Stochastic embedding of dynamical systems

Jacky Cresson, Sébastien Darses

Jacky CRESSON
Sébastien DARSES

STOCHASTIC EMBEDDING OF DYNAMICAL SYSTEMS

Jacky CRESSON

Université de Franche-Comté, Équipe de Mathématiques de Besançon, CNRS-UMR 6623, 16 route de Gray, 25030 Besançon cedex, France..

E-mail : cresson@math.univ-fcomte.fr

Sébastien DARSES

Université de Franche-Comté, Équipe de Mathématiques de Besançon, CNRS-UMR 6623, 16 route de Gray, 25030 Besançon cedex, France..

E-mail : darses@math.univ-fcomte.fr

2000 ***Mathematics Subject Classification.*** — Primary 54C40, 14E20; Secondary 46E25, 20C20.

Key words and phrases. — Stochastic calculus, Dynamical systems, Lagrangian systems, Hamiltonian systems.

July 2, 1991

STOCHASTIC EMBEDDING OF DYNAMICAL SYSTEMS

Jacky CRESSON, Sébastien DARSES

Abstract. — Most physical systems are modelled by an ordinary or a partial differential equation, like the n-body problem in celestial mechanics. In some cases, for example when studying the long term behaviour of the solar system or for complex systems, there exist elements which can influence the dynamics of the system which are not well modelled or even known. One way to take these problems into account consists of looking at the dynamics of the system on a larger class of objects, that are eventually stochastic. In this paper, we develop a theory for the stochastic embedding of ordinary differential equations. We apply this method to Lagrangian systems. In this particular case, we extend many results of classical mechanics namely, the least action principle, the Euler-Lagrange equations, and Noether's theorem. We also obtain a Hamiltonian formulation for our stochastic Lagrangian systems. Many applications are discussed at the end of the paper.

CONTENTS

INTRODUCTION

Ordinary as well as partial differential equations play a fundamental role in most parts of mathematical physics. The story begins with Newton's formulation of the law of attraction and the corresponding equations which describe the motion of mechanical systems. Regardless the beauty and usefulness of these theories in the study of many important natural phenomena, one must keep in mind that they are based on experimental facts, and as a consequence are only an approximation of the real world. The basic example we have in mind is the motion of the planets in the solar system which is usually modelled by the famous n-body problem, i.e. n points of mass m_i which are only submitted to their mutual gravitational attraction. If one looks at the behaviour of the solar system for finite time then this model is a very good one. But this is not true when one looks at the long term behaviour, which is for instance relevant when dealing with the so called *chaotic* behaviour of the solar system over billions years, or when trying to predict ice ages over a very large range of time. Indeed, the n-body problem is a conservative system (in fact a Lagrangian system) and many non-conservative effects, such as tidal forces between planets, will be of increasing importance along the computation. These non-conservative effects push the model outside the category of Lagrangian systems. You can go further by considering effects due to the changing in the oblateness of the sun. In this case, we do not even know how to model such kind of perturbations, and one is not sure of staying in the category of differential equations[(1)].

[(1)]Note that in the context of the solar system we have two different problems: first, if one uses only Newton's gravitational law, one must take into account the entire universe to model the behaviour of the planets. This by itself is a problem which can be studied by using the classical perturbation theory of ordinary differential equations. This is different if we want to speak of the "real" solar system for which we must consider effects that we ignore. In that case, even the validation of the law of gravitation as a real law of nature is not clear. I refer to [**16**] for more details on this point.

As a first step, this paper proposes tackling this problem by introducing a natural stochastic embedding procedure for ordinary or partial differential equations. This consists of looking for the behaviour of stochastic processes submitted to constraints induced by the underlying differential equation[(2)]. We point out that this strategy is different from the standard approach based on stochastic differential equations or stochastic dynamical systems, where one gives a meaning to ordinary differential equations perturbed by a small random term. In our work, no perturbations of the underlying equation are carried out.

A point of view that bears some resemblance to ours is contained in V.I. Arnold's *materialization of resonances* ([**6**],p.303-304), whose main underlying idea can be briefly explained as follows: the divergence of the Taylor expansion of the $\arctan x$ function at 0 for $\mid x \mid > 1$ can be proved by computing the coefficients of this series. However, this does not explain the reason for this divergence behaviour. One can obtain a better understanding by extending the function to the complex plane and by looking at its singularities at $\pm i$. The same idea can be applied in the context of dynamical systems. In this case, we look for the obstruction to linearization of a real systems in the complex plane. Arnold has conjectured that this is due to the accumulation of periodic orbits in the complex plane along the real axis. In our case, one can try to understand some properties of the trajectories of dynamical systems by using a suitable extension of its domain of definition. In our work, we give a precise sense to the concept of differential and partial differential equations in the class of stochastic processes. This procedure can be viewed as a first step toward the general "stochastic programme" as described by Mumford in [**51**].

Our embedding procedure is based on a simple idea: in order to write down differential or partial differential equations, one uses derivatives. An ordinary differential equation is nothing else but a differential operator of order one[(3)]. In order to embed ordinary differential equations, one must first extend the notion of derivative so that it makes sense in the context of stochastic processes. By extension, we mean that our stochastic derivative reduces to the classical derivative for deterministic differentiable processes. Having this extension, one easily defines in a unique way, the stochastic analogue of a differential operator, and as a consequence, a natural embedding of an ordinary differential

[(2)]This strategy is part of a general programme called the embedding procedure in [**15**] and which can be used to embed ordinary differential equations not only on stochastic processes but on general functional spaces. A previous attempt was made in [**13**],[**14**] in the context of the non-differentiable embedding of ordinary differential equations.
[(3)]In this case, we can also speak of vector fields.

equation on stochastic processes.

Of course, one can think that such a simple procedure will not produce anything new for the study of classical differential equations. This is not the case. The main problem that we study in this paper is the embedding of natural Lagrangian systems which are of particular interest for classical mechanics. In this context, we obtain some numerous surprising results, from the existence of a coherent least action principle with respect to the stochastic embedding procedure, to a derivation of a stochastic Noether theorem, and passing by a new derivation of the Schrödinger equation. All these points will be described with details in the following.

Two companion papers ([**18**],[**9**]) give an application of this method to derive new results on the formation of planets in a protoplanetary nebulae, in particular a proof of the existence of a so called Titus-Bode law for the spacing of planets around a given star.

The plane of the paper is as follow:

In a first part, we develop our notion of a stochastic derivative and study in details all its properties.

Chapter 1 gives a review of the stochastic calculus developed by Nelson [**53**]. In particular, we discuss the classical definition of the backward and forward Nelson derivatives, denoted by D and D_*, with respect to dynamical problems. We also define a class of stochastic process called good diffusion processes for which one can compute explicitly the Nelson derivatives.

In Chapter 2 we define what we call an abstract extension of the classical derivative. Using the Nelson derivatives, we define an extension of the ordinary derivative on stochastic processes, which we call the *stochastic derivative*. As pointed out previously, one imposes that the stochastic derivative reduces to the classical derivative on differentiable deterministic processes. This constraint ensures that the stochastic analogue of a PDE contains the classical PDE. Of course such a gluing constraint is not sufficient to define a rigid notion of stochastic derivative. We study several natural constraints which allow us to obtain a unique extension of the classical derivative on stochastic processes as

$$\mathcal{D}_\mu = \frac{D + D_*}{2} + i\mu \frac{D - D_*}{2}, \ \mu = \pm 1. \tag{0.1}$$

By extending this operator to complex valued stochastic processes, we are able to define the iterate of $\mathcal{D}$, i.e. $\mathcal{D}^2 = \mathcal{D} \circ \mathcal{D}$ and so on. The main surprise is that the real part of $\mathcal{D}^2$ correspond to the choice of Nelson for acceleration in his dynamical theory of Brownian motion. However, this result depends on the way we extend the stochastic derivative to complex valued stochastic processes. We discuss several alternative which covers well known variations on the Nelson acceleration.

In Chapter 3 we study the product rule satisfied by the stochastic derivative which is a fundamental ingredient of our stochastic calculus of variation. We also introduce an important class of stochastic processes, called Nelson differentiable, which have the property to have a real valued stochastic derivative. These processes play a fundamental role in the stochastic calculus of variation as they define the natural space of variations for stochastic processes.

The second part of this article deals specifically with the definition of a stochastic embedding procedure for ordinary differential equations.

Chapter 4 associate to a differential operator of a *given form* acting on sufficiently regular functions a unique operator acting on stochastic processes and defined simply by replacing the classical derivative by the stochastic derivative. This is this procedure that we call the stochastic embedding procedure. Note that the form of this procedure acts on differential operators of a given form. Although the procedure is canonical for a given form of operator, it is not canonical for a given operator.

The previous embedding is formal and does not take constraints which are of dynamical nature, like the reversibility of the underlying differential equation. As reversibility plays a central role in physics, especially in celestial mechanics which is one domain of application of our theory, we discuss this point in details. We introduce an embedding which respect the reversibility of the underlying equation. Doing this, we see that we must restrict attention to the real part of our operator, which is the unique one to possess this property in our setting. We then recover under dynamical and algebraic arguments studies dealing with particular choice of stochastic derivatives in order to derive quantum mechanics from classical mechanics under Nelson approach.

The third part is mainly concerned with the application of the stochastic embedding to Lagrangian systems.

