R}^k$ such that F_t^i denotes the (locally defined) flow of X_i, $i \in \{1, \ldots, k\}$, t_i; the elements $\mathcal{F}_T$ of $G_{\mathcal{D}}$ are the locally defined diffeomorphisms of the form $\mathcal{F}_T = F_{t_1}^1 \circ F_{t_2}^2 \circ \cdots \circ F_{t_k}^k$. Two points x and y in M are said to be $G_{\mathcal{D}}$-equivalent, if there exists a diffeomorphism $\mathcal{F}_T \in G_{\mathcal{D}}$ such that $\mathcal{F}_T(x) = y$. The relation $G_{\mathcal{D}}$–equivalent is an equivalence relation whose equivalence classes are called the $G_{\mathcal{D}}$-orbits, that are sometimes referred to as the ***accessible sets*** associated to the family $\mathcal{D}$.

Given the family $\mathcal{D}$ and the associated pseudogroup $G_{\mathcal{D}}$ we can define another family $\mathcal{D}'$ of vector fields as

$$\mathcal{D}' := \{T\mathcal{F}_T \cdot X \mid X \in \mathcal{D}, \mathcal{F}_T \in G_{\mathcal{D}}\},$$

that clearly extends $\mathcal{D}$, that is, $\mathcal{D} \subset \mathcal{D}'$. The distribution D' spanned by the elements of $\mathcal{D}'$ is by construction $G_{\mathcal{D}}$-invariant. That is, for each $\mathcal{F}_T \in G_{\mathcal{D}}$ and for each $z \in M$ in the domain of $\mathcal{F}_T$,

$$T_z\mathcal{F}_T(D'(z)) = D'(\mathcal{F}_T(z)). \tag{4.12}$$

Moreover, since $(\mathcal{D}')' = \mathcal{D}'$ by construction, the Stefan-Sussmann Theorem guarantees that it is ***completely integrable*** in the sense that for every point $z \in M$, there exists an integral manifold of D' everywhere of maximal dimension which contains z. The maximal integral manifolds of a completely integrable generalized distribution on M form a ***generalized foliation*** of M (see for instance [D85]). A leaf of a generalized foliation is ***regular*** if it has a neighborhood where the singular foliation induces a regular foliation by restriction. A point is regular if it belongs to a regular leaf. Regular points are open and dense in M ([D85, Théorème 2.2]). We will refer to D' (respectively $\mathcal{D}'$) as the ***Stefan-Sussmann extension*** of D (respectively $\mathcal{D}$). The Stefan-Sussmann's Theorem also establishes an equivalence between the $G_{\mathcal{D}}$-invariance of D ($D' = D$) and its complete integrability; additionally, if D is a completely integrable distribution, then its integral manifolds are the G_D-orbits. When the distribution D has constant dimension, the Stefan-Sussmann Theorem reduces to the celebrated and especially convenient Frobenius Theorem which states the D is integrable if and only if D is involutive. Recall that D is ***involutive*** if $[X, Y]$ takes values in D whenever X and Y are vector fields with values in D.

In the sequel, we will use the following notation in order to be able to handle diagonal extensions of different dimensions. Given $l \in \mathbb{N}$ and $X \in \mathfrak{X}(\mathbb{R}^n)$, we will denote by $\widetilde{X}^l \in \mathfrak{X}(\mathbb{R}^{ln})$

the diagonal extension of X to $\mathbb{R}^{ln}$. For the sake of consistency with the previous section $\widetilde{X}$ means $\widetilde{X}^{m+1}$.

Theorem 4.7 (Lie-Scheffers' Theorem for SDE) *Let*

$$\delta\Gamma = S\left(X,\Gamma\right)\delta X \tag{4.13}$$

be a stochastic differential equation on $\mathbb{R}^n$, where $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^l$ is a given $\mathbb{R}^l$-valued semimartingale and $S\left(x,z\right) : T_x\mathbb{R}^l \longrightarrow T_z\mathbb{R}^n$ is a Stratonovich operator from $\mathbb{R}^l$ to $\mathbb{R}^n$. Let V be an arbitrary open neighborhood of $\mathbb{R}^n$.

(i) *Let V be an arbitrary open neighborhood of $\mathbb{R}^n$. If the X-dependent vector fields $\{S_1\left(X,\cdot\right), \ldots, S_l\left(X,\cdot\right)\}$ can be expressed on V as*

$$S_j\left(X,z\right) = \sum_{i=1}^{r} b_j^i\left(X\right)Y_i\left(z\right) \in T_z\mathbb{R}^n, \quad b_j^i \in C^\infty(\mathbb{R}^l), \quad z \in V, \tag{4.14}$$

and the distribution D spanned by the vector fields $\mathcal{D} = \{Y_1,\ldots,Y_r\} \subset \mathfrak{X}\left(V\right)$ is involutive, then (4.13) admits a local superposition rule.

(ii) *Conversely, suppose that (4.13) admits a superposition rule $(\Phi, \{\Gamma_1,\ldots,\Gamma_m\})$ and that the diagonal extensions $\{\widetilde{S}_1\left(X,\cdot\right),\ldots,\widetilde{S}_l\left(X,\cdot\right)\}$ to $\mathbb{R}^{n(m+1)}$ are tangent to the family $\mathcal{G}$ of nm-dimensional submanifolds of $\mathbb{R}^{n(m+1)}$ associated to this superposition rule (see Proposition 4.4). Let $\widetilde{D}(q) := \operatorname{span}\{\widetilde{S}_j(X_t,q) \mid j \in \{1,\ldots,l\}, t \in \mathbb{R}_+\}$, $q \in \mathbb{R}^{n(m+1)}$, $\widetilde{D}'$ the Stefan-Sussmann extension of $\widetilde{D}$, and $\mathcal{G}_0$ its associated generalized foliation. Let $z \in \mathbb{R}^n$, $p_i = (\Gamma_i)_{t=0}$, and suppose that $p = (z,p_1,\ldots,p_m) \in \mathbb{R}^{n(m+1)}$ belongs to a regular leaf $\left(\mathcal{G}_0\right)_z$ of $\mathcal{G}_0$. Then, there exists an open neighborhood V of z, a family of vector fields $\{Y_1,\ldots,Y_r\} \subset \mathfrak{X}\left(V\right)$, and a family of functions $\{b_j^i\}_{j=1,..,l}^{i=1,..,r} \subset C^\infty\left(\mathbb{R}^l\right)$ such that*

$$S_j\left(X,v\right) = \sum_{i=1}^{r} b_j^i\left(X\right)Y_i\left(v\right), \tag{4.15}$$

for any $v \in V$. Moreover, the vector fields $\{Y_1,\ldots,Y_r\}$ form a real Lie algebra.

Proof. **(i)** Given $l \in \mathbb{N}$, we define $V^l := V \times \overset{l)}{\ldots} \times V$ and $d_l := \max_{q\in V^l}\{\dim(\operatorname{span}\{\widetilde{Y}_1^l(q),\ldots, \widetilde{Y}_r^l(q)\})\}$. Notice that for any $l \in \mathbb{N}$ one has $d_l \leq d_{l+1}$ and $d_l \leq r$. Let $m \in \mathbb{N}$ be the smallest number for which $d_m = d_{m+1}$ and let $q_0 \in V^{m+1}$ be such that

$$\dim\left(\operatorname{span}\{\widetilde{Y}_1^{m+1}(q_0),\ldots,\widetilde{Y}_r^{m+1}(q_0)\}\right) = d_{m+1}. \tag{4.16}$$

The maximality of the dimension of $\operatorname{span}\{\widetilde{Y}_1^{m+1},\ldots,\widetilde{Y}_r^{m+1}\}$ at q_0 implies that there exists a neighborhood U of q_0 in V^{m+1} for which $\dim(\operatorname{span}\{\widetilde{Y}_1^{m+1}(q),\ldots,\widetilde{Y}_r^{m+1}(q)\}) = d_{m+1}$, for all $q \in U$. Indeed, the expression (4.16) is equivalent to saying that the $r \times n(m+1)$ matrix $M(q)$ with entries $M_{ij}(q) := (\widetilde{Y}_i^{m+1}(q))^j$ has rank d_m when evaluated at q_0 which, in turn, amounts to the existence of a non-vanishing minor $M_{d_{m+1}}(q_0)$ of $M(q_0)$ of order d_{m+1}. Since the minor $M_{d_{m+1}}(q)$ depends smoothly on q and $M_{d_{m+1}}(q_0) \neq 0$, there exists an open neighborhood U of q_0 in V^{m+1} for which $M_{d_{m+1}}(q) \neq 0$, for any $q \in U$. This implies that $\dim(\mathrm{span}\{\widetilde{Y}_1^{m+1}(q), \ldots, \widetilde{Y}_r^{m+1}(q)\}) \geq d_{m+1}$, for all $q \in U$. However, the maximality used in the definition of d_{l+1} implies that the previous inequality is necessarily an equality.

Consequently, we have found an open set $U \subset V^{m+1}$ in which the distribution D spanned by the family $\{\widetilde{Y}_1^{m+1}, \ldots, \widetilde{Y}_r^{m+1}\}$ has constant rank. Moreover, (4.4) and the hypothesis on $\{Y_1, \ldots, Y_r\}$ being in involution imply by the classical Frobenius Theorem that D is integrable. Let $\mathcal{G}_0$ be the family of maximal integrable leaves of D that form a foliation of U^{m+1}. Now, shrinking U if necessary and using foliation coordinates for $\mathcal{G}_0$, we extend the distribution D to another integrable distribution $\overline{D} \supset D$ of rank nm whose integrable leaves $\mathcal{G}$ contain those of $\mathcal{G}_0$, and for which the restrictions of $\pi_m : \mathbb{R}^{n(m+1)} \to \mathbb{R}^{mn}$ to the leaves in $\mathcal{G}$ are diffeomorphisms onto their images.

Let now $\{p_1, \ldots, p_m\}$ be a set of m distinct points in V such that $(p_1, \ldots, p_m) \in \pi_m(U)$ and $\{\Gamma_1, \ldots, \Gamma_m\}$ the solutions of of (4.13) such that $(\Gamma_1, \ldots, \Gamma_m)_{t=0} = (p_1, \ldots, p_m)$ a.s.. Let $\Gamma := (\Gamma_1, \ldots, \Gamma_m)$ and τ the stopping time defined as $\tau := \inf\{t > 0 \mid \Gamma_t \neq \pi_m(U)\}$. Since the vector fields

$$\widetilde{S}_j^{m+1}(X, \Gamma) = \sum_{i=1}^{r} b_j^i(X)\, \widetilde{Y}_i^{m+1}(\Gamma)$$

are tangent to the integral leaves of $\mathcal{G}_0$ and hence to those of $\mathcal{G}$, at least up to time τ, Proposition 4.4 guarantees the existence of a local superposition rule.

(ii) We start the proof by providing a lemma that will be needed in our argument.

Lemma 4.8 *Let $\{Y_1, \ldots, Y_r\} \subset \mathfrak{X}(\mathbb{R}^n)$ with $r \leq mn$ and let $\{\widetilde{Y}_1, \ldots, \widetilde{Y}_r\}$ be the corresponding diagonal extensions to $\mathbb{R}^{n(m+1)}$. Suppose that $\{T_q\pi_m(\widetilde{Y}_1(q)), \ldots, T_q\pi_m(\widetilde{Y}_r(q))\}$ are linearly independent for any q in a neighborhood $U \subseteq \mathbb{R}^{n(m+1)}$. If the sum $\sum_{i=1}^r b^i \widetilde{Y}_i$ with $b^i \in C^\infty(U)$, $i = 1, \ldots, r$, is again a diagonal extension then the functions b^i are necessarily the pull-back by π_m of a family functions in $C^\infty(\pi_m(U))$. More specifically, if $(q_a^j; j = 1, \ldots, n; a = 0, \ldots, m)$ are coordinates for $\mathbb{R}^{n(m+1)}$, then the functions $\{b^i\}_{i=1,..,r}$ do not depend on $(q_0^j; j = 1, \ldots, n)$.*

Proof. Using the coordinates $\left(q^j; j = 1, \ldots, n\right)$ for $\mathbb{R}^n$, there exists a family of functions $A_i^j \in C^\infty(\mathbb{R}^n)$, $i \in \{1, \ldots, r\}$, $j \in \{1, \ldots, n\}$, such that the vector fields $\{Y_1, \ldots, Y_r\} \subset \mathfrak{X}(\mathbb{R}^n)$ can be written as

$$Y_i(q) = \sum_{j=1}^{n} A_i^j(q) \frac{\partial}{\partial q^j}$$

which implies that the diagonal extensions have the expression

$$\widetilde{Y}_i(q_0, \ldots, q_m) = \sum_{a=0}^{m} \sum_{j=1}^{n} A_i^j(q_a) \frac{\partial}{\partial q_a^j}.$$

Then, if we assume that

$$\sum_{i=1}^{r} b^i(q_0, \ldots, q_m)\, \widetilde{Y}_i(q_0, \ldots, q_m) = \sum_{i=1}^{r} \sum_{a=0}^{m} \sum_{j=1}^{n} b^i(q_0, \ldots, q_m)\, A_i^j(q_a) \frac{\partial}{\partial q_a^j}$$

is a diagonal extension on U, then there exist some functions $\{B^i\}_{i=1,\ldots,r} \subset C^\infty(\mathbb{R}^n)$ such that

$$\sum_{i=1}^{r} b^i(q_0,\ldots,q_m)\, A_i^j(q_a)\Big|_U = B^j(q_a)\big|_U, \quad a=0,\ldots,m, \quad j=1,\ldots,n.$$

That is, the r functions $b^i(q_0,\ldots,q_m)$ solve the following subsystem of linear equations

$$\begin{pmatrix} \mathcal{A}(q_0) \\ \mathcal{A}(q_1) \\ \vdots \\ \mathcal{A}(q_m) \end{pmatrix} \begin{pmatrix} b^1(q_0,\ldots,q_m) \\ \vdots \\ b^r(q_0,\ldots,q_m) \end{pmatrix} = \begin{pmatrix} \mathcal{B}(q_0) \\ \mathcal{B}(q_1) \\ \vdots \\ \mathcal{B}(q_m) \end{pmatrix} \tag{4.17}$$

where $\mathcal{A}$ and $\mathcal{B}$ are the $n(m+1)\times r$ and $n(m+1)\times 1$ matrices, respectively, defined as $\mathcal{A}(q_a)_{ij} = A_j^i(q_a)$ and $\mathcal{B}(q_a)_i = B^i(q_a)$, $a=0,\ldots,m$. Now, the hypothesis on the linear independence of $\{T\pi_m(\widetilde{Y}_1),...,T\pi_m(\widetilde{Y}_r)\}$ implies that the rank of the matrix $(\mathcal{A}(q_1),\ldots,\mathcal{A}(q_m))$ is $r \le nm$ and hence (4.17) has a unique solution which coincides with the unique solution of the system

$$\begin{pmatrix} \mathcal{A}(q_1) \\ \vdots \\ \mathcal{A}(q_m) \end{pmatrix} \begin{pmatrix} b^1(q_0,\ldots,q_m) \\ \vdots \\ b^r(q_0,\ldots,q_m) \end{pmatrix} = \begin{pmatrix} \mathcal{B}(q_1) \\ \vdots \\ \mathcal{B}(q_m) \end{pmatrix}. \tag{4.18}$$

Since there is no dependence on the coordinates q_0 in the augmented matrix associated to the system (4.18), its solution $(b^1,\ldots,b^r)$ does not therefore depend on q_0, as required. ▼

Suppose now that the stochastic differential equation (4.13) admits a superposition rule and that we are in the hypotheses of the theorem. We start by emphasizing that since the vector fields $\{\widetilde{S}_1(X,\cdot),\ldots,\widetilde{S}_l(X,\cdot)\}$ are, by hypothesis, tangent to the elements of the family $\mathcal{G}$ then their flows leave invariant those submanifolds and hence, the Stefan-Sussmann extension $\widetilde{D}'$ of $\widetilde{D}$ is also tangent to the elements of $\mathcal{G}$. This argument guarantees that, given the regular leaf $(\mathcal{G}_0)_z$ of $\mathcal{G}_0$, then there exists an element $\mathcal{G}_z$ in $\mathcal{G}$ that contains it.