We consider autonomous[4] Lagrangian systems $L(x,v)$, $(x,v) \in U \subset \mathbb{R}^d \times \mathbb{R}^d$, where U is an open set, which satisfy a number of conditions, one of it being that it must be holomorphic with respect to the second variable which represent the derivative of a given function. Such kind of Lagrangian functions are called admissible. Using the stochastic embedding procedure we can associate to the classical Euler-Lagrange equation a stochastic one which has the form

$$\frac{\partial L}{\partial x}(X(t), \mathcal{D}X(t)) = \mathcal{D}\left[\frac{\partial L}{\partial v}(X(t), \mathcal{D}X(t))\right], \qquad (SEL)$$

where X is a real valued stochastic process.

At this point, our manipulation is only formal and one can ask if this embedding is significant or not. We then remark that the Lagrangian function L keep sense on stochastic processes and can be considered as a functional. As a consequence, we can search for the existence of a least action principle which gives the stochastic Euler-Lagrange equation (SEL). The existence of such a stochastic least action principle is far from being trivial with respect to the embedding procedure. Indeed, it must follows from a stochastic calculus of variations which is not developed apart from this procedure. Our problem can then be formalize as the following diagram:

$$\begin{array}{ccc} L(x, dx/dt) & \xrightarrow{\text{LAP}} & EL \\ \big\downarrow \mathcal{S} & & \big\downarrow \mathcal{S} \\ L(X, \mathcal{D}X) & \xrightarrow{\text{SLAP }?} & (SEL), \end{array} \qquad (0.2)$$

where LAP is the least action principle, $\mathcal{S}$ is the stochastic embedding procedure, (EL) is the classical Euler-Lagrange equation associated to L and $SLAP$ the at this moment unknown stochastic least action principle. The existence of such a principle is called the coherence problem.

Chapter 7 develop a stochastic calculus of variations for functionals of the form

$$E\left[\int_a^b L(X(t), \mathcal{D}X(t))\, dt\right], \qquad (0.3)$$

where E denotes the classical expectation. Introducing the correct notion of extremals and variations we obtain two different stochastic analogue of the least action principle depending on the regularity class we choose for the admissible variations. The main point is that for variations in the class of Nelson differentiable process, the extremals of our functional coincide with the stochastic Euler-Lagrange equation obtained via the stochastic embedding procedure. This result is called the coherence lemma. In the

[4] This restriction is due to technical difficulties.

reversible case, i.e. taking as a stochastic derivative only the real part of our operator, we obtain the same result but in this case one can consider general variations.

In chapter 8 we provide a first study of what dynamical data remain from the classical dynamical system under the stochastic embedding procedure. We have focused on symmetries of the underlying equation and as a consequence on first integrals. We prove a stochastic analogue of the Noether theorem. This allows us to define a natural notion of first integral for stochastic differential equations. This part also put in evidence the need for a geometrical setting governing Lagrangian systems which is the analogue of symplectic manifolds.

Chapter 9 deals with the stochastic Euler-Lagrange equation for natural Lagrangian systems, i.e. associated to Lagrangian functions of the form

$$L(x,v) = T(v) - U(x), \tag{0.4}$$

where U is a smooth function and T is a quadratic form. In classical mechanics U is the potential energy and T the kinetic energy. The main result of this chapter is that by restricting our attention to good diffusion processes, and up to a a well chosen function ψ, called the wave function, the stochastic Euler-Lagrange equation is equivalent to a non linear Schrödinger equation. Moreover, by specializing the class of stochastic processes, we obtain the classical Schrödinger equation. In that case, we can give a very interesting characterization of stochastic processes which are solution of the stochastic Euler-Lagrange equation. Indeed, the square of the modulus of ψ is equal to the density of the associated stochastic process solution.

In chapter 10, we define a natural notion of stochastic Hamiltonian system. This result can be seen as a first attempt to put in evidence the stochastic analogue of a symplectic structure. We define a stochastic momentum process and prove that, up to a suitable modification of the stochastic embedding procedure called the Hamiltonian stochastic embedding, and reflecting the fact that the "speed" of a given stochastic process is complex, we obtain a coherent picture with the classical formalism of Hamiltonian systems. This first result is called the Legendre coherence lemma as it deals with the coherence between the Hamiltonian stochastic embedding procedure and the Legendre transform. Secondly, we develop a Hamilton least action principle and we prove again a coherence lemma, i.e.

that the following diagram commutes

$$\begin{array}{ccc} H(x(t),p(t)) & \xrightarrow{S_H} & H(X(t),P(t)) \\ \text{Hamilton least action principle}\downarrow & & \downarrow\text{Stochastic Hamilton least action principle} \\ (HE) & \xrightarrow[S_H]{} & (SHE) \end{array}$$

where S_H denotes the Hamiltonian stochastic embedding procedure.

The last chapter discuss many possible developments of our theory from the point of view of mathematics and applications.

PART I

THE STOCHASTIC DERIVATIVE

CHAPTER 1

ABOUT NELSON STOCHASTIC CALCULUS

1.1. About measurement and experiments

In this section, we explain what we think are the basis of all possible extensions of the classical derivative. The setting of our discussion is the following:

We consider an experimental set-up which produces a dynamics. We assume that each dynamics is observed during a time which is fixed, for example $[0, T]$, where $T \in \mathbb{R}^{*+}$. For each experiment i, $i \in \mathbb{N}$, we denote by $X_i(t)$ the dynamical variable which is observed for $t \in [0, T]$.

Assume that we want to describe the *kinematic* of such a dynamical variable. What is the strategy ?

The usual idea is to *model* the dynamical behaviour of a variable by ordinary differential equations or partial differential equations. In order to do this, we must first try to have access to the *speed* of the variable. In order to compute a significant quantity we can follow at least two different strategies:

– We do not have access to the variable $X_i(t)$, $t \in [0, T]$, but to a collection of measurements of this dynamical variable. Assume that we want to compute the speed at time t. We can only compute an approximation of it for a given resolution h greater than a given threshold h_0. Assume that for each experiment we are able to compute the quantity

$$v_{i,h}(t) = \frac{X_i(t+h) - X_i(t)}{h}. \tag{1.1}$$

We can then try to look for the behaviour of this quantity when h varies. If the underlying dynamics is not too irregular, then we can expect a limit for $v_{i,h}(t)$ when

h goes to zero that we denote by $v_i(t)$.

We then compute the mean value

$$\bar{v}(t) = \frac{1}{n}\sum_{i=1}^{n} v_i(t). \tag{1.2}$$

If the underlying dynamics is not too irregular then $\bar{v}(t)$ can be used to model the problem. In the contrary the basic idea is to introduce a random variable.

Remark that due to the intrinsic limitation for h we never have access to $v_i(t)$ so that this procedure can not be implemented.

– Another idea is to look directly for the quantity

$$\bar{v}_{h,n}(t) = \frac{1}{n}\sum_{i=1}^{n} v_{i,h}(t). \tag{1.3}$$

Contrary to the previous case, if there exists a well defined mean value $\bar{v}_h(t)$ when n goes to infinity then we can have a as close as we want approximation. Indeed it suffices to do sufficiently many experiences. We then look for the limit of $\bar{v}_h(t)$ when h goes to zero.

For regular dynamics these two procedures lead to the same result as all these quantities are well defined and converge to the same quantity. This is not the case when we deal with highly irregular dynamics. In that case the second procedure is easily implemented contrary to the first one. The only problem is that we loose the geometrical meaning of the resulting limit quantity with respect to individual trajectories as one directly take a mean on all trajectories before taking the limit in h.

This second alternative can be formalized using stochastic processes and leads to the Nelson backward and forward derivatives that we define in the next section.

We have take the opportunity to discuss these notions because the previous remarks proves that one can not justify the form of the Nelson derivatives using a geometrical argument like the non differentiability of trajectories for a Brownian motion. This is however the argument used by E. Nelson ([**54**],p.1080) in order to justify the fact that we need a substitute for the classical derivative when studying Wiener processes. This misleadingly suggest that the forward and backward derivative capture this non differentiability in their definition, which is not the case.

1.2. The Nelson derivatives

Let $X(t)$, $0 \leqslant t \leqslant 1$ be d-dimensional continuous random process defined on a probability space $(\Omega, \mathcal{A}, P)$, where $\mathcal{A}$ is the σ-algebra of all measurable events and P is a probability measure defined on $\mathcal{A}$. We denote by I the open interval $(0, 1)$.

Definition 1.1. — *The random process $X(t)$, $a \leqslant t \leqslant b$, is an SO-process if each $X(t)$ belongs to $L^1(\Omega)$ and the mapping $t \to X(t)$ from $\mathbb{R}$ to $L^1(\Omega)$ is continuous.*

Let $\mathcal{P} = \{\mathcal{P}_t\}$ and $\mathcal{F} = \{\mathcal{F}_t\}$ be an increasing and a decreasing family of sub-σ-algebras, respectively, such that $X(t)$ is $\mathcal{F}_t$-measurable and $\mathcal{P}_t$-measurable. In other words, $\mathcal{F}$ and $\mathcal{P}$ are two filtration to which $X(t)$ is adapted. We let $E[\bullet \mid \mathcal{B}]$ denote the conditional expectation with respect to any sub-σ-algebra $\mathcal{B} \subset \mathcal{A}$.

Definition 1.2. — *The random process $X(t)$, $a \leqslant t \leqslant b$, is an S1-process if it is an SO-process such that*

$$DX(t) = \lim_{h \to 0^+} E\left[\frac{X(t+h) - X(t)}{h} \mid \mathcal{P}_t\right], \tag{1.4}$$

and

$$D_*X(t) = \lim_{h \to 0^+} E\left[\frac{X(t) - X(t-h)}{h} \mid \mathcal{F}_t\right], \tag{1.5}$$

exist in $L^1(\Omega)$ and the mappings $t \mapsto DX(t)$ and $t \mapsto D_*X(t)$ are both continuous from $\mathbb{R}$ to $L^1(\Omega)$.

Definition 1.3. — *The random process $X(t)$, $a \leqslant t \leqslant b$, is an S2-process if it is an S1-process, and*

$$\sigma^2 X(t) = \lim_{h \to 0^+} E\left[\frac{(X(t+h) - X(t))^2}{h} \mid \mathcal{P}_t\right], \tag{1.6}$$

and

$$\sigma_*^2 X(t) = \lim_{h \to 0^+} E\left[\frac{(X(t+h) - X(t))^2}{h} \mid \mathcal{F}_t\right], \tag{1.7}$$

exist in $L^1(\Omega)$.

Definition 1.4. — *We denote by $\mathcal{C}^1(I)$ the totality of S2-processes with continuous sample paths, such that $X(t)$, $DX(t)$ and $D_*X(t)$, $a \leqslant t \leqslant b$, all lie in the Hilbert space $L^2(\Omega)$ and are continuous functions of t in $L^2(\Omega)$.*

A completion of $\mathcal{C}^1(I)$ in the norm

$$\| X \| = \sup_{t \in I} (\| X(t) \|_{L^2(\Omega)} + \| DX(t) \|_{L^2(\Omega)} + \| D_*X(t) \|_{L^2(\Omega)}), \tag{1.8}$$

is also denoted by $\mathcal{C}^1(I)$, where $\| \cdot \|_{L^2(\Omega)}$ denotes the norm of Hilbert space $L^2(\Omega)$.

Remark 1.1. — *The main point in the previous definitions for a forward and backward derivative of a stochastic process, is that the forward and backward filtration are fixed by the problem. As a consequence, we have not an intrinsic quantity only related to the stochastic process. A possible alternative definition is the following:*

Definition 1.5. — *Let X be a stochastic process, and $\sigma(X)$ (resp. $\sigma_*(X)$) the forward (resp. backward) adapted filtration. We define*

$$\mathfrak{d}X(t) = \lim_{h\to 0^+} h^{-1}E[X(t+h) - X(t) \mid \sigma(X_s, 0 \leqslant s \leqslant t)], \tag{1.9}$$

$$\mathfrak{d}_*X(t) = \lim_{h\to 0^+} h^{-1}E[X(t) - X(t-h) \mid \sigma(X_s, t \leqslant s \leqslant 1)]. \tag{1.10}$$

In this case, we obtain intrinsic quantities, only related to the stochastic process. However, these new operators behave very badly from an algebraic view point. Indeed, without stringent assumptions on stochastic processes, we do not have linearity of $\mathfrak{d}$ or $\mathfrak{d}_*$.

This difficulty is not apparent as long as one restrict attention to a single stochastic process.

1.3. Good diffusion processes

We introduce a special class of diffusion processes for which we can explicitly compute the derivative D, D_*, DD_*, D_*D, D^2 and D_*^2.

Definition 1.6. — *We denote by Λ_d the space of diffusion processes X satisfying the following conditions:*

i- X solves a stochastic differential equation :

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \quad X(0) = X_0, \tag{1.11}$$

where $X_0 \in L^2(\Omega)$, $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d\otimes\mathbb{R}^d$ are Borel measurable functions satisfying the hypothesis : there exists a constant K such that for every $x, y \in \mathbb{R}^d$ we have

$$\sup_t \left(|\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)|\right) \leqslant K\,|x-y|\,, \tag{1.12}$$

$$\sup_t \left(|\sigma(t,x)| + |b(t,x)|\right) \leqslant K(1+|x|). \tag{1.13}$$

ii- For any $t > 0$, $X(t)$ has a density $p_t(x)$ at point x.

iii- Setting $a_{ij} = (\sigma\sigma^*)_{ij}$*, for any* $i \in \{1, \cdots, n\}$*, for any* $t_0 > 0$*, for any bounded open set* $D \subset \mathbb{R}^d$*,*

$$\int_{t_0}^{1} \int_D |\partial_j(a_{ij}(t,x)p_t(x))|\, dxdt < +\infty. \tag{1.14}$$

iv- b *and* $(t,x) \to \dfrac{1}{p_t(x)}\partial_j(a_{ij}(t,x)p_t(x))$ *are continuous and bounded functions.*

Remark 1.2. —
- *Hypothesis iii) ensures that (1.11) has a unique* $t-$*continuous solution* $X(t)$*.*
- *Hypothesis i), ii) and iii) allow to apply theorem* 2.3 *p.*217 *in* [**49**]*.*
- *We may wonder in which cases hypothesis ii) holds. Theorem* 2.3.2 *p.*111 *of* [**58**] *gives the existence of a density for all* $t > 0$ *under the Hörmander hypothesis which is involved by the stronger condition that the matrix diffusion* $\sigma\sigma^*$ *is elliptic at any point* x*. A simple example is given by a SDE where* b *is a* $\mathcal{C}^\infty(I \times \mathbb{R}^d)$ *function with all its derivatives bounded, and where the diffusion matrix is a constant equal to* cId*. In this case,* $p_t(x)$ *belongs to* $\mathcal{C}^\infty(I \times \mathbb{R}^d)$*; moreover, if* X_0 *has a differentiable and everywhere positive density* $p_0(x)$ *with respect to Lebesgue measure such that* $p_0(x)$ *and* $p_0(x)^{-1}\nabla p_0(x)$ *are bounded, then* $b(t,x) - c\nabla log(p_t(x))$ *is bounded as noticed in the proof of proposition* 4.1 *in* [**64**]*. So hypothesis ii) seems not to be such a restrictive condition.*
- *Assumption iv) is necessary to compute explicitly the second order operators of* D *and* D_**. The existence of* D *and* D_* *is ensured under a weaker condition, the finite entropy condition equivalent to*

$$E\left[\int_0^1 (b(t,X(t))^2\, dt\right] < \infty. \tag{1.15}$$

We refer to Föllmer ([**25**]*,proposition 2.5 p.121 and lemma 3.1 p.123) for more details.*

According to the theorem 2.3 of [**49**] and thanks to iv), we will see that $\Lambda_d \subset \mathcal{C}^1([0,T])$ and that we can compute DX and D_*X for $X \in \Lambda_d$ (see Theorem 1.1).

1.4. The Nelson derivatives for good diffusion processes

A useful property of good diffusions processes is that their Nelson's derivatives can be explicitly computed. Precisely, we have:

Theorem 1.1. — *Let $X \in \Lambda_d$ which writes $dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t)$. Then X is Markov diffusion with respect to an increasing filtration $(\mathcal{P}_t)$ and a decreasing filtration $(\mathcal{F}_t)$. Moreover, DX and D_*X exists w.r.t. these filtration and :*

$$DX(t) = b(t, X(t)) \tag{1.16}$$

$$D_*X(t) = b_*(t, X(t)) \tag{1.17}$$

where $x \to p_t(x)$ denotes the density of $X(t)$ at x and

$$b^i_*(t,x) = b^i(t,x) - \frac{1}{p_t(x)}\partial_j(a^{ij}(t,x)p_t(x))$$

with the convention that the term involving $\frac{1}{p_t(x)}$ is 0 if $p_t(x) = 0$.

Proof. — The proof uses essentially theorem 2.3 of Millet-Nualart-Sanz [**49**] and the techniques of M. Thieullen for the proof of proposition 4.1 in [**64**].

(1) Let $X \in \Lambda_d$. Then X is a Markov diffusion w.r.t. the increasing filtration $(\mathcal{P}_t)$ generated by the Brownian Motion $W(t)$ and so :

$$E\left[\frac{X(t+h) - X(t)}{h}\,|\mathcal{P}_t\right] = E\left[\frac{1}{h}\int_t^{t+h} b(s, X(s))ds\,|\mathcal{P}_t\right],$$

and

$$E\left[\left|E\left[\frac{X(t+h) - X(t)}{h}\,|\mathcal{P}_t\right] - b(t, X(t))\right|\right] \leqslant E\left[\frac{1}{h}\int_t^{t+h} |b(s, X(s)) - b(t, X(t))|\,ds\right].$$

We can apply the dominated convergence theorem since b is bounded and

$$\frac{1}{h}\int_t^{t+h} |b(s, X(s)) - b(t, X(t))|\,ds \overset{h\to 0}{\longrightarrow} 0 \ a.s.$$

(for b is continuous and X has a.s. continuous paths).

Therefore DX exists and $DX(t) = b(t, X(t))$.

(2) As $X \in \Lambda_d$, we can apply theorem 2.3 in [**49**]. So $\overline{X}(t) = X(1-t)$ is a diffusion process w.r.t. an increasing filtration $(\overline{\mathcal{P}}_t)$ and whose generator reads $\overline{L}_t f = \overline{b}^i\partial_i f + \frac{1}{2}\overline{a}^{ij}\partial_{ij} f$ with $\overline{a}^{ij}(1-t,x) = a^{ij}(t,x)$ and $\overline{b}^i(1-t,x) = -b^i(t,x) + \dfrac{1}{p_t(x)}\partial_j(a^{ij}(t,x)p_t(x))$. Setting $\mathcal{F}_t = \overline{\mathcal{P}}_{1-t}$, X is a Markov diffusion w.r.t. the decreasing filtration $(\mathcal{F}_t)$. We have :

$$\begin{aligned} E\left[\frac{X(t) - X(t-h)}{h}\,|\mathcal{F}_t\right] &= E\left[\frac{\overline{X}(1-t) - \overline{X}(1-t+h)}{h}\,|\overline{\mathcal{P}}_{1-t}\right] \\ &= -E\left[\frac{1}{h}\int_{1-t}^{1-t+h} \overline{b}(s, \overline{X}(s))ds\,|\overline{\mathcal{P}}_{1-t}\right]. \end{aligned} \tag{1.18}$$

Using the same calculations and arguments as above (since hypothesis iv) in the definition of class Λ_d implies that $\overline{b}$ is continuous and bounded), we obtain that $D_*X(t)$ exists and is equal to $-\overline{b}(1-t, \overline{X}(1-t))$. □

In the case of *fractional Brownian motion* of order $H \neq 1/2$, the Nelson derivatives do not exist. However, one can define new operators using the so-called quasi conditional expectation introduced by [**1**]. We refer to the work of Darses and Sausserau [**19**] for more details.