Now since $p=(z,p_1,\ldots,p_m)\in\mathbb{R}^{n(m+1)}$ belongs to a regular leaf $(\mathcal{G}_0)_z$ of $\mathcal{G}_0$, then there is an open neighborhood U of p where we can choose (taking regular foliation coordinates) a family of linearly independent vector fields $\{\widetilde{Y}_1,\ldots,\widetilde{Y}_r\}\subset\mathfrak{X}(\mathbb{R}^{n(m+1)})$ that span the tangent spaces to the leaves of $\mathcal{G}_0\cap U$. The vector fields $\{\widetilde{Y}_1,\ldots,\widetilde{Y}_r\}$ can be chosen as the diagonal extensions of r vector fields $\{Y_1,\ldots,Y_r\}\subset\mathfrak{X}(\mathbb{R}^n)$, since the Stefan-Sussmann extension $\widetilde{D}'=\operatorname{span}\{T\widetilde{\mathcal{F}}_T\cdot\widetilde{S}_i(X,\cdot)\mid i\in\{1,\ldots,l\},\widetilde{\mathcal{F}}_T\in G_{\mathcal{D}}\}$ of $\widetilde{D}$ is made of diagonal extensions. Indeed, in order to see that $\widetilde{D}'$ is spanned by diagonal extensions, it suffices to notice that the flow $\widetilde{F}_t$ of the diagonal extension $\widetilde{Y}\in\mathfrak{X}(\mathbb{R}^{n(m+1)})$ of a vector field $Y\in\mathfrak{X}(\mathbb{R}^n)$ is $\widetilde{F}_t(q_0,\ldots,q_m)=(F_t(q_0),\ldots,F_t(q_m))$, with F_t the flow of Y; hence

$$\begin{aligned} T_q\widetilde{F}_t(\widetilde{Y}(q)) &= (T_{q_0}F_t\times\ldots\times T_{q_m}F_t)\,(\widetilde{Y}(q)) \\ &= (T_{q_0}F_t(Y(q_0)),\ldots,T_{q_m}F_t(Y(q_m)) = (\widetilde{TF_t(Y)})(q) \end{aligned}$$

is again a diagonal extension. Given that by (4.4) diagonal extensions form an algebra, the statement follows.

Moreover, since the distribution $\widetilde{D}'|_U$ is regular and integrable then it is necessarily integrable in the sense of Frobenius, that is, there exist functions $\{c_{ij}^k\}_{i,j,k=1,..,r} \subset C^\infty(\mathbb{R}^{n(m+1)})$ such that

$$\left[\widetilde{Y}_j, \widetilde{Y}_i\right] = \sum_{k=1}^{r} c_{ji}^k \widetilde{Y}_k. \tag{4.19}$$

Now, as $[\widetilde{Y}_j, \widetilde{Y}_i] = \widetilde{[Y_j, Y_i]}$, we conclude that $\sum_{k=1}^r c_{ji}^k \widetilde{Y}_k$ is a diagonal extension. Also, as the projection π_m is a local diffeomorphism when restricted to $U \cap \mathcal{G}_z$, the family of vectors $\{T\pi_m(\widetilde{Y}_1), \ldots, T\pi_m(\widetilde{Y}_r)\}$ is necessarily linearly independent. In these circumstances Lemma 4.8 implies that the coefficients $\{c_{ij}^k\}_{i,j,k=1,..,r}$ do not depend on the first n coordinates q_0^j, $j = 1, \ldots, n$. We now apply $\pi_{\mathbb{R}^n}^0$ (see (4.5)) on both sides of (4.19) and we obtain

$$\left[Y_j, Y_i\right](v) = \sum_{k=1}^{r} c_{ji}^k(q_1, \ldots, q_m) Y_k(v) \tag{4.20}$$

where $v \in V := \pi_{\mathbb{R}^n}^0(U)$ and $(q_1, \ldots, q_m) \in \mathbb{R}^{nm}$ is any arbitrary point such that $(v, q_1, \ldots, q_m) \in U$. Since the left hand side of (4.20) does not depend on $(q_1, \ldots, q_m)$ then the dependence of the coefficients $c_{ji}^k(q_1, \ldots, q_m)$ on those coordinates is necessarily trivial which allows us to conclude that $\{Y_1, \ldots, Y_r\}$ close a Lie algebra.

Finally, since the vector fields $\widetilde{S}_j(X, \cdot)$ are tangent to $\mathcal{G}_0$, $j = 1, \ldots, l$, then there is a family of X-dependent functions $b_j^i(X, \cdot) \in C^\infty(U)$ such that

$$\widetilde{S}_j(X, q) = \sum_{i=1}^{r} b_j^i(X, q) \widetilde{Y}_i(q),$$

for any $q \in U$. As $\widetilde{S}_j(X, \cdot)$ is also a diagonal extension, we can use again Lemma 4.8 in order to prove that the functions $\{b_j^i\}_{j=1,..,l}^{i=1,..,r}$ do not depend on q_0. Consequently,

$$\widetilde{S}_j(X, q) = \sum_{i=1}^{r} b_j^i(X, (q_1, \ldots, q_m)) \widetilde{Y}_i(p). \tag{4.21}$$

As we did in the previous paragraph, we apply $\pi_{\mathbb{R}^n}^0$ on both sides of (4.21)

$$S_j(X, v) = \sum_{i=1}^{r} b_j^i(X, (q_1, \ldots, q_m)) Y_i(v),$$

for any $v \in V$. Again, we realize that since the left hand side of this equation is independent of $(q_1, \ldots, q_m)$, the dependence of the functions b_j^i on the coordinates $(q_1, \ldots, q_m)$ is necessarily trivial, which yields expression (4.15). ■

Remark 4.9 Theorem 4.7 is a generalization for stochastic differential equations of the classical Lie-Scheffers Theorem stated for time-dependent ordinary differential equations. That theorem claims that a differential equation $\dot{y} = Y(t, y)$ on $\mathbb{R}^n$ given by a time-dependent vector field $Y(t,\cdot) \in \mathfrak{X}(\mathbb{R}^n)$, $t \in \mathbb{R}$, admits a superposition rule if and only if Y can be locally written in the form $Y(t,y) = \sum_{i=1}^{r} f^i(t) Y_i(y)$, where $\{f^i\}_{i=1,\ldots,r} \subset C^\infty(\mathbb{R})$ and $\{Y_1,\ldots,Y_r\} \subset \mathfrak{X}(\mathbb{R}^n)$ form a (real) Lie subalgebra of $(\mathfrak{X}(M),[\cdot,\cdot])$ (see [CGM07] and [CGM00]). In relation to the traditional presentation of the Lie-Scheffers Theorem, our Theorem 4.7:

(i) weakens the hypotheses under which we can guarantee the existence of superposition rules. The involutivity of the vector fields $\{Y_1,\ldots,Y_r\}$ is, in general, less restrictive than requiring that they form a Lie algebra over the reals. We know *a posteriori* by the second part of Theorem 4.7 that, around regular points, if there exists a superpositon rule, the components $\{S_1,\ldots,S_l\}$ of the Stratonovich operator can also be expressed in terms of a family of vector fields that close a Lie algebra.

(ii) carefully spells out the regularity conditions under which we have a converse; those conditions are only vaguely evoked in the already cited deterministic papers.

(iii) It is worth noticing that, apart from the two points that we just explained, Theorem 4.7 cannot be seen as a mere transcription of the deterministic Lie-Scheffers Theorem into the context of Stratonovich stochastic integration by using the so called Malliavin's Transfer Principle [Ma78] due to the purely stochastic conditions that appear in the statement of the theorem. Those additional requirements have to do with the tangency of the diagonal extensions of the components of the Stratonovich operator to the family of submanifolds associated to the superposition rule (see also Remark 4.5).

Remark 4.10 An interesting research problem would be the formulation of a Lie-Schffers Theorem in the context of Rough Paths Theory [CLT04]. Such result seems to us plausible and would yield Theorem 4.7 as a particular case.

In the next corollary, we show for the sake of completeness how the classical statement of the Lie-Scheffers Theorem (generalized to SDEs) can be easily obtained out of Theorem 4.7.

Corollary 4.11 *Using the notation in Theorem 4.7, suppose that the X-dependent family of vector fields $\{S_1(X,\cdot),\ldots,S_l(X,\cdot)\}$ that define the stochastic differential equation (4.1) can be expressed as*

$$S_j(X,z) = \sum_{i=1}^{r} b_j^i(X)\, Y_i(z) \in T_z\mathbb{R}^n, \quad b_j^i \in C^\infty(\mathbb{R}^l), \quad z \in \mathbb{R}^n.$$

Let $\mathrm{Lie}\{Y_1,\ldots,Y_r\}$ *be the real Lie subalgebra of* $(\mathfrak{X}(\mathbb{R}^n),[\cdot,\cdot])$ *generated by the family* $\{Y_1,\ldots,Y_r\} \subset \mathfrak{X}(\mathbb{R}^n)$*. If* $\mathrm{Lie}\{Y_1,\ldots,Y_r\}$ *is finite dimensional then (4.1) has a superposition rule.*

Proof. Let D and D_2 be the generalized distributions associated to the families of vector fields $\mathcal{D} = \{Y_1,\ldots,Y_r\}$ and $\mathcal{D}_2 = \mathrm{Lie}\{Y_1,\ldots,Y_r\}$, respectively. Observe that if $D(z) \subsetneq D_2(z)$, $z \in \mathbb{R}^n$, then since $\mathrm{Lie}\{Y_1,\ldots,Y_r\}$ is finite dimensional, we can always complete the family $\{Y_1,\ldots,Y_r\}$ with a finite number of vectors $\{Z_1,\ldots,Z_s\} \subset \mathcal{D}$ such that $D(z) = D_2(z)$. We then write the X-dependent vector fields $\{S_1(X,\cdot),\ldots,S_l(X,\cdot)\}$ as

$$S_j(X,z) = \sum_{i=1}^{r} b_j^i(X)\, Y_i(z) + \sum_{k=1}^{s} a_j^k(X) Z_k(z), \quad z \in \mathbb{R}^n,$$

with $a_j^k = 0$ for any $j = 1, \ldots, l$ and any $k = 1, \ldots, s$. Therefore, we may simply suppose that $D(z) = \operatorname{span}\{\operatorname{Lie}\{Y_1, \ldots, Y_r\}(z)\}$, $z \in \mathbb{R}^n$ and since D_2 is trivially involutive, the corollary follows from Theorem 4.7 (i). ■

4.3 Lie-Scheffers systems and stochastic differential equations on Lie groups and homogeneous spaces

The Lie-Scheffers systems that are defined by a set of vector fields that generate a finite dimensional Lie algebra, that is, those that satisfy the hypothesis of Corollary 4.11 or of Theorem 4.16 can be reformulated in the language of group actions. More specifically, as we see in the next proposition, such systems come down locally to studying the solutions of an equivalent Lie-Scheffers system on a Lie group.

Proposition 4.12 *Consider a stochastic differential equation that satisfies the hypotheses of Corollary 4.11. Let $z \in M$ be a point such that there exists a neighborhood V of z in which the dimension of $\operatorname{Lie}\{Y_1, \ldots, Y_r\}$ is constant. Then, shrinking V if necessary, there exists a Lie group G such that $\dim(G) = \dim(\operatorname{Lie}\{Y_1, \ldots, Y_r\}|_V)$, a group action $\Xi : G \times V \to V$, and Lie algebra elements $\{\xi_1, \ldots, \xi_r\} \subset \mathfrak{g}$ such that*

$$Y_i(z) = \xi_i^M(z) := \left.\frac{d}{dt}\right|_{t=0} \Xi\left(\exp\left(t\xi_i\right), z\right), \quad z \in V. \tag{4.22}$$

Moreover, the solution starting at $z \in M$ of the restriction to V of the stochastic differential equation can be expressed as

$$\Gamma_t^z = \Xi\left(g_t, z\right), \tag{4.23}$$

where $g_t : \mathbb{R}_+ \times \Omega \to G$ is the semimartingale solution of the stochastic differential equation on G

$$\delta g_t = \sum_{i=1}^{r} \xi_i^G\left(g_t\right) \delta X_t^i \tag{4.24}$$

with initial condition $g_{t=0} = e$ a.s.

Proof. Since the statement of the proposition is local we can always assume that the vector fields $\{Y_1, \ldots, Y_r\}$ are complete by multiplying them by a compactly supported bump function and by restricting ourselves to an open neighborhood V consistent with that construction. In that situation and if $\dim(\operatorname{Lie}\{Y_1, \ldots, Y_r\}|_V) < \infty$, Palais showed in [P57] (see Corollary in page 97 and Theorem III in page 95) that there exists a unique connected Lie group G contained in the group of diffeomorphisms of M and a left action $\Xi : G \times M \to M$ such that (4.22) holds and $T_e\Xi_z : \mathfrak{g} \to \operatorname{Lie}\{Y_1, \ldots, Y_r\}(z)$ is an isomorphism, for any $z \in V$.

Let now $g_t : \mathbb{R}_+ \times \Omega \to G$ be the solution semimartingale of the stochastic differential equation on G

$$\delta g_t = \sum_{i=1}^{r} \xi_i^G\left(g_t\right) \delta X_t^i, \tag{4.25}$$

where $\xi_i^G \in \mathfrak{X}(G)$ denotes the right invariant infinitesimal generator associated to $\xi_i \in \mathfrak{g}$ via the left translations of G on G. Given that any two infinitesimal generators ξ^G and ξ^M, $\xi \in \mathfrak{g}$, are related by the formula $T_g \Xi_z(\xi^G) = \xi^M(\Xi(g,z))$, $g \in G$, $z \in V$, it is straightforward to verify that if g_t is a solution of (4.24) with initial condition $g_{t=0} = e$ a.s., then

$$\Gamma_t^z = \Xi(g_t, z),$$

is the solution of $\delta\Gamma_t = \sum_{i=1}^r Y_i(\Gamma_t)\,\delta X_t^i$ such that $\Gamma_0 = z$, a.s. ■

Remark 4.13 Observe that (4.23) may be understood as a general reformulation of (4.38) (see also [B89, Théorème 19]). Processes of the type $\Gamma_t^z = \Xi(g_t, z)$ defined using a group action are sometimes called ***one point motions*** ([L04]).