1.5. A remark about reversed processes

This part reviews basic results about reversed processes, with a special emphasis to diffusion processes. We use Nelson's stochastic calculus.

Let X be a process in the class $\mathcal{C}^1([0,1])$. We denote by $\widetilde{X}$ the reversed process : $\widetilde{X}(t) = X(1-t)$, with his "past" $\widetilde{\mathcal{P}}_t$ and his "future" $\widetilde{\mathcal{F}}_t$. As a consequence, we also have $\widetilde{x} \in \mathcal{C}^1([0,1] \to H)$.

Using the operators $\mathfrak{d}$ and $\mathfrak{d}_*$ defined in definition 1.5,we have:

Lemma 1.1. — $\mathfrak{d}_* x(t) = -\mathfrak{d}\widetilde{x}(1-t) = -\widetilde{\mathfrak{d}\widetilde{x}}(t)$.

Proof. — The definition of $\mathfrak{d}_*$ gives immediately:

$$\mathfrak{d}_* x(t) = \lim_{\epsilon \to 0^+} E\left[\frac{\widetilde{x}(1-t) - \widetilde{x}(1-t+\epsilon)}{\epsilon} \middle| \mathcal{F}_t \right].$$

But $\mathcal{F}_t = \sigma\{x(s), t \leqslant s \leqslant 1\} = \sigma\{\widetilde{x}(u), 0 \leqslant u \leqslant 1-t\} = \widetilde{\mathcal{P}}_{1-t}$.
Thus:

$$\mathfrak{d}_* x(t) = \lim_{\epsilon \to 0^+} -E\left[\frac{\widetilde{x}(1-t+\epsilon) - \widetilde{x}(1-t)}{\epsilon} \middle| \widetilde{\mathcal{P}}_{1-t} \right] = -\mathfrak{d}\widetilde{x}(1-t) = -\widetilde{\mathfrak{d}\widetilde{x}}(t).$$

□

The same computation is not at all possible when dealing with the operators D and D_*.

CHAPTER 2

STOCHASTIC DERIVATIVE

In this part, we construct a natural *extension*[1] of the classical derivative on real stochastic processes as a unique solution to an algebraic problem. This stochastic derivative turns out to be necessarily complex valued. Our construction relies on Nelson's stochastic calculus [**53**]. We then study properties of our stochastic derivative and establish a number of technical results, including a generalization of Nelson's product rule [**53**] as well as the stochastic derivative for functions of diffusion processes . We also compute the stochastic derivative in some classical examples. The main point is that, after a natural extension to complex processes, the real part of the second derivative of a real stochastic process coincide with Nelson's mean acceleration. We define a special class of processes called Nelson differentiable, which will be of importance for the stochastic calculus of variations developed in chapter 7. This part is self contained and all basic results about Nelson's stochastic calculus are reminded.

2.1. The abstract extension problem

In this section, we discuss in a general abstract setting, what kind of analogue of the classical derivative we are waiting for on stochastic processes.

We first remark that real[2] valued functions naturally embed in stochastic processes.

Indeed, let $f : \mathbb{R} \to \mathbb{R}$ be a given function. We denote by X_f the deterministic stochastic process defined by

$$X_f(\omega) = f \ \forall \omega \in \Omega. \tag{2.1}$$

[1] A precise meaning to this word will be given in the following. It should be noted that Malliavin calculus is not an extension of the ordinary differential calculus (see below).

[2] Our aim was first to study dynamical systems over $\mathbb{R}^n$. However, as we will see we will need to consider complex valued objects.

We denote by $\iota : \mathbb{R}^{\mathbb{R}} \to \mathcal{P}$ the map associating to $f \in \mathbb{R}^{\mathbb{R}}$ the stochastic process X_f.

We denote by $\mathcal{P}_{\text{det}}$ the subset of $\mathcal{P}$ consisting of deterministic processes, and by $\mathcal{P}^k_{\text{det}}$ the set $\iota(C^k)$, $k \geqslant 1$.

As a consequence, we have a natural action of the classical derivative on the set of differentiable deterministic processes, that we denote again d/dt.

Let $K = \mathbb{R}$ or $\mathbb{C}$. In the sequel, we denote by $\mathcal{P}_K \subset \mathcal{S}_K$ a subset of the set of K-valued stochastic processes[(3)].

Let $K = \mathbb{R}$ or $\mathbb{C}$.

Definition 2.1. — *Let $K = \mathbb{R}$ or $\mathbb{C}$. An extension of d/dt on $\mathcal{P}_K$ is an operator δ, i.e. a map $\delta : \mathcal{P}_K \to \mathcal{S}_K$ such that:*

i) *δ coincides with d/dt on $\mathcal{P}^1_{\text{det}}$,*
ii) *δ is $\mathbb{R}$-linear.*

Condition i), which is a gluing condition on the classical derivative is necessary as long as one wants to relate classical differential equations with their stochastic counterpart.

Condition ii) is more delicate. Of course, one has linearity of δ on Diff. A natural idea is then to preserve fundamental algebraic properties of d/dt, $\mathbb{R}$-linearity being one of them. This condition is not so stringent, if for example we consider $K = \mathbb{C}$. But, following this point of view, one can ask for more precise properties like the Leibniz rule

$$d/dt(X \cdot Y) = d/dt(X) \cdot Y + X \cdot d/dt(Y), \quad \forall X, Y \in \mathcal{P}^1_{\text{det}}. \tag{2.2}$$

In what follows, we construct a stochastic differential calculus based on Nelson's derivatives.

2.2. Stochastic differential calculus

In this part, we extend the classical differential calculus to stochastic processes using a previous work of Nelson [**53**] on the dynamical theory of Brownian motion. We define a stochastic derivative and review its properties.

(3) We do not give more precisions on this set for the moment, the set $\mathcal{P}$ can be the whole set of real or complex valued stochastic processes, or a particular class like diffusion processes,...etc.

2.2.1. Reconstruction problem and extension. — Let us begin with some heuristic remarks supporting our definition and construction of a stochastic derivative.

Our aim is to construct a "natural" operator on $\mathcal{C}^1(I)$ which reduces to the classical derivative d/dt over differentiable deterministic processes[(4)]. The basic idea underlying the whole construction is that, for example in the case of the Brownian motion, the trajectories are non-differentiable. At least, this is the reason why Nelson [**53**] introduces the left and right derivatives DX and D_*X for a given process X. If we refer to geometry, forgetting for a moment processes for trajectories, the fundamental property of the classical derivative $dx/dt(t_0)$ of a trajectory $x(t)$ at point t_0, is to provide a first order (geometric) approximation of the curve in a *neighbourhood* of t_0. One wants to construct an operator, that we denote by $\mathcal{D}$, such that the data of $\mathcal{D}X(t_0)$ allows us to give an approximation of X in a neighbourhood of t_0. The difference is that we must know two quantities, namely DX and D_*X, in order to obtain the information[(5)]. For computational reasons, one wants an operator with values in a field F. This field must be a natural extension of $\mathbb{R}$ (as we want to recover the classical derivative) and at least of dimension 2. The natural candidate to such a field is $\mathbb{C}$. One can also recover $\mathbb{C}$ by saying that we must consider not only $\mathbb{R}$ but the doubling algebra which corresponds to $\mathbb{C}$.

This informal discussion leads us to build a complex valued operator $\mathcal{D} : \mathcal{C}^1(I) \to \mathcal{C}^1_{\mathbb{C}}(I)$, with the following constraints:

i) (*Gluing property*) For $X \in \mathcal{P}^1_{\text{det}}$, $\mathcal{D}X(t) = dX/dt$,

ii) The operator $\mathcal{D}$ is $\mathbb{R}$-linear,

iii) (*Reconstruction property*) For $X \in \mathcal{C}^1(I)$, let us denote by

$$\mathcal{D}X = A(DX, D_*X) + iB(DX, D_*X),$$

where A and B are linear $\mathbb{R}$-valued mappings by ii). We assume that the mapping

$$(DX, D_*X) \mapsto (A(DX, D_*X), B(DX, D_*X))$$

is invertible.

(4) A rigourous meaning to this sentence will be given in the sequel.
(5) This remark is only valid for general stochastic processes. Indeed, as we will see, for diffusion processes, there is a close connection between DX and D_*X, which allows to simplify the definition of $\mathcal{D}$.

Lemma 2.1. — *The operator $\mathcal{D}$ has the form*

$$\mathcal{D}_\mu X = [aDX + (1-a)D_*X] + i\mu b\,[DX - D_*X]\,,\ \mu = \pm 1,$$

where $a, b \in \mathbb{R}$ and $b \neq 0$.