The proposition that we just proved shows that for Lie-Scheffers systems defined by vector fields that generate a finite dimensional Lie algebra $\mathfrak{g}$, it is the associated Lie-Scheffers system on the Lie group G (4.24) that really matters. This is the subject of the rest of this section.

Stochastic differential equations on Lie groups. Let now G be an arbitrary connected Lie group and $\mathfrak{g}$ its Lie algebra. Let $\{\xi_1, \ldots, \xi_l\}$ and $\{\epsilon^1, \ldots, \epsilon^l\}$ be dual bases of $\mathfrak{g}$ and $\mathfrak{g}^*$, respectively. Left (respectively, right) translations on G will be denoted by $L : G \times G \to G$ (respectively, $R : G \times G \to G$). With the same notation that we have used so far, let

$$\begin{array}{rcl} S(\mu, g) : T_\mu \mathfrak{g} \simeq \mathfrak{g} & \longrightarrow & T_g G \\ \eta & \longmapsto & \sum_{i=1}^l \xi_i^G(g) \left\langle \epsilon^i, \eta \right\rangle = \eta^G(g) \end{array} \tag{4.26}$$

be a Stratonovich operator from $\mathfrak{g}$ to G, where η^G denotes the infinitesimal generator associated to the G-action on itself by left translations. Consider the stochastic differential equation associated to (4.26),

$$\delta g_t = \sum_{i=1}^l \xi_i^G(g_t)\,\delta X_t^i, \tag{4.27}$$

for some driving noise (semimartingale) $X : \mathbb{R}_+ \times \Omega \to \mathfrak{g}$. Using the equivariance of the vector fields $\xi^G \in \mathfrak{X}(G)$ with respect to right translations, that is, $T_h R_g(\xi^G(h))) = \xi^G(R_g(h))$ for any g, $h \in G$, and $\xi \in \mathfrak{g}$, it is immediate to check that if Γ^e is the solution of (4.27) with initial condition $\Gamma_{t=0}^e = e$ a.s., then the solution Γ_t^g starting at $g \in G$ is given by

$$\Gamma_t^g = L_{\Gamma_t^e} g = R_g(\Gamma_t^e) \tag{4.28}$$

In other words, the stochastic differential equation (4.27) *has a superposition rule in the sense of Definition 4.1 and the superposition function* Φ *is given by*

$$\begin{array}{rcl} \Phi : G \times G & \longrightarrow & G \\ (h, g) & \longmapsto & L_h g = R_g h. \end{array}$$

It is also worth noticing that (4.27) is ***stochastically complete*** ([E82, Chapter VII §6]) since it is a left-invariant system. Therefore any solution of (4.27) is defined for all $(t, \omega) \in \mathbb{R}_+ \times \Omega$

and, consequently, so is any one point motion and, in particular, any solution of any Lie-Scheffers system on a manifold M which can be globally considered as induced by a group action $\Xi : G \times M \to M$.

Lévy processes and Lie-Scheffers systems. This is an important class of Lie group valued stochastic processes and, as we will now see, a class of examples of Lie-Scheffers systems. Recall that a continuous process $g : \mathbb{R}_+ \times \Omega \to G$ is called a right Lévy process if, for any $0 = t_0 < t_1 < t_2 < \ldots < t_n$, the increments

$$g_{t_0}, g_{t_0} g_{t_1}^{-1}, g_{t_1} g_{t_2}^{-1}, \ldots, g_{t_{n-1}} g_{t_n}^{-1} \tag{4.29}$$

are *independent* and *stationary*. This means that the random variables in (4.29) are mutually independent and that their distributions only depend on the differences $t_i - t_{i-1}$, $i \in \{1, \ldots, n\}$. If $g_{t_0} \neq e$ a.s., we define $g_t^e = g_t g_{t_0}^{-1}$, which is a right Lévy process starting at the identity.

We are now going to see that continuous Lévy processes and Lie-Scheffers systems are closely related. First of all, recall that any right Lévy process on a locally compact topological group with a countable basis of open sets is a Markov process with a right invariant Feller transition semigroup $\{P_t\}_{t \in \mathbb{R}_+}$ given by $P_t f(g) := E\left[f(g_t^e g)\right]$, $g \in G$, where $f : G \to \mathbb{R}$ is any measurable function. Conversely, any right invariant *continuous* Markov process is a right Lévy process ([L04, Proposition 1.2]). Moreover, if $g : \mathbb{R}_+ \times \Omega \to G$ is a right Lévy process, then there exists a l-dimensional Brownian motion $B : \mathbb{R}_+ \times \Omega \to \mathbb{R}^l$ with respect to the natural filtration $\{\mathcal{F}_t^e\}_{t \in \mathbb{R}_+}$ of the process g_t^e, $l = \dim(\mathfrak{g})$, with covariance matrix $(a_{ij})_{i,j=1,\ldots,l}$ and some constants $\{c_i\}_{i=1,\ldots,l}$ such that

$$f(g_t) = f(g_0) + \sum_{i=1}^{l} \int_0^t \xi_i^G[f](g_s)\, \delta B_s^i + \sum_{i=1}^{l} c_i \int_0^t \xi_i^G[f](g_s)\, ds,$$

for any $f \in C^2(G)$ and where, as before, $\{\xi_1, \ldots, \xi_l\}$ is a basis of $\mathfrak{g}$ ([L04, Theorem 1.2]). This expression amounts to saying that the Lévy process $g : \mathbb{R}_+ \times \Omega \to G$ satisfies the stochastic differential equation

$$\delta g_t = \sum_{i=1}^{l} c_i \xi_i^G(g_s)\, \delta s + \sum_{i=1}^{l} \xi_i^G(g_s)\, \delta B_s^i,$$

and hence by Corollary 4.11 we can conclude that *any continuous right Lévy process is a solution of a right invariant Lie-Scheffers system.* Additionally, it can be shown in this context (see [L04, Theorem 1.2]) that one point motions obtained out of a G-action $\Xi : G \times M \to M$ are Markov processes with Feller transition semigroup $\left\{P_t^M\right\}_{t \in \mathbb{R}_+}$

$$P_t^M f(z) = E\left[f(\Xi(g_t^e, z))\right], \ z \in M, \ f \in C(M).$$

Lie-Scheffers systems on homogeneous spaces. Let $H \subset G$ be a closed subgroup of G and consider the ***homogeneous space*** $G/H = \{gH \mid g \in G\}$ with the unique smooth structure that makes the projection $\pi_H : G \to G/H$ into a submersion. The group G acts on G/H via the map $\lambda : G \times G/H \to G/H$ on G/H defined by $(h, gH) \mapsto (hg)H$. It is immediate to check that the infinitesimal generators associated to the left G-actions on G and on G/H are π_H-related, that is,

$$T_g\pi_H\left(\xi^G(g)\right) = \xi^{G/H}(\pi_H(g))$$

for any $g \in G$, any $\xi \in \mathfrak{g}$, and where $\xi^{G/H}(gH) = \left.\frac{d}{dt}\right|_{t=0} \lambda_{\exp(t\xi)}(gH)$. This straightforward observation has as an immediate consequence the next proposition:

Proposition 4.14 *Let $X : R_+ \times \Omega \to \mathfrak{g}$ be a $\mathfrak{g}$-valued semimartingale, G a Lie group, and $H \subset G$ a closed subgroup. Let Γ be a solution of the Lie-Scheffers system defined by X and the Stratonovich operator (4.26) with initial condition $\Gamma_{t=0}$. Then, $\pi_H(\Gamma)$ is a solution of the Lie-Scheffers system on G/H*

$$\delta\overline{\Gamma} = \sum_{j=1}^{l} \xi_j^{G/H}\left(\overline{\Gamma}_t\right)\delta X_t^j \tag{4.30}$$

with initial condition $\pi_H(\Gamma_{t=0})$.

Observe that since the Stratonovich operator (4.26) is right invariant by the action of G, and therefore H-invariant, and that since this action is free and proper, the previous proposition can be seen as a particular case of the ***Reduction Theorem*** 3.9. The next theorem is a transcription of the ***Reconstruction Theorem*** 3.10 into the present context and describes how to construct solutions in the opposite direction, that is, it tells us *how to construct a solution Γ of the Lie-Scheffers system (4.27) out of the solutions of two other dimensionally smaller Lie-Scheffers systems*: first, a solution of the reduced system (4.30) and second, another solution of a new Lie-Scheffers system, now on H.

Theorem 4.15 *Let $X : \mathbb{R}_+ \times \Omega \to \mathfrak{g}$ be a $\mathfrak{g}$-valued semimartingale, G a Lie group, $H \subset G$ a closed subgroup, and S the Stratonovich operator defined in (4.26). Let $R : H \times G \to G$ be the (right) action of H on G by right translations and A an auxiliary principal connection on $\pi_H : G \to G/H$. Then, any solution Γ of the system (4.27) can be written in the form*

$$\Gamma_t = R_{h_t} g_t = g_t h_t.$$

In this statement, $g : \mathbb{R}_+ \times \Omega \to G$ is a G-valued semimartingale horizontal with respect to A, i.e. $\int \langle A, \delta g_t\rangle = 0 \in \mathfrak{g}$, $g_{t=0} = \Gamma_{t=0}$, and such that $\pi_H(g_t)$ is a solution of the reduced system (4.30). On the other hand, $h : \mathbb{R}_+ \times \Omega \to H$ is a H-valued semimartingale that satisfies the stochastic differential equation

$$\delta h_t = \widetilde{R}(Y_t, h_t)\,\delta Y_t \tag{4.31}$$

with initial condition $h_{t=0} = e$, and associated to the Stratonovich operator

$$\begin{array}{rcl} \widetilde{R}(\xi, h) : T_\xi\mathfrak{h} & \longrightarrow & T_hH \\ \eta & \longmapsto & T_eR_h(\eta) = \eta^H(h), \end{array} \tag{4.32}$$

and the stochastic component $Y : \mathbb{R}_+ \times \Omega \to \mathfrak{h}$ given by

$$Y = \sum_{i=1}^{l} \int A_{g_t}\left(\xi_i^G(g_t)\right)\delta X^i.$$

Proof. See Chapter 3 Theorem 3.10 and Proposition 3.12. ■

4.3.1 The Wei-Norman method for solving stochastic Lie-Scheffers systems

The method that we are going to develop in this subsection is a generalization to stochastic systems of the one proposed by Wei and Norman in [WN63, WN64] in order to solve by quadratures time evolution equations of the form $\frac{dU_t}{dt} = H_t U_t$ that appear in quantum mechanics, where both U_t and H_t are bounded linear operators on a suitable Hilbert space. This method has already been adapted by Cariñena and Ramos [CR01] to the study of deterministic Lie-Scheffers systems on Lie groups and it is their approach that we will follow. As we will see later on, the power of this method and the ease of its implementation depends strongly on the algebraic structure of the Lie algebra $\mathfrak{g}$ of the group G where the solutions of the stochastic differential equation take values.

Let $\Gamma : \mathbb{R} \times \Omega \to G$ be the solution of (4.27) such that $\Gamma_{t=0} = e \in G$ a.s.; we write it down in terms of second kind canonical coordinates with respect to a basis $\{\xi_1, \ldots, \xi_l\}$ of the Lie algebra $\mathfrak{g}$. That is,

$$\Gamma_t = \exp(d_t^1 \xi_1) \cdots \exp(d_t^l \xi_l), \tag{4.33}$$

where $\{d_t^1, \ldots, d_t^l\}$ is a family of real-valued semimartingales, $d^i : \mathbb{R}_+ \times \Omega \to \mathbb{R}$, such that $d_{t=0}^i = 0$ a.s. for any $i = 1, \ldots, l$. Notice that the expression (4.33) is only valid up to the exit time of Γ from the neighborhood U_e of $e \in G$ where the second kind canonical coordinates for G around the origin are valid. The key idea in this method is that if the functions d^i were differentiable then

$$\frac{d\Gamma_t}{dt} = T_e R_{\Gamma_t} \left(\sum_{i=1}^l \dot{d}_t^i \left(\prod_{j<i} \mathrm{Ad}_{\exp\left(d_t^j \xi_j\right)} \right) \xi_i \right)$$

(see [CR01, Eq. (33) and (34)]), where $\mathrm{Ad}_g(\eta) \in \mathfrak{g}$ is the adjoint representation of G on $\mathfrak{g}$, $g \in G$, $\eta \in \mathfrak{g}$. In our setup we obviously cannot invoke the differentiability of the functions d^i, however applying the Stratonovich differentiation rules to (4.33) with d^i our real-valued semimartingales, $i = 1, \ldots, l$, we have

$$\delta\Gamma_t = T_e R_{\Gamma_t} \left(\sum_{i=1}^l \delta d_t^i \left(\prod_{j<i} \mathrm{Ad}_{\exp\left(d_t^j \xi_j\right)} \right) \xi_i \right).$$

This expression implies that for any right invariant one-form $\mu^G \in \Omega(G)$, that is, $\mu^G(g) = T_g^* R_{g^{-1}}(\mu)$ for any $g \in G$ and a fixed $\mu \in \mathfrak{g}^*$,

$$\int \left\langle \mu^G, \delta\Gamma \right\rangle = \langle \mu, \sum_{i=1}^r \int \left(\prod_{j<i} \mathrm{Ad}_{\exp\left(\sum_{j=1}^l d_t^j \nu_j\right)} \right) \xi_i \delta d_t^i \rangle. \tag{4.34}$$

At the same time, it is clear that $\int \left\langle \mu^G, \delta\Gamma \right\rangle = \langle \mu, X \rangle$ and hence (4.34) implies that

$$X = \sum_{i=1}^l \int \left(\prod_{j<i} \mathrm{Ad}_{\exp\left(d_t^j \xi_j\right)} \right) \xi_i \delta d_t^i.$$

Using the identity $\mathrm{Ad}_{\exp(\eta)} = \mathrm{e}^{\mathrm{ad}(\eta)} = \sum_{n \geq 0} \frac{1}{n!} \mathrm{ad}(\eta) \circ \overset{n}{\cdots} \circ \mathrm{ad}(\eta)$, for any $\eta \in \mathfrak{g}$, and writing $X = \sum_{i=1}^l X^i \xi_i$, we get the relation

$$\sum_{i=1}^l X^i \xi_i = \sum_{i=1}^l \int \left(\prod_{j<i} \mathrm{e}^{\,\mathrm{ad}\left(d_t^j \xi_j\right)} \right) \xi_i \delta d_t^i. \tag{4.35}$$

The system of stochastic differential equations (4.35) can be solved for the semimartingales d_t^i, $i = 1, \ldots, l$ by quadratures if the Lie algebra $\mathfrak{g}$ is solvable (see [WN63, WN64]) and, in particular, for nilpotent Lie algebras. The solvable case was extensively studied in [K80] where similar conclusions were presented using a different approach.

As a simple example consider the affine group in one dimension $\mathcal{A}_1$,that is, the group of affine transformations of the real line. Any element of $\mathcal{A}_1$ can be expressed as a pair of real numbers (a_0, a_1) with $a_1 \neq 0$ defining the affine transformation $x \mapsto a_1 x + a_0$. The product $* : \mathcal{A}_1 \times \mathcal{A}_1 \to \mathcal{A}_1$ in $\mathcal{A}_1$ is

$$(a_0, a_1) * (b_0, b_1) = (a_0 + a_1 b_0, a_1 b_1) .$$

If $\{\xi_0 = (1,0), \xi_1 = (0,1)\}$ is a basis of the Lie algebra $\mathfrak{a}_1$ of $\mathcal{A}_1$, it is immediate to check that

$$[\xi_0, \xi_1] = \mathrm{ad}_{\xi_0}(\xi_1) = -\xi_0. \tag{4.36}$$

Furthermore, the infinitesimal generators associated to the left action of $\mathcal{A}_1$ on itself are

$$\xi_0^{\mathcal{A}_1}(x,y) = \frac{\partial}{\partial x} \quad \text{and} \quad \xi_1^{\mathcal{A}_1}(x,y) = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.$$

A typical Lie system on $\mathcal{A}_1$ would be, for instance, the following Stratonovich differential equation on the upper half-plane $H_+ = \left\{(x,y) \in \mathbb{R}^2 \mid y > 0\right\}$,

$$\delta\Gamma_x = dt + \Gamma_x \delta B_t, \quad \delta\Gamma_y = \Gamma_y \delta B_t$$

obtained as a particular case of (4.26) when $G = \mathcal{A}_1$,

$$\delta\Gamma_t = \xi_0^{\mathcal{A}_1} dt + \ \xi_1^{\mathcal{A}_1} \delta B_t$$

where $X = (t, B)$ and $B : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is a Brownian motion. More generally, let $X : \mathbb{R}_+ \times \Omega \to \mathfrak{a}_1$ be an $\mathfrak{a}_1$-valued semimartingale and write $X = X^0\xi_0 + X^1\xi_1$, with X^0 and X^1 real semimartingales. Then, using (4.36), (4.35) reads in this particular case

$$X^0\xi_0 + X^1\xi_1 = \int \xi_0 \delta d_t^0 + \int \left(\xi_1 - d_t^0\xi_0\right)\delta d_t^1 = \left(\int \delta d_t^0 - \int d_t^0 \delta d_t^1\right)\xi_0 + \left(\int \delta d_t^1\right)\xi_1.$$

Putting together the terms that go both with ξ_1 and ξ_0 respectively, we obtain

$$d_t^1 = X_t^1, \quad d_t^0 = X_t^0 + \int_0^t d_s^0 \delta X_s^1,$$

and hence $\delta d_t^0 = \delta X_t^0 + d_s^0 \delta X_t^1$, whose solution is

$$d_t^0 = \mathrm{e}^{X_t^1}\left(\int_0^t \delta X_s^0\, \mathrm{e}^{-X_s^1}\right).$$

4.4 The flow of a stochastic Lie-Scheffers system

Theorem 4.7 claims, roughly speaking, that the stochastic system (4.1) admits a superposition rule $(\Phi, \{\Gamma_1, \ldots, \Gamma_m\})$ if the components of the Stratonovich operator $S(x,z) : T_x\mathbb{R}^l \longrightarrow T_z\mathbb{R}^n$, $x \in \mathbb{R}^l$, $p \in \mathbb{R}^n$, that define it may be written as $S_j(X,z) = \sum_{i=1}^{r} b_j^i(X) Y_i(z)$, where $b_j^i \in C^\infty(\mathbb{R}^l)$ and $\{Y_1, \ldots, Y_r\} \subseteq \mathfrak{X}(\mathbb{R}^n)$ span an involutive distribution. The converse of this statement is also true provided that, for a given initial condition $z \in \mathbb{R}^n$, the point $(z, (\Gamma_1, \ldots, \Gamma_m)_{t=0})$ is a regular point of the foliation $\mathcal{G}_0$ generated by the diagonal extensions of $\{S_1(X, \cdot), \ldots, S_m(X, \cdot)\}$. Notice that this is a reasonable condition since the set of regular points of a generalized foliation is open and dense ([D85, Théorème 2.2]). Moreover, when this happens, the vector fields $\{Y_1, \ldots, Y_r\}$ form a real Lie algebra.

The condition on the vector fields $\{Y_1, \ldots, Y_r\}$ forming a real finite dimensional Lie algebra or, more generally, $\dim(\mathrm{Lie}\{Y_1, \ldots, Y_r\}) < \infty$, are particularly appealing since these are algebraic requirements that we may expect to be easily verified for stochastic differential equations of a certain type. Moreover, these conditions have consequences that go beyond Corollary 4.11. More specifically, we will show that if $\dim(\mathrm{Lie}\{Y_1, \ldots, Y_r\}) < \infty$, then the general solution of a stochastic differential equation *can be written by composing a deterministic function with a suitable noise*. In the following paragraphs we are going to give a precise meaning to this statement and to put it in the context of well known results available in the literature.

Traditionally, stochastic differential equations on a manifold M have been presented as

$$\delta\Gamma_t = Y_0(\Gamma_t)dt + \sum_{i=1}^{r} Y_i(\Gamma_t)\, \delta B_t^i, \tag{4.37}$$

where $\{Y_0, \ldots, Y_r\} \subseteq \mathfrak{X}(M)$ and $B : \mathbb{R}_+ \times \Omega \to \mathbb{R}^r$ is a r-dimensional Brownian motion defined on a standard filtered probability space $(\Omega, \mathcal{F}_t, P)$. For the sake of having a more compact notation, we write $B_t^0 := t$. The flow of such a stochastic differential equation may be locally written, that is, up to a given stopping time τ, by means of a Taylor series expansion that comes out of Picard's iterative method for solving stochastic differential equations. In order to be more explicit we introduce some notation. Let $J = \{j_1, \ldots, j_n\}$, $j_i \in \{0, \ldots, r\}$, $1 \leq i \leq n$, be a multi-index of ***size*** n. $\|J\|$ will denote the ***degree*** of J that, by definition, is the size of J plus the number of zeros in the n-tuple $(j_1, \ldots, j_n)$. For any $J = \{j_1, \ldots, j_n\}$, we consider the iterated Stratonovich multiple integral

$$B_t^J = \underset{0<t_1<\ldots<t_n<t}{\int_0^t \cdots \int_0^{t_3} \int_0^{t_2}} \delta B_{t_1}^{j_1} \cdots \delta B_{t_n}^{j_n}.$$

In addition, Y_J will denote

$$Y_J := [Y_{j_1}, [Y_{j_2}, \ldots, [Y_{j_{n-1}}, Y_{j_n}]].$$

If $Y \in \mathfrak{X}(M)$ is a vector field on the manifold M, we will use the following notation for its flow: $\exp(sY)(z)$ denotes the solution at time s of the ordinary differential equation $\dot{\gamma} = Y(\gamma)$ with initial condition $\gamma(0) = z$. Then,

Theorem 4.16 ([B89, Théorème 20]) *With the notation introduced so far, if* $\dim(\mathrm{Lie}\{Y_0, \ldots, Y_r\}) < \infty$ *and* $\mathrm{span}\{\mathrm{Lie}\{Y_0, \ldots, Y_r\}\}$ *has constant dimension on a neighborhood* V *of the point* $z \in M$*, then there exists a stopping time* τ *such that the solution of (4.37) with initial condition* z *can be expressed as*

$$\Gamma_t^z = \exp\left(\sum_{n=1}^{\infty} \sum_{\|J\|=n} \beta_J B_t^J\right)(z) \tag{4.38}$$

up to time τ*. In this expression,*

$$\beta_J := \sum_{\sigma \in S_n} \frac{(-1)^{e(\sigma)}}{n^2 \binom{n-1}{e(\sigma)}} Y_{\sigma(J)},$$

S_n *denotes the permutation group of* n *elements, and* $e(\sigma)$ *is the cardinality of the set* $\{j \in \{1, \ldots, n-1\} \mid \sigma(j) > \sigma(j+1)\}$*.*

If the finiteness condition on the dimensionality of the Lie algebra generated by the vector fields is not available but, nevertheless, $\{Y_0, \ldots, Y_r\}$ are Lipschitz vector fields, then the solution of (4.37) starting at $z \in M$ can always be approximated by a process like (4.38): if ζ_t^N denotes the *finite* sum $\sum_{n=1}^{N} \sum_{\|J\|=n} \beta_J B_t^J$, then

$$\Gamma_t^z = \exp\left(\zeta_t^N\right)(z) + t^{N/2} R_N(t)$$

where the error term $R_N(t)$ is bounded in probability when t tends to 0 ([C93, Theorem 2.1]). The expression (4.38) also holds if instead of the hypotheses of Theorem 4.16 we require M to be an analytic manifold and $\{Y_0, \ldots, Y_r\}$ a family of real analytic vector fields ([B89, Théorème 10]). An important consequence of Theorem 4.16 lies in the fact that the general solution of the stochastic differential equation (4.37) may be written, at least locally and up to a suitable stopping time τ, as the composition of a deterministic and smooth function, namely, the flow exponential, with the diffusion that defines the stochastic differential equation (see [H92] for a complementary reading). From this point of view, there is a strong resemblance between Theorem 4.16 and Theorem 4.7:

- First, by Corollary 4.11, all the systems that satisfy the hypotheses of Theorem 4.16 admit a superposition rule.
- Second, the superposition rule allows us to write any solution as the composition of the deterministic function Φ and the set of solutions $\{\Gamma_1, \ldots, \Gamma_m\}$ that are responsible for the stochastic behavior of the resulting flow.

We conclude by quoting two references that study the nilpotent case (that is, the Lie algebra $\mathrm{Lie}\{Y_0, ..., Y_r\}$ is nilpotent); this case has deserved special attention in the literature (see, for example, [K80]) because in that situation the Taylor series expansion of the flow in terms of iterated integrals in (4.38) becomes finite. We also recommend the excellent exposition in [B04] for a complementary approach to the subject of Taylor series approximation of the general solution of (4.37); in this book it is shown that, for instance, the Carnot group of depth

$N = \dim(\mathrm{Lie}\{Y_0, \ldots, Y_r\})$ can be used in the nilpotent case to integrate the Lie algebra action of $\mathrm{Lie}\{Y_0, \ldots, Y_r\}$ when one writes, as we did in the previous section, a Lie-Scheffers system as a stochastic differential equation on a Lie group that acts on the manifold in question.

4.5 Examples.

4.5.1 Inhomogeneous linear systems.

Let $A_k : \mathbb{R} \to M_n(\mathbb{R})$ a $n \times n$ time-dependent real matrix and $B_k : \mathbb{R} \to \mathbb{R}^n$ a time-dependent vector for any $k = 1, \ldots, l$. Let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}^l$ be a semimartingale. An inhomogeneuous linear system is a system of stochastic differential equations on $\mathbb{R}^n$ that may be written as

$$\delta\Gamma_t = \sum_{k=1}^{l} (A_k(t)(\Gamma_t) - B_k(t))\delta X_t^k \tag{4.39}$$

Let $\left(q^1, \ldots, q^n\right)$ be coordinates for $\mathbb{R}^n$. It is an exercise to check that (4.39) can be equivalently written as

$$\delta\Gamma_t = \sum_{k=1}^{l} \sum_{i,j=1}^{n} (A_k)_i^j (t) Y_j^i(\Gamma_t)\delta X_t^k + \sum_{k=1}^{l} \sum_{i,j=1}^{n} (B_k)^j(t) Z_j(\Gamma_t)\delta X_t^k$$

where the vector fields Y_j^i, $Z_j \in \mathfrak{X}(\mathbb{R}^n)$, $i, j, k = 1, \ldots, n$, are given by

$$Y_j^i = q^i \frac{\partial}{\partial q^j}, \qquad Z_j = \frac{\partial}{\partial q^j}.$$

Given that

$$[Y_j^i, Y_l^k] = \delta_j^k Y_l^i - \delta_l^i Y_j^k, \quad [Y_j^i, Z_k] = -\delta_k^i Z_j, \text{ and } [Z_i, Z_j] = 0$$

we see that the vectors $\{Y_j^i, Z_k \mid i, j, k = 1, \ldots, n\} \subset \mathfrak{X}(\mathbb{R}^n)$ span a Lie algebra isomorphic to the $(n^2 + n)$-dimensional Lie algebra of the group of affine transformations of $\mathbb{R}^n$. Therefore, the system (4.39) satisfies the hypotheses of Theorem 4.7 and hence it admits a superposition rule. In order to explicitly construct the superposition rule, let Γ^{e_j} be the solution of the homogeneous part of (4.39),

$$\delta\Gamma_t = \sum_{k=1}^{l} A_k(t)(\Gamma_t)\delta X_t^k$$

with initial solution $\Gamma_{t=0}^{e_j} = e_j \in \mathbb{R}^n$ a.s., where $e_j = (0, \overset{j-1}{\ldots}, 0, 1, 0, \ldots, 0)$ for any $j = 1, \ldots, n$. Let $\overline{\Gamma}$ be a particular solution of (4.39) with initial condition $\overline{\Gamma}_{t=0} = 0 \in \mathbb{R}^n$ a.s.. Then,

$$\Gamma_t = \sum_{j=1}^{n} z^j \Gamma_t^{e_j} + \overline{\Gamma}_t$$

is the general semimartingale solution of (4.39) starting at $z = (z^1, \ldots, z^n) \in \mathbb{R}^n$.

4.5.2 The stochastic exponential of a Lie group.

Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. Let $\{\xi_1, \ldots, \xi_l\}$ a basis of $\mathfrak{g}$ and $X : \mathbb{R}_+ \times \Omega \to \mathfrak{g}$ be a $\mathfrak{g}$-valued semimartingale. Observe that X can be written as $X = \sum_{i=1}^{r} a_t^i \xi_i$ for a family of real semimartingales $a^i : \mathbb{R}_+ \times \Omega \to \mathbb{R}$, $i = 1, \ldots, l$. Following [HL86] and [EP01], we define the (left) ***stochastic exponential*** $\mathcal{E}(X) : \mathbb{R}_+ \times \Omega \to G$ of X as the unique solution of the Lie-Scheffers system on G given by

$$\delta\Gamma_t = \sum_{i=1}^{l} (\xi_i)^G(\Gamma_t)\delta a_t^i$$

with initial condition $\Gamma_{t=0} = e \in G$ a.s.. Unlike the conventions used in Section 4.3, the vector fields $(\xi_i)^G \in \mathfrak{X}(G)$ here are not the right-invariant vector fields built from ξ_i, $i = 1, \ldots, l$, but the left-invariant ones. That is,

$$(\xi_i)^G(g) = T_e L_g(\xi_i), \quad g \in G.$$

Except for the fact that $(\xi_i)^G \in \mathfrak{X}(G)$, $i = 1, \ldots, l$, are now left-invariant, solving a Lie-Scheffers system on a Lie group such as those presented in Section 4.3 amounts to computing the stochastic exponential of a given $\mathfrak{g}$-valued semimartingale X.