Proof. — We denote by $A(X) = aDX + bD_*X$ and $B(X) = cDX + dD_*X$. If $X \in C^1(I)$, we have $DX = D_*X = dX/dt$, and i) implies

$$a + b = 1,\ c + d = 0.$$

We then obtain the desired form. By iii), we must have $b \neq 0$ in order to have invertibility. □

In order to rigidify this operator, we impose a constraint coming from the analogy with the construction of the scale-derivative for non-differentiable functions in [**13**].

iv) If $D_* = -D$, then $A(X) = 0$, $B(X) = D$.

We then obtain the following result:

Lemma 2.2. — *An operator $\mathcal{D}$ satisfying conditions i), ii), iii) and iv) is of the form*

$$\mathcal{D}_\mu = \frac{D + D_*}{2} + i\mu\frac{D - D_*}{2},\ \mu = \pm 1. \tag{2.3}$$

Proof. — Using lemma 2.1, iii) implies the relations: $2a - 1 = 0$ and $2b = 1$, so $a = b = 1/2$. □

We then introduce the following notion of *stochastic derivative*:

Definition 2.2. — *We denote by $\mathcal{D}_\mu$ the operators defined by*

$$\mathcal{D}_\mu = \frac{D + D_*}{2} + i\mu\frac{D - D_*}{2},\ \mu = \pm 1.$$

2.2.2. Extension to complex processes. — In order to embed second order differential equations, we need to define the meaning of $\mathcal{D}^2$, and more generally of $\mathcal{D}^n$, $n \in \mathbb{N}$. The basic problem is that, contrary to what happens for the ordinary differential operator d/dt, even if we consider real valued processes X, the derivative $\mathcal{D}X$ is a complex one. As a consequence, one must extend $\mathcal{D}$ to *complex processes*.

For the moment, let us denoted by $\mathcal{D}_{\mathbb{C}}$ the extension to be define of $\mathcal{D}$, to complex processes. Let F be a field containing $\mathbb{C}$ to be defined, and $\mathcal{D}_{\mathbb{C}} : \mathcal{C}^1_{\mathbb{C}}(I) \to F$. There are essentially two possibilities to extend the stochastic derivative leading to the same definition: an algebraic and an analytic one.

2.2.2.1. Algebraic extension. — Let us assume that:

i) the operator $\mathcal{D}_{\mathbb{C}}$ is $\mathbb{R}$-linear.

Let $Z = X + iY$ be a complex process, where X and Y are two real processes. By $\mathbb{R}$-linearity, we have

$$\mathcal{D}_{\mathbb{C}}(Z) = \mathcal{D}_{\mathbb{C}}X + \mathcal{D}_{\mathbb{C}}(iY).$$

As $\mathcal{D}_{\mathbb{C}}$ reduce to $\mathcal{D}$ on real processes, we obtain

$$\mathcal{D}_{\mathbb{C}}(Z) = \mathcal{D}X + \mathcal{D}_{\mathbb{C}}(iY),$$

which reduce the problem of the extension to find a suitable definition of $\mathcal{D}$ on *purely imaginary processes*.

We now make an assumption about the image of $\mathcal{D}_{\mathbb{C}}$:

ii) The operator $\mathcal{D}_{\mathbb{C}}$ is $\mathbb{C}$-valued.

This assumption is far from being trivial, and has many consequences. One of them is that, whatever the definition of $\mathcal{D}_{\mathbb{C}}(iY)$ is, we will obtain a complex quantity which *mixes* with the quantity $\mathcal{D}X$ in a non trivial way.

Remark 2.1. — *One can wonder if another choice is possible, as for example, using quaternions in order to avoid this mixing problem. However, a heuristic idea behind the complex nature of $\mathcal{D}$ is that it corresponds to a fundamental property of Nelson processes, the (in general) non-differentiable character of trajectories. Then, the doubling of the underlying algebra is related to a symmetry breaking*[(6)]*. The computation of $\mathcal{D}^2$ is not related to such phenomenon.*

In the following, we give two different extensions of $\mathcal{D}$ to complex processes under hypothesis i) and ii). The basic problem is the following:

Let Y be a real process. We denote

$$\mathcal{D}Y = S(Y) \pm iA(Y), \tag{2.4}$$

where

$$S(Y) = \left[\frac{D + D_*}{2}\right](Y), \text{ and } A(Y) = \left[\frac{D - D_*}{2}\right](Y), \tag{2.5}$$

[(6)]This reduces to $DX = D_*X$ for deterministic differentiable processes, namely the invariance under $h \to -h$.

and the letters S and A stand for the symmetric and antisymmetric operators with respect to the exchange of D with D_*.

We denote

$$\mathcal{D}_{\mathbb{C}}(iY) = R(Y) + iI(Y), \tag{2.6}$$

where $R(Y)$ and $I(Y)$ are two real processes.

One can ask if we expect for special relations between $R(Y)$, $I(Y)$ and $S(Y)$, $A(Y)$.

2.2.2.1.1. $\mathbb{C}$-linearity. — If no relations are expected for, the natural hypothesis is to assume $\mathbb{C}$-linearity of $\mathcal{D}_{\mathbb{C}}$, i.e.

$$\mathcal{D}_{\mathbb{C}}(iY) = i\mathcal{D}Y. \tag{2.7}$$

As a consequence, we obtain the following definition for the operator $\mathcal{D}_{\mathbb{C}}$:

We denote by $\mathcal{C}^1_{\mathbb{C}}(I)$ the set of stochastic processes of the form $Z = X + iY$, with $X, Y \in \mathcal{C}^1(I)$.

Definition 2.3. — *The operator $\mathcal{D}_{\mathbb{C}} : \mathcal{C}^1_{\mathbb{C}} \to \mathcal{C}^1_{\mathbb{C}}$ is defined by*

$$\mathcal{D}_{\mathbb{C},\mu}(X + iY) = \mathcal{D}_\mu X + i\mu\mathcal{D}_\mu Y,\ \mu = \pm 1,$$

where $X, Y \in \mathcal{C}^1$.

In the sequel, we denote $\mathcal{D}_{\mathbb{C}}$ for $\mathcal{D}_{\mathbb{C},\sigma}$.

The following lemma gives a strong reason to choose such a definition of $\mathcal{D}_{\mathbb{C}}$. We denote by

$$\mathcal{D}^n_{\mathbb{C}} = \mathcal{D}_{\mathbb{C}} \circ \cdots \circ \mathcal{D}_{\mathbb{C}}.$$

Lemma 2.3. — *We have*

$$\mathcal{D}^2_{\mathbb{C}} = \left[\frac{DD_* + D_*D}{2}\right] + i\left[\frac{D^2 - D_*^2}{2}\right]. \tag{2.8}$$

Proof. — One use the $\mathbb{C}$-linearity of operator $\mathcal{D}$. □

We note that the real part of $\mathcal{D}^2$ is the *mean acceleration* as defined by Nelson [**53**].

Remark 2.2. — *In ([**53**],p.81-82), Nelson discusses natural candidates for the stochastic analogue of acceleration. More or less, the idea is to consider quadratic combinations of D and D^*, respecting a gluing property with the classical derivative:*

Let $Q_{a,b,c,d}(x,y) = ax^2 + bxy + cyx + dy^2$ be a real non-commutative quadratic form such that $a+b+c+d = 1$. A possible definition for a stochastic acceleration is $Q(D, D^*)$.

We remark that the condition $a + b + c + d = 1$ implies that when $D = D^*$, we have $Q(D, D^*) = D = D_*$.

The simplest examples of this kind are: D^2, D_*^2, DD_* and D_*D.

We can also impose a symmetry condition in order to take into account that we do not want to give a special importance to the mean-forward or mean-backward derivative, by assuming that $Q(x,y) = Q(y,x)$, so that Q is of the form

$$Q_a(x,y) = a(x^2+y^2) + (1-2a)\frac{xy+yx}{2}, a \in \mathbb{R}.$$

The simplest example in this case is obtained by taking $a = 0$, i.e.

$$Q_0(D, D_*) = \frac{DD_* + D_*D}{2}.$$

This last one corresponds to Nelson's mean acceleration and coincide with the real part of our stochastic derivative.

It must be pointed out that Nelson discuss only five possible candidates where at least a three parameters family can be defined by $Q_{a,b,c,1-a-b-c}(D, D_*)$. His five candidates correspond to the simplest cases we have described.

The choice of $Q_0(D, D_*)$ as a mean acceleration is justified by Nelson using a Gaussian Markov process $X(t)$ in equilibrium, satisfying the stochastic differential equation

$$dX(t) = -\omega X(t)dt + dW(t).$$

We will return to this problem below.

2.2.2.2. Analytic extension. — We first remark that D and D_* possess a natural extension to complex processes. Indeed, let $X = X_1 + iX_2$, with $X_i \in \mathcal{C}^1(I)$ then

$$D(X_1 + iX_2) = D(X_1) + iD(X_2) \text{ and } D_*(X_1 + iX_2) = D_*(X_1) + iD_*(X_2).$$

As a consequence, the quantities $S(Y)$ and $A(Y)$ introduced in the previous section for real valued processes make sense for complex processes, and the quantity $A(X) + iS(X)$ is well defined for the complex process $X \in \mathcal{C}^1_{\mathbb{C}}(I)$. As a consequence, we can naturally extend $\mathcal{D}(X)$ to complex processes by simply posing

$$\mathcal{D}(X) = \frac{D + D_*}{2} + \mu i \frac{D - D_*}{2},$$

with the natural extension of D and D_*.