The stochastic exponential establishes a bijection between $\mathfrak{g}$-valued local martingales and martingales on G with respect to certain connections. Recall that, given an affine connection $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ on a manifold M, a M-valued semimartingale $\Gamma : \mathbb{R}_+ \times \Omega \to M$ is said to be a ∇-martingale (or a martingale with respect to ∇) provided that

$$f(\Gamma) - f(\Gamma_{t=0}) - \frac{1}{2} \int \operatorname{Hess} f \,(d\Gamma, d\Gamma)$$

is a real local martingale for any $f \in C^\infty(M)$, where $\operatorname{Hess} f : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ is the bilinear form defined as

$$\operatorname{Hess} f \,(Y, Z) = Y\,[Z\,[f]] - \nabla_Z Y\,[f]$$

for any Y, $Z \in \mathfrak{X}(M)$ (see [E89, Chapter IV]). When $M = G$ is a Lie group, one can construct left invariant connections ∇ by using bilinear skew-symmetric forms $\alpha : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ on the Lie algebra $\mathfrak{g}$ via the definition

$$\nabla_{\xi^G} \eta^G := \alpha(\xi, \eta), \quad \xi, \eta \in \mathfrak{g}.$$

The curves $\exp(t\xi) \in G$, where $\xi \in \mathfrak{g}$ and $\exp : \mathfrak{g} \to G$ is the Lie algebraic exponential, coincide with the geodesics $c(t)$ with respect to these connections that start at $e \in G$ and that satisfy $\dot{c}(0) = \xi$. It can be shown ([EP01, Lemma 1.4]) that the connections built from $\alpha = 0$ and $\alpha(\xi, \eta) = \frac{1}{2}[\xi, \eta]$ induce the same ∇-martingales on G. Moreover, with respect to these two connections, the set of ∇-martingales consists precisely of the processes of the form $\Gamma_0 \mathcal{E}(X)$ where X is a $\mathfrak{g}$-valued local martingale and Γ_0 a G-valued $\mathcal{F}_0$-measurable random variable ([EP01, Proposition 1.9]). This expression provides the bijection between $\mathfrak{g}$-valued local martingales and ∇-martingales on G that we announced above.

4.5.3 Geometric Brownian motion.

Let $(\mathbb{R}_+, \cdot)$ be the Abelian Lie group of strictly positive real numbers endowed with the standard product. Its Lie algebra is simply $\mathbb{R}$ and, for any $\xi \in \mathbb{R}$, the Lie algebra exponential coincides with the standard exponential, that is $\exp \xi = e^{\xi}$; consequently, the infinitesimal generator (right or left-invariant) is

$$\xi^{\mathbb{R}_+}(q) = \xi q, \text{ for any } q \in \mathbb{R}.$$

Let $G = \mathbb{R}_+ \times \overset{n}{\ldots} \times \mathbb{R}_+$ be the Lie group constructed as the direct product of n copies of $(\mathbb{R}_+, \cdot)$. Its product map $\cdot : G \times G \to G$ is obviously $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_m) = (a_1 b_1, \ldots, a_n b_n)$, a_i, $b_i \in \mathbb{R}_+$ for any $i = 1, \ldots, n$, and its Lie algebra is $\mathfrak{g} = T_1\mathbb{R}_+ \times \overset{n}{\ldots} \times T_1\mathbb{R}_+ \simeq \mathbb{R} \times \overset{n}{\ldots} \times \mathbb{R} = \mathbb{R}^n$. Let $\{\xi_i = (0, \overset{i-1}{\ldots}, 0, 1, 0, \ldots, 0) \mid i = 1, \ldots, n\}$ be the canonical basis of $\mathfrak{g} = \mathbb{R}^n$, $\mu = (\mu^1, \ldots, \mu^n)$, $\sigma = (\sigma^1, \ldots, \sigma^n) \in \mathfrak{g}$ a couple of elements of $\mathfrak{g}$, $B : \mathbb{R}_+ \times \Omega \to \mathfrak{g}$ a n-dimensional Brownian motion on some filtered probability space $\left(\Omega, P, \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}\right)$, and consider the following Lie-Scheffers system on G

$$\delta\Gamma_t = \left(\mu - \frac{1}{2}\sigma^2\right)^G (\Gamma_t)dt + \sum_{i=1}^n \sigma^i \xi_i^G(\Gamma_t)\delta B_t^i, \tag{4.40}$$

where $\sigma^2 = ((\sigma^1)^2, \ldots, (\sigma^n)^2)$. Using coordinates $\left(q^1, \ldots, q^n\right)$ in G we can rewrite (4.40) as

$$\delta q_t^i = \left(\mu^i - \frac{1}{2}(\sigma^i)^2\right) q_t^i dt + \sigma^i q_t^i \delta B_t^i, \quad i = 1, \ldots, n,$$

which may be rewritten in terms of Itô integrals as

$$dq_t^i = \mu^i q_t^i dt + \sigma^i q_t^i dB_t^i, \quad i = 1, \ldots, n. \tag{4.41}$$

The solutions of the n-dimensional system of stochastic differential equations (4.41) are usually referred to as the geometric Brownian motion which is well-known for its use in the Black-Scholes theory of derivatives pricing as a model for the time evolution of the prices of n assets in a complete and arbitrage-free financial market.

The well-known solution of the differential equation (4.41) can be easily obtained by using the stochastic version of the Wei-Norman method that we introduced in Section 4.3.1. Indeed, let $q_t = \exp(a_t^1 \xi_1) \cdots \exp(a_t^n \xi_n)$ be the solution of (4.41) starting at $e = (1, \ldots, 1) \in G$ as in the, where $a^i : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ are real semimartingales such that $a_{t=0}^i = 0$ a.s. for any $i = 1, \ldots, n$. Since the Lie algebra $\mathfrak{g}$ of G is Abelian, and (4.40) is written in Lie-Scheffers form

$$\delta\Gamma_t = \sum_{i=1}^l \xi_i^G (\Gamma_t)\, \delta X_t^i$$

by taking the noise semimartingale $X := \left(\left(\mu^1 - \frac{(\sigma^1)^2}{2}\right) t + \sigma^1 B_t^1, \ldots, \left(\mu^n - \frac{(\sigma^n)^2}{2}\right) t + \sigma^n B_t^n\right)$, the equation (4.35) in the Wei-Norman method reduces to

$$\left(\mu^1 - (\sigma^1)^2/2, \ldots, \mu^n - (\sigma^n)^2/2\right) t + \left(\sigma^1 B_t^1, \ldots, \sigma^n B_t^n\right) = \sum_{i=1}^n \xi_i a_t^i,$$

which implies that $a_t^i = (\mu^i - (\sigma^i)^2/2)t + \sigma^i B_t^i$ for any $i = 1, \ldots, n$. Now, since the exponential map is given by

$$\begin{array}{rcl} \exp : \mathfrak{g} & \longrightarrow & G = \mathbb{R}_+^n \\ \xi = \sum_{i=1}^n \xi^i \xi_i & \longmapsto & \left(\mathrm{e}^{\xi^1}, \ldots, \mathrm{e}^{\xi^n}\right) \end{array}$$

where e^x is the standard exponential function, we recover the well-known result that the general solution q_t of (4.41) starting at $q_0 \in \mathbb{R}_+^n$ is

$$q_t = \left(q_0^1 \, \mathrm{e}^{(\mu^1 - (\sigma^1)^2/2)t + \sigma^1 B_t^1}, \ldots, q_0^n \, \mathrm{e}^{(\mu^n - (\sigma^n)^2/2)t + \sigma^n B_t^n}\right).$$

4.5.4 Brownian motion on reductive homogeneous spaces and symmetric spaces.

Let G a Lie group and $H \subseteq G$ a closed subgroup. We say that the homogeneous space $M = G/H$ is ***reductive*** if the Lie algebra $\mathfrak{g}$ of G may be decomposed into as a direct sum $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ where $\mathfrak{h}$ is the Lie algebra of H and $\mathfrak{m}$ is a subspace invariant under the action of Ad_H. That is, $\mathrm{Ad}_h(\mathfrak{m}) \subseteq \mathfrak{m}$ for any $h \in H$ and, consequently, $[\mathfrak{h}, \mathfrak{m}] \subseteq \mathfrak{m}$. The symmetric spaces introduced in the Example 3.17 are a particular case of reductive homogeneous spaces. Suppose now that the reductive homogeneous space M is Riemann manifold with Riemannian metric η and that the transitive action of G leaves the metric η invariant. As in the Example 3.17, we want to define Brownian motions on (M, η) by reducing a suitable process defined on G.

Let $o \in M$ denote the equivalent class of H in M. We have assumed that (M, η) is a Riemann manifold with a (left) G-invariant metric η. Since η is G-invariant and Φ is transitive, the only thing that really matters as far as the characterization of η is concerned is the symmetric bilinear form $\eta_o : T_oM \times T_oM \to T_oM$. It can be easily proved that there is a natural one-to-one correspondence between the G-invariant Riemannian metrics η on $M = G/H$ and the Ad_H-invariant positive definite symmetric bilinear forms B on $T_oM = \mathfrak{g}/\mathfrak{h}$ ([KN69, Chapter X Proposition 3.1]). The correspondence is given by

$$\eta\left(\xi_1^M, \xi_2^M\right) = B\left(T_e\pi\left(\xi_1\right), T_e\pi\left(\xi_2\right)\right),$$

where $\xi_1, \xi_2 \in \mathfrak{g}$, $\pi : G \to G/H$ is the canonical submersion, and $\xi^M \in \mathfrak{X}(M)$ denotes the infinitesimal generator associated to $\xi \in \mathfrak{g}$. In addition, if M is reductive then the bilinear form B may be regarded as defined on $\mathfrak{m}$, $B : \mathfrak{m} \times \mathfrak{m} \to \mathbb{R}$, since T_oM is naturally isomorphic to $\mathfrak{m}$, which is an Ad_H-invariant subspace of $\mathfrak{g}$. The Riemannian connection ∇ of the metric η associated to such a bilinear form B is given by

$$\nabla_{\xi_1^M} \xi_2^M = \frac{1}{2}\left[\xi_1^M, \xi_2^M\right] + \left(U\left(\xi_1, \xi_2\right)\right)^M, \tag{4.42}$$

([KN69, Chapter X Theorem 3.3]). In this expression ξ_1 and ξ_2 belong to $\mathfrak{m}$ and $U : \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ is the bilinear mapping defined by

$$2B\left(U\left(\xi_1, \xi_2\right), \xi_3\right) = B\left(\xi_1, \left[\xi_3, \xi_2\right]_{\mathfrak{m}}\right) + B\left(\left[\xi_3, \xi_1\right]_{\mathfrak{m}}, \xi_2\right),$$

where $[\cdot,\cdot]_{\mathfrak{m}}$ is such that $[\cdot,\cdot] = [\cdot,\cdot]_{\mathfrak{h}} + [\cdot,\cdot]_{\mathfrak{m}}$ with $[\cdot,\cdot]_{\mathfrak{h}} \in \mathfrak{h}$ and $[\cdot,\cdot]_{\mathfrak{m}} \in \mathfrak{m}$. A consequence of (4.42) is that the Laplacian Δ takes the expression

$$\Delta\left(f\right)(m) = \sum_{i=1}^{r} (\mathcal{L}_{\xi_i^M} \circ \mathcal{L}_{\xi_i^M} + U\left(\xi_i,\xi_i\right)^M)\left(f\right)(m), \quad m \in M = G/K,$$

where $\{\xi_1^M, \ldots, \xi_r^M\}$ is an orthonormal basis of $T_m M$.

As we said, the most important examples of reductive homogeneous spaces are symmetric spaces. In that case, G is the connected component of the isometric group $I(M) \subseteq \mathrm{Diff}(M)$ of the symmetric space (M,η) containing $e = \mathrm{Id}$. We saw in the Example 3.17 that, for a symmetric space, the Lie algebra $\mathfrak{g}$ can be written as $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ such that

$$[\mathfrak{h},\mathfrak{h}] \subseteq \mathfrak{h}, \quad [\mathfrak{h},\mathfrak{m}] \subseteq \mathfrak{m}, \;\; [\mathfrak{m},\mathfrak{m}] \subseteq \mathfrak{h},$$

and $\mathrm{Ad}_H\left(\mathfrak{m}\right) \subseteq \mathfrak{m}$ ([KN69, Chapter XI Proposition 2.1 and 2.2]). Moreover, the symmetric space G/K has a unique affine connection ∇ invariant under the action of G. This is actually the Riemannian connection ([KN69, Chapter XI Theorem 3.3]) so that (4.42) reads

$$\nabla_{\xi_1^M} \xi_2^M = 0$$

for any pair of left-invariant vector fields ξ_1^M and ξ_2^M.

Returning to the general case, let $\{\xi_1, \ldots, \xi_r\}$ be a basis of $\mathfrak{m}$ such that $\{T_e\pi\left(\xi_1\right) \ldots, T_e\pi\left(\xi_r\right)\}$ is an orthonormal basis of $T_o(G/K)$ with respect to η_o and let $\{\xi_1^G, \ldots, \xi_r^G\} \subset \mathfrak{X}\left(G\right)$ be now the corresponding family of *right-invariant* vector fields built from $\{\xi_1, \ldots, \xi_r\}$. Observe that $\{\xi_1^M(m), \ldots, \xi_r^M(m)\}$ is an orthonormal basis of $T_m(G/K)$ due to the transitivity of the action and to the G-invariance of the metric η. Consider now the Stratonovich stochastic differential equation

$$\delta g_t = \sum_{i=1}^{r} \xi_i^G(g_t)\delta B_t^i + \sum_{i=1}^{r} U\left(\xi_i,\xi_i\right)^G(g_t)dt, \tag{4.43}$$

where $\left(B_t^1, \ldots, B_t^r\right)$ is a $\mathbb{R}^r$-valued Brownian motion. The stochastic system (4.43) is by definition K-invariant with respect to the natural *right action* $R : K \times G \to G$, $R_k\left(g\right) = gk$ for any $g \in G$ and $k \in K$. In addition, it is straightforward to check that the projection $\pi : G \to G/K$ send any right-invariant vector field $\xi^G \in \mathfrak{X}\left(G\right)$, $\xi \in \mathfrak{g}$, to the infinitesimal generator $\xi^M \in \mathfrak{X}\left(M\right)$ of the G-action $\Phi : G \times M \to M$. Hence (4.43) projects to the stochastic system

$$\delta\Gamma_t = \sum_{i=1}^{r} \xi_i^M(\Gamma_t)\delta B_t^i + \sum_{i=1}^{r} U\left(\xi_i,\xi_i\right)^M\left(\Gamma_t\right)dt \tag{4.44}$$

on M by Proposition 4.14. It is evident that the solutions of (4.44) have as a generator the second order differential operator $\frac{1}{2}\sum_{i=1}^{r}(\mathcal{L}_{\xi_i^M} \circ \mathcal{L}_{\xi_i^M} + U\left(\xi_i,\xi_i\right)^M)$ and they are therefore Brownian motions.

5
Conclusions and Outlook

The main goal of this thesis is taking advantage of tools coming from differential geometry in the qualitative study of stochastic differential equations. Some questions arose during this work which are still open and may constitute topics of further research. A brief summary and conclusions of the results achieved, as well as hints about the possible directions for future work are presented in what follows.

1. We have generalized the concept of stochastic Hamiltonian system in two aspects. First of all, allowing the driving noise of such a system to be an arbitrary continuous semimartingale instead of a standard Brownian motion and, secondly, defining them on general Poisson manifolds in terms of intrinsically defined structures. It is worth pointing out that the geometric language that we use improves the presentation of stochastic Hamiltonian systems that Bismut made in his seminal work [B81]. This geometric framework turns out to be extremely convenient at the time of studying the qualitative behavior of Hamiltonian systems. Indeed, we could for example give simple criteria that characterize conserved quantities, or to provide sufficient conditions for the stability (almost sure or in probability) of a given equilibrium point (see the Stochastic Dirichlet Criterion Theorem 2.15 and the Stochastic Lyapunov Theorem 2.17). However, the most important contribution in Chapter 2 is the Variational Principle introduced in Theorem 2.34. Not only do we define a stochastic action which can be naturally seen as a generalization of that of a deterministic Hamiltonian systems; more importantly, we also prove that the solutions of the stochastic Hamiltonian equations are in one-to-one correspondence with the semimartingales that make that action critical under a suitable set of (pathwise) variations. This improves other attempts found in the literature. As far as we know, this is the first time that such a characterization is proved.

2. The stochastic action is not only important because it characterizes the solutions of the stochastic Hamiltonian equations but also because, when appropriately regarded as a

function on the configuration space, it satisfies a stochastic version of the Hamilton-Jacobi equation. As a consequence, we showed in Example 2.41 that the stochastic action of some particular Hamiltonian systems can be used to build solutions of the heat equation modified with a potential term. As we saw, these solutions are obtained by exponentiation of the action and then taking expectations. It is worth noticing that except for the absence of the imaginary complex unit i, this equation is formally equivalent to the Schrödinger equation. We must confess that we tried to conveniently insert the complex numbers in the derivation of that heat equation and therefore obtain solutions of the Schrödinger's, obviously following the ideas of the Feynman path integral quantization procedure. Unfortunately, we did not obtain any satisfactory result. However, we think that Example 2.41 and its utility to find solutions of the Schrödinger equation needs further investigation. In particular, it may be useful to write down the wave function of stationary states since, in this particular case, time dependence, and therefore the role of the complex unit i in the Schrödinger equation, can be removed.

3. In Chapter 2, we provide several examples of stochastic Hamiltonian systems. Despite their intrinsic interest, they are rather illustrative and we have the feeling that more physical consequences could be obtained if they were pushed forward. For example, the stochastic models for a randomly perturbed rigid body introduced in Subsection 3.6.3 could be further explored as far as its stability properties is concerned; on the other hand, some numerical schemes such as Monte Carlo methods could be implemented and tested to compute expected values of observable quantities.

   We also have the impression that stochastic Hamiltonian systems may play an important role in the description of some systems arising in statistical physics which, by their own nature, are described in probabilistic terms. Indeed, statistical physics tries to build a bridge between classical mechanics and thermodynamics using probability theory and statistics. Therefore, as the example in Subsection 2.2.3 suggests, the appearance of the Langevin equation as a consequence of an underlying stochastic Hamiltonian equation seems not to be a coincidence. In any case, this relation should be investigated more carefully. It is worth noticing that the use of stochastic processes in the statistical study of some, a priori, deterministic Hamiltonian systems is nothing new. One of the most illustrative examples are the Kac-Zwanzig heat bath models which still deserve a lot of attention from both physicists and mathematicians (see [AV08, K04]). We plan to consider these models in the future.

4. Nowadays, mathematical finance is probably the field where stochastic processes find a larger number of practical applications. Consequently, one lack of the present thesis is the absence of any comment about the stochastic models used in finance, whether Hamiltonian or not. We hope to fix this situation in a near future. In particular, one should consider the stochastic volatility models widely used in the industry. Unlike the standard Black-Scholes model for the price of an asset which assumes constant volatility, stochastic volatility models allow it to depend on time (see for instance [HW78, H93]). Then, volatility evolves with time satisfying a suitable stochastic differential equation and,

consequently, can be regarded as a generalized momentum for the asset price. In other words, stochastic volatility models could be good candidates for Hamiltonian systems.

5. A substantial contributions of this thesis is the systematic treatment of symmetries of stochastic differential equations carried out in Chapter 3, a subject that had so far received little attention in the literature, when compared with its deterministic counterpart. Using mechanics as a model, we showed in the Reduction Theorem 3.9 and the Reconstruction Theorem 3.10 how to cope with the degeneracies of a stochastic differential equation invariant under the (free) action of a Lie group of symmetries. In some cases, we related the presence of such symmetries to the skew-product decomposition of the infinitesimal generators associated to the symmetric stochastic equations under study. As we pointed out in Chapter 3, skew-product decompositions have deserved special attention by some authors in the last decades. For a non-free Lie group action, more interesting decompositions were obtained in Theorem 3.20.

6. The theory of Lie algebroids has recently proved to be extremely fruitful in tackling some problems in the context of geometric mechanics ([CM01]). Nowadays, the formulation of Lagrangian and Hamiltonian mechanics in the context of Lie algebroids is fully understood. Recall that the dual of a Lie algebroid admits a canonical Poisson structure and, therefore, one can naturally consider Hamiltonian systems on them. According to the results and the acceptance of this new formalism, one should investigate the consequences of having stochastic processes taking values on Lie algebroids or their duals for mechanical purposes.

7. In Chapter 4, we introduce the notion of superposition rule for stochastic differential equations. Roughly speaking, a superposition rule exists for stochastic differential equations if (up to a non-zero stopping time) any solution of the equation can be expressed using a given set of particular solutions. In this context, our contribution consist in formulating the stochastic version of the Lie-Scheffers Theorem. Indeed, Theorem 4.7 gives sufficient and necessary conditions for a given stochastic differential equation to admit a superposition rule. The proof of the theorem, which includes the classical Lie-Scheffers Theorem as a particular case, fills the gaps and clarifies some points in the proofs of the previous versions of the deterministic Lie-Scheffers Theorem. Connections with existing results, applications and illustrations of this theorem are given in the setting of Lie groups and homogeneous spaces.

8. Finally, we would like to say a few words about *rough paths*. The theory of rough paths was introduced to define integration with respect to non-differentiable functions ([CLT04]). Unlike stochastic integration, rough paths integration does not refer to any probabilistic concept. However, it extends it in the sense that paths need not have finite quadratic variation. We are convinced that most of the results of the present thesis can be reformulated replacing stochastic differential equations with differential equations driven by rough paths, something that would considerable reduce the difficulty of some proofs. Indeed, dealing with a single path instead of a whole process seems a priori more manageable. It would be advisable to carry out this generalization in a future work.

Appendix A

Auxiliary results about integrals and stopping times

In the following paragraphs we collect some results that are used in Chapter 2 in relation with the interplay between stopping times and integration limits.

Proposition A.1 *Let X be a continuous semimartingale defined on $[0, \zeta_X)$ and Γ a continuous semimartingale. Let $\tau,\ \xi$ be two stopping times such that $\tau \leq \xi < \zeta_X$. Then,*

$$(X \cdot \Gamma)^\tau = \left(\mathbf{1}_{[0,\tau]} X\right) \cdot \Gamma = (X \cdot \Gamma^\tau) \qquad \textit{and} \qquad (X \cdot \Gamma)^\xi - (X \cdot \Gamma)^\tau = \left(\mathbf{1}_{(\tau,\xi]} X\right) \cdot \Gamma$$

An equivalent result holds when dealing with the Stratonovich integral, namely

$$\left(\int X \delta \Gamma\right)^\tau = \int X \delta \Gamma^\tau = \left(\int X^\tau \delta \Gamma\right)^\tau .$$

Proof. By [P05, Theorem 12, page 60] we have that $\mathbf{1}_{[0,\tau]} X \cdot \Gamma = (X \cdot \Gamma)^\tau = (X \cdot \Gamma^\tau)$. Therefore,

$$(X \cdot \Gamma)^\xi - (X \cdot \Gamma)^\tau = \mathbf{1}_{[0,\xi]} X \cdot \Gamma - \mathbf{1}_{[0,\tau]} X \cdot \Gamma = \left[\left(\mathbf{1}_{[0,\xi]} - \mathbf{1}_{[0,\tau]}\right) X\right] \cdot \Gamma = \left(\mathbf{1}_{(\tau,\xi]} X\right) \cdot \Gamma.$$

As to the Stratonovich integral, since X and Γ are semimartingales, we can write [P05, Theorem 23, page 68] that

$$\left(\int X \delta \Gamma\right)^\tau = (X \cdot \Gamma)^\tau + \frac{1}{2} [X, \Gamma]^\tau = (X \cdot \Gamma^\tau) + \frac{1}{2} [X, \Gamma^\tau] = \int X \delta \Gamma^\tau .$$

Finally, observe that for any process, $(X^\tau)^\tau = X^\tau$. On the other hand, taking into account that $\mathbf{1}_{[0,\tau]} X = \mathbf{1}_{[0,\tau]} X^\tau$ and $[\Gamma, X] = [X, \Gamma]$, we have

$$\begin{aligned}
\left(\int X \delta \Gamma\right)^\tau &= \mathbf{1}_{[0,\tau]} X \cdot \Gamma + \frac{1}{2} [X, \Gamma]^\tau = \mathbf{1}_{[0,\tau]} X^\tau \cdot \Gamma + \left(\frac{1}{2} [X, \Gamma]^\tau\right)^\tau \\
&= (X^\tau \cdot \Gamma)^\tau + \left(\frac{1}{2} [X^\tau, \Gamma]\right)^\tau = \left(X^\tau \cdot \Gamma + \frac{1}{2} [X^\tau, \Gamma]\right)^\tau = \left(\int X^\tau \delta \Gamma\right)^\tau .
\end{aligned}$$

■

Proposition A.2 *Let $X : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ be a real valued process. Let $\{\tau_n\}_{n\in\mathbb{N}}$ be a sequence of stopping times such that a.s. $\tau_0 = 0$, $\tau_n \leq \tau_{n+1}$, for all $n \in \mathbb{N}$, and $\sup_{n\in\mathbb{N}} \tau_n = \infty$. Then,*

$$X = \lim_{\substack{ucp \\ n\to\infty}} X^{\tau_n}.$$

In particular, if $\Gamma : \mathbb{R}_+ \times \Omega \to M$ is a continuous M-valued semimartingale and $\eta \in \Omega_2(M)$ then,

$$\int \langle \eta, d\Gamma \rangle = \lim_{\substack{ucp \\ k\to\infty}} \left(\int \langle \eta, d\Gamma \rangle \right)^{\tau_k} = \lim_{\substack{ucp \\ k\to\infty}} \sum_{n=0}^{k-1} \int \mathbf{1}_{(\tau_n, \tau_{n+1}]} \langle \eta, d\Gamma \rangle .$$

Proof. Let $\varepsilon > 0$ and $t \in \mathbb{R}_+$. Then for any $s \in [0, t]$ one has

$$\{|X^{\tau_n} - X|_s > \varepsilon\} \subseteq \{\tau_n < s\} \subseteq \{\tau_n < t\}.$$

Hence for any $t \in \mathbb{R}_+$

$$P(\{|X^{\tau_n} - X|_s > \varepsilon\}) \leq P(\{\tau_n < t\}).$$

The result follows because $P(\{\tau_n < t\}) \to 0$ as $n \to \infty$ since $\tau_n \to \infty$ a.s., and hence in probability. Let now Γ be a M-valued continuous semimartingale and $\eta \in \Omega_2(M)$. Notice first that $\left(\int \langle \eta, d\Gamma \rangle \right)^{\tau_0} = 0$ because $\tau_0 = 0$. Consequently, by Proposition A.1 we can write

$$\left(\int \langle \eta, d\Gamma \rangle \right)^{\tau_k} = \sum_{n=0}^{k-1} \left(\int \langle \eta, d\Gamma \rangle \right)^{\tau_{n+1}} - \left(\int \langle \eta, d\Gamma \rangle \right)^{\tau_n} = \sum_{n=0}^{k-1} \int \mathbf{1}_{(\tau_n, \tau_{n+1}]} \langle \eta, d\Gamma \rangle$$

and the result follows. ■

Lemma A.3 *Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of real processes converging in ucp to a process X. Let τ be a stopping time such that $\tau < \infty$ a.s.. Then, the sequence of random variables $\{(X_n)_\tau\}_{n\in\mathbb{N}}$ converge in probability to $(X)_\tau$.*

Proof. First of all we show that since $\tau < \infty$ a.s., then $P(\{\tau > t\})$ converges to zero as $t \to \infty$. By contradiction, suppose that this is not the case. Then, denoting $A_n := \{\tau > n\}$, we have that $A_{n+1} \subseteq A_n$, so $P(A_n)$ forms a non-increasing sequence of real numbers in the interval $[0, 1]$. Since this sequence is bounded below, it must have a limit. This limit corresponds to the probability of the event $\{\tau = \infty\}$. If it is strictly positive then there is a contradiction with the fact that $\tau < \infty$ a.s.. So $P(\{\tau > t\})$ tends to zero as $t \to \infty$.