2.2.2.3. Symmetry. — A possible way to extend $\mathcal{D}$ is to assume that the regular part of $\mathcal{D}_{\mathbb{C}}(iY)$ is equal the imaginary part of $\mathcal{D}(Y)$, i.e. that the geometric meaning of the complex and real part of $\mathcal{D}Y$ is exchanged. We then impose the following relation:

$$R(Y) = \sigma A(Y).$$

This leads to the following extension:

Definition 2.4. — *The operator $\mathcal{D}_{\mathbb{C}} : \mathcal{C}^1_{\mathbb{C}} \to \mathcal{C}^1_{\mathbb{C}}$ is defined by*

$$\mathcal{D}_{\mathbb{C},\mu}(X + iY) = \mathcal{D}_\mu X - i\mu \mathcal{D}_\mu Y,\ \mu = \pm 1,$$

where $X, Y \in \mathcal{C}^1$.

2.2.3. Stochastic derivative for functions of diffusion process. — In the following, we need to compute the stochastic derivative of $f(t, X_t)$ where X_t is a diffusion process and f is a smooth function. Our main result is the following lemma:

Lemma 2.4. — *Let $X \in \Lambda_d$ and $f \in C^{1,2}(I \times \mathbb{R}^d)$ such that $\partial_t f$, ∇f and $\partial_{ij} f$ are bounded. Then, we have:*

$$Df(t, X(t)) = \left[\partial_t f + DX(t) \cdot \nabla f + \frac{1}{2} a^{ij} \partial_{ij} f\right](t, X(t)), \tag{2.9}$$

$$D_* f(t, X(t)) = \left[\partial_t f + D_* X(t) \cdot \nabla f - \frac{1}{2} a^{ij} \partial_{ij} f\right](t, X(t)). \tag{2.10}$$

Proof. — Let $X \in \Lambda_d$ and $f \in C^{1,2}(I \times \mathbb{R}^d)$ such that $\partial_t f$, ∇f and $\partial_{ij} f$ are bounded. Thus f belongs to the domain of the generators L_t and $\overline{L}_t$ of the diffusions $X(t)$ and $\overline{X}(t)$. Moreover these regularity assumptions allow us to use the same arguments as in the proof of theorem 1.1 in order to write :

$$\begin{aligned} Df(t, X(t)) &= \partial_t f(t, X(t)) + L_t(f(t, \cdot))(X(t)) \\ &= \left[\partial_t f + b^i \partial_i f + \frac{1}{2} a^{ij} \partial_{ij} f\right](t, X(t)) \\ &= \left[\partial_t f + DX(t) \cdot \nabla f + \frac{1}{2} a^{ij} \partial_{ij} f\right](t, X(t)) \end{aligned}$$

and

$$\begin{aligned} D_* f(t, X(t)) &= \partial_t f(t, X(t)) - \overline{L}_{1-t}(f(t, \cdot))(X(t)) \\ &= \left[\partial_t f + D_* X(t) \cdot \nabla f - \frac{1}{2} a^{ij} \partial_{ij} f\right](t, X(t)) \end{aligned}$$

□

We deduce immediately the following corollary :

Corollary 2.1. — *Let $X \in \Lambda_d$ and $f \in C^{1,2}(I \times \mathbb{R}^d)$ such that $\partial_t f$, ∇f and $\partial_{ij} f$ are bounded. Then, we have:*

$$\mathcal{D}_\mu f(t, X(t)) = \left[\partial_t f + \mathcal{D}_\mu X(t) \cdot \nabla f + \frac{i\mu}{2} a^{ij} \partial_{ij} f\right](t, X(t)). \tag{2.11}$$

and

Corollary 2.2. — *Let $X \in \Lambda_d$ with a constant diffusion coefficient σ and $f \in C^{1,2}(I \times \mathbb{R}^d)$ such that $\partial_t f$, ∇f and $\partial_{ij} f$ are bounded. Then, we have:*

$$\mathcal{D}_\mu f(t, X(t)) = \left[\partial_t f + \mathcal{D}_\mu X(t) \cdot \nabla f + \frac{i\mu\sigma^2}{2} \Delta f\right](t, X(t)). \tag{2.12}$$

2.2.4. Examples. — We compute the stochastic derivative in some famous examples, like the Ornstein-Uhlenbeck process and a Brownian mation in an external force.

2.2.4.1. The Ornstein-Uhlenbeck process. — A good model of the Brownian motion of a particle with friction is provided by the Ornstein-Uhlenbeck equation:

$$\begin{cases} X''(t) = -\alpha X'(t) + \sigma\xi(t) \\ X(0) = X_0,\ X'(0) = V_0, \end{cases} \tag{2.13}$$

where X(t) is the position of the particle at time, α is the friction coefficient, σ is the diffusion coefficient, X_0 and V_0 are given Gaussian variables, ξ is "white noise". The term $-\alpha X'(t)$ represents a frictional damping term.

The stochastic differential equation satisfied by the velocity process $V(t) := Y'(t)$ is given by:

$$\begin{cases} dV(t) = -\alpha V(t)dt + \sigma dW(t) \\ V(0) = V_0, \end{cases} \tag{2.14}$$

We can explicitly compute $\mathcal{D}V$ and $\mathcal{D}^2 V$:

Lemma 2.5. — *Let $V(\cdot)$ be a solution of*

$$\begin{cases} dV(t) = -\alpha V(t)dt + \sigma dW(t) \\ V(0) = V_0, \end{cases} \tag{2.15}$$

where V_0 has a normal distribution with mean zero and variance $\frac{\sigma^2}{2\alpha}$.

Then $V \in \mathcal{C}^2(]0, +\infty))$ and:

$$\mathcal{D}V(t) = -i\alpha V(t) \tag{2.16}$$

$$\mathcal{D}^2 V(t) = -\alpha^2 V(t). \tag{2.17}$$

Proof. — The solution is a Gaussian process explicitly given by:

$$\forall t \geqslant 0,\ V(t) = V_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW(s). \tag{2.18}$$

Therefore, we can compute the expectation and the variance of the normal variable $V(t)$:

$$\begin{cases} E[V(t)] = E[V_0]e^{-\alpha t} \\ \mathrm{Var}(V(t)) = \frac{\sigma^2}{2\alpha} + \left(\mathrm{Var}(V_0)) - \frac{\sigma^2}{2\alpha}\right) e^{-2\alpha t}, \end{cases} \tag{2.19}$$

We notice, as in [**30**], that if V_0 has a normal distribution with mean zero and variance $\frac{\sigma^2}{2\alpha}$, then X is a stationary gaussian process which distribution $p_t(x)$ at each time t reads

$$p_t(x) = \frac{\sqrt{\alpha}}{\sqrt{\pi}\sigma} e^{-\frac{\alpha x^2}{\sigma^2}}. \tag{2.20}$$

As a consequence, we have

$$\forall t \geqslant 0, \ln(p_t(x)) = \ln(\frac{\sqrt{\alpha}}{\sqrt{\pi}\sigma}) - \frac{\alpha x^2}{\sigma^2}, \tag{2.21}$$

and

$$\sigma^2 \partial_x \ln(p_t(x)) = \sigma^2 \frac{-2\alpha x}{\sigma^2} = -2\alpha x. \tag{2.22}$$

Moreover, we have

$$DV(t) = -\alpha V(t), \tag{2.23}$$

and according to theorem 1.1, we obtain

$$D_* V(t) = -\alpha V(t) - \sigma^2 \partial_x \ln(p_t(V(t))) = \alpha V(t). \tag{2.24}$$

Therefore $\mathcal{D}V(t) = -i\alpha V(t)$, and using the $\mathbb{C}-$linearity of $\mathcal{D}$, we obtain $\mathcal{D}^2 V(t) = -\alpha^2 V(t)$, which concludes the proof. □

2.2.4.2. Brownian particle submitted to an external force. — In some examples of random mechanics, one has to consider the stochastic differential system:

$$\begin{cases} dX(t) = V(t)dt \\ dV(t) = -\alpha V(t)dt + K(X(t))dt + \sigma dW(t) \\ X(0) = X_0, \ V(0) = V_0, \end{cases} \tag{2.25}$$

X and V may represent the position and the velocity of a particle of mass m being under the influence of an external force $F = -\nabla U$ where U is a potential. Set $K = F/m$. The "free" case $K = 0$ is the above example.

When $K(x) = -\omega^2 x$ (a linear restoring force), the system can also be seen as the random harmonic oscillator. In this case, it can be shown that if (X_0, V_0) has an appropriate gaussian distribution then $(X(t), V(t))$ is a stationary gaussian process in the same way as before.

Let us come back to the general case.

First, we remark that X is Nelson-differentiable and we have $DX(t) = D_*X(t) = V(t)$. Moreover, Nelson claims in ([**53**],p.83-84) that, when the particle is in equilibrium with a special stationary density,

$$DV(t) = -\alpha V(t) + K(X(t)), \tag{2.26}$$

$$D_*V(t) = \alpha V(t) + K(X(t)). \tag{2.27}$$

We can summarize these results with the computation of $\mathcal{D}$:

$$\mathcal{D}X(t) = V(t), \tag{2.28}$$

$$\mathcal{D}^2X(t) = K(X(t)) - i\alpha V(t). \tag{2.29}$$

CHAPTER 3

PROPERTIES OF THE STOCHASTIC DERIVATIVES

3.1. Product rules

In chapter 7, we develop a stochastic calculus of variations. In many problems,we will need the analogue of the classical formula of *integration by parts*, based on the following identity, called the *product or Leibniz rule*

$$\frac{d}{dt}(fg) = \frac{df}{dt}g + f\frac{dg}{dt}, \tag{P}$$

where f, g are two given functions.