We now prove the statement of the lemma. Take some $\varepsilon > 0$ and an auxiliary $t \in \mathbb{R}_+$. The set $\{|(X_n)_\tau - X_\tau| > \varepsilon\}$ can be decomposed as the disjoint union of the following two events,

$$(\{|(X_n)_\tau - X_\tau| > \varepsilon\} \cap \{\tau \leq t\}) \bigcup (\{|(X_n)_\tau - X_\tau| > \varepsilon\} \cap \{\tau > t\}).$$

The first one is contained in the set $\left\{\sup_{0\leq s\leq t} |(X_n)_s - X_s| > \varepsilon\right\}$ whose probability, by hypothesis, converges to zero as $n \to \infty$. Regarding the second one,

$$P(\{|(X_n)_\tau - X_\tau| > \varepsilon\} \cap \{\tau > t\}) \leq P(\{\tau > t\}).$$

But $P(\{\tau > t\})$ can be made arbitrarily small by taking the auxiliary t big enough. In conclusion, for any $\varepsilon > 0$,

$$P(\{|(X_n)_\tau - X_\tau| > \varepsilon\}) \underset{n\to\infty}{\longrightarrow} 0$$

in probability. ■

Lemma A.4 *Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of real processes converging in ucp to a real process X and τ a stopping time. Then, the stopped sequence $\{X_n^\tau\}_{n\in\mathbb{N}}$ converges in ucp to X^τ as well.*

Proof. We just need to observe that, for any $t \in \mathbb{R}_+$,

$$\sup_{0\le s\le t} |(X_n^\tau)_s - X_s^\tau| = \sup_{0\le s\le t} |(X_n)_{\tau\wedge s} - X_{\tau\wedge s}| \le \sup_{0\le s\le t} |(X_n)_s - X_s|$$

and, consequently, for any $\varepsilon > 0$,

$$\left\{\sup_{0\le s\le t} |(X_n)_s - X_s| \le \varepsilon\right\} \subseteq \left\{\sup_{0\le s\le t} |(X_n^\tau)_s - X_s^\tau| \le \varepsilon\right\}.$$

Hence, since by hypothesis $P\left(\left\{\sup_{0\le s\le t} |(X_n)_s - X_s| \le \varepsilon\right\}\right)$ converges to 1 as $n \to \infty$, then so does $P\left(\left\{\sup_{0\le s\le t} |(X_n^\tau)_s - X_s^\tau| \le \varepsilon\right\}\right)$. ■

Proposition A.5 *Let X and Y be two real semimartingales. Suppose that X is continuous and $X_0 = 0$. Then, for any $t \in \mathbb{R}_+$, the Stratonovich integral $\int \left(\mathbf{1}_{[0,t]}Y\right)\delta X$ is well defined and equal to $\left(\int Y\delta X\right)^t$.*

Proof. If $\int \left(\mathbf{1}_{[0,t]}Y\right)\delta X$ was well defined, it should be equal to $\int \left(\mathbf{1}_{[0,t]}Y\right)dX + \frac{1}{2}\left[\mathbf{1}_{[0,t]}Y, X\right]$. Since $\int \left(\mathbf{1}_{[0,t]}Y\right)dX$ is well defined, the only thing that we need to check is that $\left[\mathbf{1}_{[0,t]}Y, X\right]$ exists. On the other hand, recall that ([P05, Theorem 12 page 60 and Theorem 23 page 68])

$$\left(\int Y\delta X\right)^t = \int \left(\mathbf{1}_{[0,t]}Y\right)dX + \frac{1}{2}[Y,X]^t = \int \left(\mathbf{1}_{[0,t]}Y\right)dX + \frac{1}{2}\left[Y^t, X\right].$$

Hence, what we are actually going to proceed by showing that $\left[\mathbf{1}_{[0,t]}Y, X\right]$ is equal to $\left[Y^t, X\right]$. Let $\sigma_n = \left\{0 = T_0^n \le T_1^n \le ... \le T_{k_n}^n < \infty\right\}$ be a sequence of random partitions tending to the identity (in the sense of [P05, page 64]). Given two real processes X and Y, their quadratic variation, if it exists, can be defined as the limit in *ucp* when $n \to \infty$ of the following sums

$$[Y,X] = \lim_{\substack{ucp \\ n\to\infty}} \sum_{i=0}^{k_n-1} \left(Y^{T_{i+1}^n} - Y^{T_i^n}\right)\left(X^{T_{i+1}^n} - X^{T_i^n}\right).$$

Let now

$$H_n := \sum_{i=0}^{k_n-1} \left(\left(Y^t\right)^{T_{i+1}^n} - \left(Y^t\right)^{T_i^n}\right)\left(X^{T_{i+1}^n} - X^{T_i^n}\right),$$

$$G_n := \sum_{i=0}^{k_n-1} \left(\left(\mathbf{1}_{[0,t]}Y\right)^{T_{i+1}^n} - \left(\mathbf{1}_{[0,t]}Y^t\right)^{T_i^n}\right)\left(X^{T_{i+1}^n} - X^{T_i^n}\right).$$

It is clear that the sequence $\{H_n\}_{n\in\mathbb{N}}$ converges uniformly on compacts in probability to $\left[Y^t, X\right]$. We are going to prove that there exists such a convergence for the sequence of processes $\{G_n\}_{n\in\mathbb{N}}$ by showing that the elements $(G_n)_s$ coincide with $(H_n)_s$, for any $s \in \mathbb{R}_+$, up to a set whose probability tends to zero as $n \to \infty$. We will consider two cases:

1. The case $s \leq t$. Given a specific $i \in \{0, ..., k_n - 1\}$, and recalling that by construction $T_i^n \leq T_{i+1}^n$ a.s., it is clear that $\left(\left(Y^t\right)^{T_{i+1}^n} - \left(Y^t\right)^{T_i^n}\right)_s = Y_{T_{i+1}^n \wedge s} - Y_{T_i^n \wedge s}$ is different from 0 only for those $\omega \in \Omega$ in $\{T_i^n < s\}$ in which case it takes the value

$$Y_{T_{i+1}^n \wedge s} - Y_{T_i^n}. \tag{A.1}$$

On the other hand, $\left(\left(\mathbf{1}_{[0,t]} Y\right)^{T_{i+1}^n} - \left(\mathbf{1}_{[0,t]} Y^t\right)^{T_i^n}\right)_s$ is again different from 0 only in the set $\{T_i^n < s\}$ and there it is equal to (A.1). Therefore, $(G_n)_s = (H_n)_s$ whenever $s \leq t$.

2. The case $s > t$. In this case, $\left(\left(Y^t\right)^{T_{i+1}^n} - \left(Y^t\right)^{T_i^n}\right)_s = Y_{t \wedge T_{i+1}^n} - Y_{t \wedge T_i^n}$ which is different from 0 only in the set $\{T_i^n < t\}$, where it takes the value

$$Y_{t \wedge T_{i+1}^n} - Y_{T_i^n} \tag{A.2}$$

However, in this case $\left(\left(\mathbf{1}_{[0,t]} Y\right)^{T_{i+1}^n} - \left(\mathbf{1}_{[0,t]} Y^t\right)^{T_i^n}\right)_s = \mathbf{1}_{\left\{T_{i+1}^n \leq t\right\}} Y_{t \wedge T_{i+1}^n} - \mathbf{1}_{\left\{T_i^n \leq t\right\}} Y_{t \wedge T_i^n}$, which is equal to (A.2) in the set $\left\{T_{i+1}^n \leq t\right\}$ (which contains $\{T_i^n < t\}$ since $T_i^n \leq T_{i+1}^n$), but differs from (A.2) in

$$A_i^n(t) := \left\{T_i^n \leq t < T_{i+1}^n\right\}$$

where it takes the value $-Y_{T_i^n}$. For any other $\omega \in \Omega$ not in these sets,

$$\left(\left(\mathbf{1}_{[0,t]} Y\right)^{T_{i+1}^n} - \left(\mathbf{1}_{[0,t]} Y^t\right)^{T_i^n}\right)_s (\omega) = 0.$$

Therefore, whenever $s > t$, $(G_n)_s$ and $(H_n)_s$ are different only for the $\omega \in A_i^n(t)$. Observe that, since t is fixed, only one of the sets $\{A_i^n(t)\}_{i\in\{0,\ldots,k_n-1\}}$ is non-empty and, on it,

$$(H_n)_s - (G_n)_s = Y_t \left(X_t - X_{T_i^n}\right).$$

To sum up, the analysis that we just carried out shows that for any $u \in \mathbb{R}_+$

$$\sup_{0\leq s\leq u} \left|(H_n)_s - (G_n)_s\right| = \mathbf{1}_{A_i^n(t)} \left|Y_t\right| \left|\left(X_t - X_{T_i^n}\right)\right|$$

for some $i \in \{0, ..., k_n - 1\}$. If X is continuous, this expression tells us that

$$\sup_{0\leq s\leq u} \left|(H_n)_s - (G_n)_s\right| \to 0 \text{ a.s. as } n \to \infty$$

which, in turn, implies that $\sup_{0\leq s\leq u} \left|(H_n)_s - (G_n)_s\right|$ converges to 0 in probability as well. That is, for any $\varepsilon > 0$,

$$P\left(\left\{\sup_{0\leq s\leq u} \left|(H_n)_s - (G_n)_s\right| > \varepsilon\right\}\right) \to 0, \quad \text{as } n \to \infty,$$

which is the same as saying that $H_n - G_n$ converges to 0 in *ucp*. Thus, since $G_n = H_n - (H_n - G_n)$ and the limit in *ucp* as $n \to \infty$ exist for the both sequences $\{H_n\}_{n\in\mathbb{N}}$ and $\{H_n - G_n\}_{n\in\mathbb{N}}$, so does the limit of $\{G_n\}_{n\in\mathbb{N}}$ which, by definition, is the quadratic variation $\left[\mathbf{1}_{[0,t]}Y, X\right]$. Moreover, as $(H_n - G_n) \to 0$ in *ucp* as $n \to \infty$,

$$\left[Y^t, X\right] = \lim_{\substack{ucp \\ n\to\infty}} H_n = \lim_{\substack{ucp \\ n\to\infty}} G_n = \left[\mathbf{1}_{[0,t]}Y, X\right],$$

which concludes the proof. ■

References

[AM78] Abraham, R., and Marsden, J.E. *Foundations of Mechanics*. Second edition, Addison-Wesley (1978).

[AV08] Ariel, G. and Vanden-Eijnden, E. A strong limit theorem in Kac-Zwanzig model. To appear in *Communications in Mathematical Physics* (2008/2009).

[A03] Arnold, L. *Random Dynamical Systems*. Springer Monographs in Mathematics. Springer Verlag (2003).

[A89] Arnold, V.I. *Mathematical Methods of Classical Mechanics*. Second edition. Volume **60** of *Graduate Texts in Mathematics*, Springer Verlag (1989).

[B04] Baudoin, F. *An Introduction to the Geometry of Stochastic Flows*. Imperial College Press (2004).

[B89] Ben Arous, G. Flots et series de Taylor stochastiques. *Probab. Th. Rel. Fields* **81**, 29-77 (1989).

[BB82] Bérard, L. and Bourguignon, J.-P. Laplacians and Riemannian submersions with totally geodesic fibers. *Illinois Journal of Mathematics*, **26** (2), 181-200 (1982).

[BP06] Bingham, N. H. and Peskir, G. Optimal stopping and dynamic programming. *Lecture Notes No.* **1**, Probab. Statist. Group Manchester (2006). Prepared for *Encyclopedia of Quantitative Risk Assessment* (John Wiley & Sons).

[B81] Bismut, J.-M. *Mécanique Aléatoire*. Lecture Notes in Mathematics, volume **866**. Springer-Verlag (1981).

[BG68] Blumenthal, R. M. and Getoor, R. K. *Markov Processes and Potential Theory*. Academic Press (1968).

[BO07] Bou-Rabee, N. and Owhadi, H. Stochastic variational integrators. `arXiv:0708.2187` (2007).

[BO07a] Bou-Rabee, N. and Owhadi, H. Stochastic variational partitioned Runge-Kutta integrators for constrained systems. `arXiv:0709.2222` (2007).

[BJ76] Box, G. E. P. and Jenkins, G. M. *Time Series Analysis: Forecasting and Control*. Holden-Day (1976).

[CGM00] Cariñena, J. F., Grabowski, J., and Marmo, G. *Lie-Scheffers Systems: A Geometric Approach*. Napoli Series on Physics and Astrophysics. Vol. **3** (2000).

[CGM07] Cariñena, J. F., Grabowski, J., and Marmo, G. Superposition rules, Lie theorem, and partial differential equations. *Reports on Mathematical Physics*, **60** (2), 237-258 (2007).

[CMN98] Cariñena, J. F., Marmo, G., and Nasarre, J. *Int. J. Mod. Phys.* **A 13**, 3601-27 (1998).

[CM01] Cariñena, J. F. and Martínez, E. Lie algebroid generalization of geometric mechanics. *Lie Algebroids*. Banach Centre Publications, Vol. **54** (2001).

[CR01] Cariñena, J. F. and Ramos, A. A new geometric approach to Lie systems and physical applications. *Acta Appl. Math.*, **70**(1-3), 43-69 (2002).

[CLT04] Caruana, M., Lyons, T. J., and Lévy, T. *Differential Equations Driven by Rough Paths*. École d'Été de Probabilités de Saint-Flour **XXXIV**. Springer (2004).

[C93] Castell, F. Asymptotic expansion of stochastic flows. *Probability Theory and Related Fields* **96**, 225-239 (1993).

[C01] Catuogno, P. Second order connections and stochastic horizontal lifts. *J. Geom. Phys.*, **56** (vol. 9), 1637–1653 (2006).

[CH06] Chorin, A.J. and Hald, O.H. *Stochastic Tools in Mathematics and Science*. Surveys and Tutorials in the Applied Mathematical Sciences, volume 1. Springer Verlag (2006).

[ChL00] Chossat, P. and Lauterbach, R. *Methods in Equivariant Bifurcations and Dynamical Systems*. Advanced Series in Nonlinear Dynamics, volume **15**, World Scientific (2000).

[CW90] Chung, K. L. and Williams, R. J. *Introduction to Stochastic Integration*. Second edition. Probability and its Applications. Birkhäuser Verlag (1990).

[CD06] Cresson, J. and Darses, S. Plongement stochastique des systèmes lagrangiens. *C. R. Math. Acad. Sci. Paris*, **342**, 333-336 (2006).

[D85] Dazord, P. Feuilletages à singularités. *Nederl. Akad. Wetensch. Indag. Math.*, **47**, 21-39 (1985).

[DK99] Duistermaat, J.J. and Kolk, J.A. *Lie Groups*. Universitext, Springer-Verlag (1999).

[D96] Durrett, R. *Stochastic Calculus. A Practical Introduction.* Probability and Stochastics Series. CRC Press, (1996).

[E82] Elworthy, K.D. *Stochastic Differential Equations on Manifolds*. London Mathematical Society Lecture Note Series, vol. **70** (1982).

[EK85] Elworthy, K. D. and Kendall, W. S. Factorization of harmonic maps and Brownian motions. *Local time to global geometry, control and physics*. Notes in Math. Ser. **150**, 75-83 (1985).

[ELL98] Elworthy, K. D., Le Jan, Y., and Li, X.-M. *On the Geometry of Diffusion Operators and Stochastic Flows*. Lecture Notes in Mathematics **1720**. Springer (1999).

[ELL04] Elworthy, K. D., Le Jan, Y., and Li, X.-M. Equivariant diffusions on principal bundles. *Advanced studies in pure mathematics*. Vol. **41** Stochastic analysis and related topics, 31-47 (2004).

[E89] Émery, M. *Stochastic Calculus in Manifolds*. Springer-Verlag (1989).

[E90] Émery, M. On two transfer principles in stochastic differential geometry. *Séminaire de Probabilités*, **XXIV**, 407-441, Lecture Notes in Math., **1426**, Springer Verlag, (1990). Correction: *Séminaire de Probabilités*, **XXVI**, 633, Lecture Notes in Math., **1526**, Springer Verlag, (1992).

[EP01] Estrade, A., Pontier M. Backward stochastic differential equations in a Lie group. *Séminaire de probabilités (Strasbourg)*. Tome **35**, 241-259 (2001).

[F06] Ferguson, T. *Optimal Stopping and Applications*. Available at `http://www.math.ucla.edu/~tom/`.

[F91] Field, M. J. Local structure of equivariant dynamics. In *Singularity Theory and its Applications*. M. Roberts and I. Stewart eds. Lecture Notes in Mathematics, 1463. Springer–Verlag (1991).

[G96] Gihman, I. I. Stability of solutions of stochastic differential equations. (Russian) Limit Theorems Statist. Inference (Russian) pp. 14–45 Izdat. "Fan", Tashkent. English translation: Selected Transl. Statist. and Probability, **12**, Amer. Math. Soc., pp. 125-154. MR **40**, number 944 (1966).

[GS02] Golubitsky, M. and Stewart, I. *The Symmetry Perspective*. Progress in Mathematics, volume 200. Birkhäuser (2002).

[GS85] Golubitsky, M., Stewart, I., and Schaeffer, D.G. *Singularities and Groups in Bifurcation Theory: Vol. II*. Applied Mathematical Sciences, Vol. **69**, Springer–Verlag (1985).

[GS01] Grimmet, G. and Stirzaker, D. *Probability and Random Processes*. 3rd edition. Oxford university press (2001).

[GS83] Guillemin, V., and Sternberg, S. On collective complete integrability according to the method of Thimm. *Ergod. Th. & Dynam. Sys.* **3**, 219-230 (1983).

[GS80] Guillemin, V., and Sternberg, S. The moment map and collective motion. *Annals of Physics* **127**, 220-253 (1980).

[HL86] Hakim-Dowek, M. and Lepingle D. L'exponentielle stochastique des groupes de Lie. *Séminaire de probabilités (Strasbourg).* Tome **20**, 352-374, Springer Verlag (1986).

[H80] Hasminskii, R. Z. *Stochastic Stability of Differential Equations.* Translated from the Russian by D. Louvish. Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis, **7**. Sijthoff and Noordhoff, Alphen aan den Rijn—Germantown, Md (1980).

[H78] Helgason, S. *Differential Geometry, Lie Groups, and Symmetric Spaces.* Pure and Applied Mathematics. Academic Press (1978).

[H60] Hermann, R. A sufficient condition that a mapping of Riemannian manifolds be a fibre bundle. *Proc. Amer. Math. Soc.*, **11**, 236-242 (1960).

[H93] Heston, S.L. A closed-form solution for options with stochastic volatility, with applications to bond and currency options. *The Review of Financial Studies* **6** (2) (1993).

[H02] Hsu, E. P. *Stochastic Analysis on Manifolds.* Graduate Studies in Mathematics, **38**. American Mathematical Society (2002).

[H92] Hu, Y.-Z. Série de Taylor stochastique et formule de Campbell-Hausdorff, d'après Ben Arous. *Séminaire de probabiliés (Strasbourg).* Tome **26**, 579-586 (1992).

[HW78] Hull, J. and A. White, A. The pricing of options on assets with stochastic volatilities, *The Journal of Finance*, **XLII** (2) (1987).

[IW89] Ikeda, N. and Watanabe, S. *Stochastic Differential Equations and Diffusion Processes.* Second edition. North-Holland Mathematical Library, 24. North-Holland Publishing Co (1989).

[I01] Imkeller, P. and Lederer, C.: Some formulas for Lyapunov exponents and rotation numbers in two dimensions and the stability of the harmonic oscillator and the inverted pendulum. *Dyn. Syst.*, **16**(1), 29-61 (2001).

[KS91] Karatzas, I. and Shreve S. E. *Brownian Motion and Stochastic Calculus.* 2nd edition. Graduate texts in mathematics **113**. Springer (1991).

[K97] Knight, F. B. *Essentials of Brownian Motion and Diffusions.* 2nd Edition. American Mathematical Society (1997).

[KN69] Kobayashi, S. and Nomizu, K. *Foundations of Differential Geometry* (vol. II). Tracts in mathematics **15**. Interscience Publishers. John Wiley and Sons (1969).

[KMS93] Kolář, I., Michor, P., and Slovák, J. *Naturals Operations in Differential Geometry.* Electronic edition. Originally published by Springer-Verlag (1993).

[K78] Kolmogorov, A. N. and Fomin, S. V. *Elementos de la Teoría de Funciones y del Análisis Funcional.* Third edition. Mir (1978).

[Ko53] Koszul, J. L. Sur certains groupes de transformation de Lie. In *Colloque de Géométrie Différentielle*, volume 71 in *Colloques du CNRS*, pages 137–141 (1953).

[K90] Krupa, M. Bifurcations of relative equilibria. *SIAM J. Math. Anal.* **21**(6), 1453–1486 (1990).

[K80] Kunita, H. On the representation of solutions of stochastic differential equations. *Séminaire de probabilités (Strasbourg).* Tome **14**, 282-304 (1980).

[K81] Kunita, H. Some extensions of Itô's formula. *Seminaire de probabilités de Strasbourg* XV, 118-141. Lecture Notes in Mathematics, **850**, Springer Verlag (1981).

[K90] Kunita, H. *Stochastic Flows and Stochastic Differential Equations.* Cambridge studies in advanced mathematics **24**. Cambridge university press (1990).

[K04] Kupferman, R. Fractional kinetics in Kac–Zwanzig heat bath models. *Journal of Statistical Physics*, Vol. **114** (1/2), 291-326 (2004).

[LL76] Landau, L.D., and Lifshitz, E.M. *Mechanics.* Volume 1 of *Course of Theoretical Physics.* Third Edition. Pergamon Press (1976).

[LO07] Lázaro-Camí, J.-A. and Ortega, J.-P. Stochastic Hamiltonian dynamical systems. *Reports on Mathematical Physics* Vol. **61**, No. 1, 65-122 (2008).

[LO08] Lázaro-Camí, J.-A. and Ortega, J.-P. Reduction, reconstruction, and skew-product decomposition of symmetric stochastic differential equations. To appear in *Stochastics and Dynamics.* Extended version available at `arXiv:0705.3156v2 [math.PR]`(2008).

[LO08a] Lázaro-Camí, J.-A. and Ortega, J.-P. Superposition rules and stochastic Lie-Scheffers systems. To appear in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.* `arXiv:0803.0600v1 [math.PR]`(2008).

[LO08b] Lázaro-Camí, J.-A. and Ortega, J.-P. Stochastic Hamilton-Jacobi equation (2008).

[LG97] Le Gall, J.-F. *Mouvement Brownian et Calcul Stochastique.* Notes de Cours DEA 1996-97. Available at http://www.dma.ens.fr/ legall/ (1997).

[Li08] Li, X.-M. An averaging principle for integrable stochastic Hamiltonian systems. *Nonlinearity* **21** (4), 803–822 (2008). *Preprint available at* http://www.lboro.ac.uk/departments/ma/research/preprints/papers06/06-27.pdf.

[L89] Liao, M. Factorization of diffusions on fiber bundles. *Transactions of the American Mathematical Society*, **31**(2), 813-827 (1989).

[L97] Liao, M. Random motion of a rigid body. *Journal of Theoretical Probability,* **10** (vol. 1), 201-211 (1997).

[L04] Liao, M. *Lévy Processes in Lie Groups.* Cambridge University Press (2004).

[L07] Liao, M. A decomposition of Markov processes via group actions. Preprint (2007).

[LW05] Liao, M. and Wang, L. Motion of a rigid body under random perturbation. *Elect. Comm. in Probab.* **10**, 235-243 (2005).

[Lie93] Lie, S. Vorlesungen über Continuierliche Gruppen mit Geometrischen und Anderen Andwendungen. Edited and revised by G. Scheffers, Teubner, Leipzig (1893).

[Ma78] Malliavin, P. Géométrie différentielle stochastique. *Séminaire de Mathématiques Supérieures*, **64**. Presses de l'Université de Montréal (1978).

[MR99] Marsden, J. E. and Ratiu, T. S. *Introduction to Mechanics and Symmetry.* Second edition. Springer (1999).

[MW74] Marsden, J.E., and Weinstein, A. Reduction of symplectic manifolds with symmetry. *Rep. Math. Phys.*, **5**(1), 121–130 (1974).

[M81] Meyer, P.-A. Géométrie stochastique sans larmes. *Séminaire de probabilités (Strasbourg).* Tome **15**, 44–102 (1981).

[M82] Meyer, P.-A. Géométrie différentielle stochastique. II. *Séminaire de probabilités (Strasbourg).* Tome **S16**, 165–207 (1982).

[MZ84] Meyer, P.-A. and Zheng, W. A. Quelques résultats de "mécanique stochastique". *Séminaire de probabilités (Strasbourg).* Tome **18**, 223–244 (1984).

[M99] Misawa, T. Conserved quantities and symmetries related to stochastic dynamical systems. *Ann. Inst. Statist. Math.*, **51**(4), 779–802 (1999).

[N66] Nagano, T. Linear differential systems with singularities and an application to transitive Lie algebras. *J. Math. Soc. Japan.* Vol. **18** (4), 398-404 (1966).

[N67] Nelson, E. *Dynamical Theories of Brownian Motion.* Princeton University Press (1967).

[N95] Nualart, D. *The Malliavin Calculus and Related Topics.* Springer-Verlag (1995).

[O03] Øksendal, B. *Stochastic Differential Equations.* Sixth Edition. Universitext. Springer-Verlag (2003).

[O83] O'Neill, B. *Semi-Riemannian Geometry. With Applications to Relativity.* Pure and Applied Mathematics, volume **103**. Academic Press (1983).

[O02] Ortega, J.–P. The symplectic reduced spaces of a Poisson action. *C. R. Acad. Sci. Paris Sér. I Math.*, **334**, 999–1004 (2002).

[OP04] Ortega, J.-P. and Planas-Bielsa, V.: Dynamics on Leibniz manifolds. *J. Geom. Phys.*, **52**(1), 1-27 (2004).

[OR02] Ortega, J.–P. and Ratiu, T. S. The optimal momentum map. *Geometry, Dynamics, and Mechanics: 60th Birthday Volume for J.E. Marsden*, pages 329-362. P. Holmes, P. Newton, and A. Weinstein, eds., Springer-Verlag. Available at http://arXiv.org/abs/math.SG/0203040 (2002).

[OR04] Ortega, J.-P. and Ratiu, T.S. *Momentum Maps and Hamiltonian Reduction*. Progress in Mathematics, **222**. Birkhäuser-Verlag, Boston (2004).

[OR06] Ortega, J.-P. and Ratiu, T. S. The reduced spaces of a symplectic Lie group action. *Annals of Global Analysis and Geometry*, **30**(4), 335-381 (2006).

[OR06a] Ortega, J.-P. and Ratiu, T. S. The stratified spaces of a symplectic Lie group action. *Reports on Mathematical Physics*, **58**, 51-75 (2006).

[O06] Ovseyevich, A. I. The stability of an inverted pendulum when there are rapid random oscillations of the suspension point. *Journal of Applied Mathematics and Mechanics*, **70**, 762-768 (2006).

[P57] Palais, R. A global formulation of the Lie theory on transformation groups. *Mem. Am. Math. Soc.* **22**, 95-97 (1957).

[P61] Palais, R. On the existence of slices for actions of non–compact Lie groups. *Ann. Math.*, **73**, 295–323 (1961).

[PWZ33] Paley, R. E. A. C., Wiener, N., and Zygmund, A. Note on random functions. *Math. Z.* **37**, 647-668 (1933).

[PR88] Pauwels, E. J. and Rogers, L. C. G. Skew-product decompositions of Brownian motions. *Contemporary Mathematics*, **73**, 237-262 (1988).

[P05] Protter, P. *Stochastic Integration and Differential Equations. A New Approach*. Applications of Mathematics, volume **21**. Second Edition. Springer-Verlag (2005).

[R87] Rudin, W. *Real and Complex Analysis*. 3rd edition. Higher Mathematics Series. McGraw-Hill International Editions (1987).

[S82] Schwartz, L. Géométrie différentielle du 2ème ordre, semi-martingales et équations différentielles stochastiques sur une variété différentielle. *Séminare de probabilités (Strasbourg)*. Tome **S16**, 1–148 (1982).

[S82] Shigekawa, I. On stochastic horizontal lifts. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **59**, 211-221 (1982).

[SL91] Sjamaar, R., and Lerman, E. Stratified symplectic spaces and reduction. *Ann. of Math.*, **134**, 375–422 (1991).

[St74a] Stefan, P. Accessibility and foliations with singularities. *Bull. Amer. Math. Soc.*, **80**, 1142-1145 (1974).

[St74b] Stefan, P. Accessible sets, orbits and foliations with singularities. *Proc. Lond. Math. Soc.*, **29**, 699-713 (1974).

[Su73] Sussman, H. Orbits of families of vector fields and integrability of distributions. *Trans. Amer. Math. Soc.*, **180**, 171-188 (1973).

[T92] Taylor, J. C. Skew products, regular conditional probabilities, and stochastic differential equations: a technical remark. *Séminaire de probabilités (Strasbourg).* Tome **26**, 113-126 (1992).

[TZ97] Thieullen, M. and Zambrini, J. C. Probability and quantum symmetries. I. The theorem of Noether in Schrödinger's Euclidean quantum mechanics. *Ann. Inst. H. Poincaré Phys. Théor.*, **67**(3), 297–338 (1997).

[TZ97a] Thieullen, M. and Zambrini, J. C. Symmetries in the stochastic calculus of variations. *Probab. Theory Related Fields*, **107**(3), 401–427 (1997).

[W80] Watanabe S. Differential and variation for flow of diffeomorphisms defined by stochastic differential equations on manifolds (in Japanese). Sūkaiken Kōkyuroku, **391** (1980).

[WN63] Wei, J. and Norman, E. *J. Math. Phys.* **4**, 575-81 (1963).

[WN64] Wei, J. and Norman, E. *Proc. Amer. Math. Soc.* **15**, 327-34 (1964).

[Y81] Yasue, K. Stochastic calculus of variations. *Journal of Functional Analysis*, **41**, 327-340 (1981).

[YZ82] Yasue, K. and Zambrini, J.-C. Semi-classical quantum mechanics and stochastic calculus of variations. *Annals of Physics*, **143**, 54-83 (1982).

[Y71] Yosida, K. *Functional Analysis.* Third edition. Die Grundlehren der mathematischen Wissenschaften, **123**. Springer-Verlag (1971).