Using a previous work of Nelson [**53**], we generalize this formula for our stochastic derivative. We begin by recalling the fundamental result of Nelson on a product rule formula for backward and forward derivatives:

Theorem 3.1. — *Let $X, Y \in \mathcal{C}^1(I)$, then we have:*

$$\frac{d}{dt}E[X(t)\cdot Y(t)] = E[DX(t)\cdot Y(t) + X(t)\cdot D_*Y(t)] \tag{3.1}$$

We refer to ([**53**],p.80-81) for a proof.

Remark 3.1. — *It must be pointed out that this formula mixes the backward and forward derivatives. As a consequence, even without our definition of the stochastic derivative, which takes into account these two quantities, the previous product rule suggests the construction of an operator which mixes these two terms in a "symmetrical" way.*

We now take up the various consequences of this formula regarding our operator $\mathcal{D}$. A straightforward calculation gives:

Lemma 3.1. — *Let $X, Y \in \mathcal{C}^1(I)$, we then have:*

$$\frac{d}{dt}E[X(t)\cdot Y(t)] = E[\mathrm{Re}(\mathcal{D}X(t))\cdot Y(t) + X(t)\cdot \mathrm{Re}(\mathcal{D}Y(t))] \tag{3.2}$$

$$E[\mathrm{Im}(\mathcal{D}X(t))\cdot Y(t)] = E[X(t)\cdot \mathrm{Im}(\mathcal{D}Y(t))] \tag{3.3}$$

Lemma 3.2. — *Let $X, Y \in \mathcal{C}^1_{\mathbb{C}}(I)$. We write $X = X_1 + iX_2$ and $Y = Y_1 + iY_2$ where $X_i, Y_i \in \mathcal{C}^1(I)$. Therefore :*

$$E[\mathcal{D}_\mu X \cdot Y + X \cdot \mathcal{D}_\mu Y] = \frac{d}{dt} g(X(t), Y(t)) + r(X(t), Y(t)), \tag{3.4}$$

where

$$g(X, Y) = E[X \cdot Y], \tag{3.5}$$

and

$$\begin{aligned} r(X,Y) &= -2E[Y_1 \cdot \mathrm{Im}(\mathcal{D}_\mu X_2)] - 2E[Y_2 \cdot \mathrm{Im}(\mathcal{D}_\mu X_1)] \\ &\quad + i\left(2E[Y_1 \cdot \mathrm{Im}(\mathcal{D}_\mu X_1)] - 2E[Y_2 \cdot \mathrm{Im}(\mathcal{D}_\mu X_2)]\right). \end{aligned} \tag{3.6}$$

Proof. — We have

$$\begin{aligned} Y\mathcal{D}_\mu X &= Y_1 \mathrm{Re}(\mathcal{D}_\mu X_1) - Y_1 \mathrm{Im}(\mathcal{D}_\mu X_2) \\ &\quad - Y_2 \mathrm{Im}(\mathcal{D}_\mu X_1) - Y_2 \mathrm{Re}(\mathcal{D}_\mu X_2) \\ &\quad + i\left(Y_1 \mathrm{Im}(\mathcal{D}_\mu X_1) + Y_1 \mathrm{Re}(\mathcal{D}_\mu X_2) + Y_2 \mathrm{Re}(\mathcal{D}_\mu X_1) - Y_2 \mathrm{Im}(\mathcal{D}_\mu X_2)\right). \end{aligned} \tag{3.7}$$

In a symmetrical way, we obtain

$$\begin{aligned} X\mathcal{D}_\mu Y &= X_1 \mathrm{Re}(\mathcal{D}_\mu Y_1) - X_1 \mathrm{Im}(\mathcal{D}_\mu Y_2) \\ &\quad - X_2 \mathrm{Im}(\mathcal{D}_\mu Y_1) - X_2 \mathrm{Re}(\mathcal{D}_\mu Y_2) \\ &\quad + i\left(X_1 \mathrm{Im}(\mathcal{D}_\mu Y_1) + X_1 \mathrm{Re}(\mathcal{D}_\mu Y_2) + X_2 \mathrm{Re}(\mathcal{D}_\mu Y_1) - X_2 \mathrm{Im}(\mathcal{D}_\mu Y_2)\right). \end{aligned} \tag{3.8}$$

Forming the sum of these expressions and using lemma 3.1, we obtain (3.4). □

The next lemma will be of importance in chapter 7 for the derivation of the stochastic analogue of the Euler-Lagrange equations:

Lemma 3.3. — *Let $X, Y \in \mathcal{C}^1_{\mathbb{C}}(I)$. We write $X = X_1 + iX_2$ and $Y = Y_1 + iY_2$ where $X_i, Y_i \in \mathcal{C}^1(I)$. Therefore, we have:*

$$E[\mathcal{D}_\mu X \cdot Y + X \cdot \mathcal{D}_{-\mu} Y] = \frac{d}{dt} g(X(t), Y(t)) \tag{3.9}$$

where $g(X, Y) = E[X_1 \cdot Y_1 - X_2 \cdot Y_2] + iE[Y_1 \cdot X_2 + Y_2 \cdot X_1] = E[X \cdot Y]$

Proof. — We have

$$\begin{aligned} Y\mathcal{D}_\mu X &= Y_1 \Re(\mathcal{D}_\mu X_1) - Y_1 \Im(\mathcal{D}_\mu X_2) \\ &\quad - Y_2 \Im(\mathcal{D}_\mu X_1) - Y_2 \Re(\mathcal{D}_\mu X_2) \\ &\quad + i\left(Y_1 \Im(\mathcal{D}_\mu X_1) + Y_1 \Re(\mathcal{D}_\mu X_2) + Y_2 \Re(\mathcal{D}_\mu X_1) - Y_2 \Im(\mathcal{D}_\mu X_2)\right), \end{aligned} \tag{3.10}$$

and in a symmetrical way

$$\begin{aligned} X\mathcal{D}_{-\mu} Y &= (X_1 + iX_2)(\overline{\mathcal{D}_\mu Y_1} + i\overline{\mathcal{D}_\mu Y_2}) \\ &= X_1 \Re(\mathcal{D}_\mu Y_1) + X_1 \Im(\mathcal{D}_\mu Y_2) \\ &\quad + X_2 \Im(\mathcal{D}_\mu Y_1) - X_2 \Re(\mathcal{D}_\mu Y_2) \\ &\quad + i\left(-X_1 \Im(\mathcal{D}_\mu Y_1) + X_1 \Re(\mathcal{D}_\mu Y_2) + X_2 \Re(\mathcal{D}_\mu Y_1) + X_2 \Im(\mathcal{D}_\mu Y_2)\right). \end{aligned} \tag{3.11}$$

We form the sum of these expressions and we use the lemma 3.1 to obtain (3.4). □

3.1.1. A new algebraic structure. — A convenient way to write equation (3.9) is to use the following Hermitian product:

For all $X, Y \in \mathcal{P}_{\mathbb{C}}$, we denote by $\star$ the product

$$X \star Y = X \cdot \overline{Y}, \tag{3.12}$$

where . denotes the usual scalar product.

Formula (3.9) is then equivalent to:

$$\mathcal{D}E[X \star Y] = E\left[\mathcal{D}X \star Y + X \star \mathcal{D}Y\right], \tag{3.13}$$

where we have implicitly used the fact that $\mathcal{D}$ reduces to d/dt when this quantity has a sense.

This new form leads us to the introduction of the following algebraic structure, which is, as far as we know, new. Let δ be the canonical mapping

$$\delta : \begin{array}{ccc} \mathcal{P}_{\mathbb{C}} \otimes \mathcal{P}_{\mathbb{C}} & \to & \mathcal{P}_{\mathbb{C}} \\ X \otimes Y & \mapsto & X \star \overline{Y}. \end{array} \tag{3.14}$$

We define for $\mathcal{D}$ the quantity $\Delta(\mathcal{D}) = \mathcal{D} \otimes 1 + 1 \otimes \mathcal{D}$, which we will call the *coproduct* of $\mathcal{D}$. Then, denoting by E the classical mapping which takes the expectation of a given stochastic process, we obtain the following diagram:

$$\begin{array}{ccc} \mathcal{P}_{\mathbb{C}} \otimes \mathcal{P}_{\mathbb{C}} & \xrightarrow{\Delta(D)} & \mathcal{P}_{\mathbb{C}} \otimes \mathcal{P}_{\mathbb{C}} \\ X \otimes Y & \longrightarrow & \mathcal{D}X \otimes Y + X \otimes \mathcal{D}Y \\ \big\downarrow \delta & & \big\downarrow \delta \\ X \star Y & \longrightarrow & \mathcal{D}X \star Y + X \star \mathcal{D}Y \\ \big\downarrow \mathrm{E} & & \big\downarrow \mathrm{E} \\ E[X \star Y] & \xrightarrow{\mathcal{D}} & E[\mathcal{D}X \star Y + X \star \mathcal{D}Y] \end{array} \tag{3.15}$$

This structure is similar to the classical algebraic structure of *Hopf algebra*. The difference is that we perturb the classical relations by a linear mapping, here given by E. It will be interesting to study this kind of structure in full generality.

3.2. Nelson differentiable processes

3.2.1. Definition. — We define a special class of processes, called Nelson-differentiable processes, which will play an important role in the stochastic calculus of variations of chapter 7.

Definition 3.1. — *A process $X \in \mathcal{C}^1(I)$ is called Nelson differentiable if $DX = D_*X$.*

Notation 3.1. — *We denote by $\mathcal{N}^1(I)$ the set of Nelson differentiable processes.*

A better definition is perhaps to use $\mathcal{D}$ instead of D and D_* saying that Nelson differentiable processes have a real stochastic derivative.

The main idea behind this definition is that we want to define a class $\mathcal{P}$ of processes in $\mathcal{C}^1(I)$ such that if $X \in \mathcal{C}^1(I)$ then for all $Y \in \mathcal{P}$, we have

$$\mathrm{Im}(\mathcal{D}(X+Y)) = \mathrm{Im}(\mathcal{D}X).$$

This condition imposes that $\mathrm{Im}(\mathcal{D}Y) = 0$.

This condition will appear more clearly in chapter 7 concerning the stochastic calculus of variations.

Remark 3.2. — *We must keep in mind that our definition of the stochastic derivative follows the idea of the scale calculus developed in* [**13**] *to study non-differentiable functions. In that context, the existence of an imaginary part for the scale derivative of a function is seen as a resurgence of its non-differentiability. In particular, when the underlying function is differentiable then the scale derivative is real. That is why we have chosen to call processes such that $D = D_*$ Nelson differentiable.*

The definition of Nelson differentiable processes is only given for processes in $\mathcal{C}^1(I)$. It is not at all clear to know what is the correct extension to $\mathcal{C}^1_{\mathbb{C}}(I)$. As we have no use of such kind of notion on $\mathcal{C}^1_{\mathbb{C}}(I)$ we don't discuss this point here.

Of course a difficult problem is to characterize these processes. The next section discusses some examples.

3.2.2. Examples of Nelson-differentiable process. — We give examples of Nelson-differentiable processes.

3.2.2.1. Differentiable deterministic process. — It is probably the first and the simplest example. Let $x(\cdot)$ be a differentiable deterministic process defined on $I \times \Omega$. The past $\mathcal{P}$ and the future $\mathcal{F}$ are trivial:

$$\forall t \in I,\ \mathcal{P}_t = \mathcal{F}_t = \{\emptyset, \Omega\}.$$

As a consequence, we have

$$\forall t \in I,\ Dx(t) = D_*x(t) = x'(t),$$

where x' is the usual derivative of x.

3.2.2.2. A very special random example. — Let $X \in \mathcal{C}^1(I)$. In [**53**], Nelson shows that X is a constant (i.e. X(t) is the same random variable for all t) if and only if : $\forall t \in I,\ DX(t) = D_*X(t) = 0$. So it provides us a random example of $\mathcal{N}^1(I)-$process.

3.2.2.3. Nelson-differentiable diffusion processes. — Using theorem 1.1, we can find a sufficient and necessary condition for a diffusion process to be a Nelson-differentiable process:

Lemma 3.4. — *Let $X \in \Lambda_d$ with $\sigma = const$, then $X \in \mathcal{N}^1(I)$ if and only if*

$$\nabla(\sigma^2 p)(t, X(t)) = 0. \tag{3.16}$$

When the diffusion equation is time homogeneous and the solutions have a density, we note that this density must be a stationary density. Moreover, the Fokker-Planck equation (Kolmogorov forward equation) allows us to give a necessary condition (a relation between the drift and the diffusion coefficient) for a diffusion equation to give a Nelson-differentiable solution.

3.2.2.4. The random harmonic oscillator. — The random harmonic oscillator satisfies the stochastic differential equation:

$$\begin{cases} dX(t) = V(t)dt \\ dV(t) = -\alpha V(t)dt - \omega^2 X(t)dt + \sigma dW(t) \\ X(0) = X_0,\ V(0) = V_0, \end{cases} \tag{3.17}$$

As a consequence, we have $X(t) = \int_0^t V(s)ds$ with $E\left[\int_0^b |V(s)|^2 ds\right] < \infty$ $(b > 0)$, and X has a strong derivative in L^2. We then obtain $DX(t) = D_*X(t) = V(t)$. Finally, we have $X \in \mathcal{N}^1([0, b])$ and $\mathcal{D}X(t) = V(t)$.

3.2.3. Product rule and Nelson-differentiable processes. —

Corollary 3.1. — *Let $X, Y \in \mathcal{C}^1_{\mathbb{C}}(I)$. If X is Nelson-differentiable then :*

$$E[\mathcal{D}_\mu X(t) \cdot Y(t) + X(t) \cdot \mathcal{D}_\mu Y(t)] = \frac{d}{dt} E(X(t), Y(t)) \tag{3.18}$$

Proof. — This is a simple consequence of the fact that if $X = X_1 + iX_2$ is Nelson-differentiable then $\mathrm{Im}(\mathcal{D}_\mu X_1) = \mathrm{Im}(\mathcal{D}_\mu X_2) = 0$. □

PART II

STOCHASTIC EMBEDDING PROCEDURES

CHAPTER 4

STOCHASTIC EMBEDDING OF DIFFERENTIAL OPERATORS

A natural question concerning ordinary and partial differential equations concerns their behaviour under small random perturbations. This problem is particularly important in natural phenomena where we know that models are only an approximation of the real setting. For example, the study of the long term behaviour of the solar system is usually done by running numerical computations on the n-body problem. However, many effects in the solar systems are not included in this model and can be of importance if one looks for a long term integration, as non conservative effects (due to tidal forces between planets) and the oblatness of the sun which is not yet modelled by a differential equation.

The main problem is then to find the correct analogue of a given differential equation taking into account the following facts:

i) The classical equation is a good model at least in first approximation,
ii) One must extend this equation to stochastic processes.

Using the stochastic derivative introduced in the previous part, we give a natural embedding of partial or ordinary differential equations into stochastic partial or ordinary differential equations. It must be pointed out that we do not perturb the classical equation by a random noise or anything else. In this respect we are far from the usual way of thinking underlying the fields of stochastic differential equations or stochastic dynamical systems.

Of course, having this natural embedding, we can naturally define what a stochastic perturbation of a differential equation is. This is simply a stochastic perturbation of the stochastic embedding of the given equation. The main point is that we stay in the same class of objects dealing with perturbations, which is not the case in the stochastic theory of differential equations, where we jump from classical solutions to stochastic processes in

one step using for example Ito's stochastic calculus[(1)].

In this part we first give a general embedding procedure for partial differential equations. We discuss classical examples, in particular first and second order differential equations. The case of Lagrangian systems is studied in details in chapter 7. An important part of classical differential equations coming from mechanics are *reversible*. This property is not conserved by the previous stochastic embedding procedure. We define a special embedding called reversible, which preserves this property, meaning that if X is a solution of the stochastic embedded equation, then $\tilde{X}$, the reversed process, is again a solution.

4.1. Stochastic embedding of differential operators

In this part, we first give an abstract embedding procedure based on an extension of the classical derivative defined in the previous part. We then specialize our embedding procedure using the stochastic derivative.

4.1.1. Abstract embedding. — Let $\mathcal{A}$ be a ring, we denote by $\mathcal{A}[x]$ the ring of polynomials with coefficients in $\mathcal{A}$. Let $\mathcal{A} = C^1(\mathbb{R}^d \times \mathbb{R})$.

Definition 4.1. — *A differential operator is an elements of $\mathcal{A}[d/dt]$.*

Let $O \in \mathcal{A}[d/dt]$, the differential operator O is of the form

$$O = a_0(\bullet, t) + a_1(\bullet, t)\frac{d}{dt} + \cdots + a_n(\bullet, t)\frac{d^n}{dt^n},\ a_i \in \mathcal{A},\ = 0, \ldots, n, \tag{4.1}$$

for a given $n \in \mathbb{N}$, called the *degree* of O.

The action of O on a given function $x : \mathbb{R} \to \mathbb{R}^d$, $t \mapsto x(t)$ is denoted $O \cdot x$ and defined by

$$O \cdot x = \sum_{i=0}^{n} a_i(x(t), t)\frac{dx}{dt}. \tag{4.2}$$

***Definition 4.2* (Abstract stochastization)**. — *Let $O \in \mathcal{A}[d/dt]$ be a differential operator, of the form*

$$O = a_0(\bullet, t) + a_1(\bullet, t)\frac{d}{dt} + \cdots + a_n(\bullet, t)\frac{d^n}{dt^n},\ a_i \in \mathcal{A},\ = 0, \ldots, n, \tag{4.3}$$

where $n \in \mathbb{N}$ is given.

(1) This remark is also valid for all the theories of this kind, using your favourite stochastic calculus, like Malliavin calculus for example.

The stochastic embedding of O with respect to the extension $\delta : \mathcal{P} \to \mathcal{P}$ is an element O_δ of $\mathcal{P}[\delta]$ defined by

$$O_\delta = a_0(\bullet, t) + a_1(\bullet, t)\delta + \cdots + a_n(\bullet, t)\delta^n,\ a_i \in \mathcal{P},\ i = 0, \ldots, n, \tag{4.4}$$

where $\delta^n = \delta \circ \cdots \circ \delta$.

The action of O_δ on a given stochastic process X, denoted by $O_\delta \cdot X$ is defined by

$$O_\delta \cdot X = \sum_{i=0}^{n} a_i(X, t)\delta^i X, \tag{4.5}$$

where the notation $a_i(X, t)$ stands for the stochastic process defined for all $\omega \in \Omega$ by

$$a_i(X, y)(\omega) = a_i(X(\omega, t), t). \tag{4.6}$$

The main property of this embedding is the fact that

$$O_\delta \mid_{\mathcal{P}^n_{\text{det}}} = O, \tag{4.7}$$

so that the classical differential equation associated to O, and given by

$$O \cdot x = 0, \tag{E}$$

is contained in the stochastic differential equation

$$O_\delta \cdot X = 0. \tag{SE}.$$

4.1.2. Nelson Stochastic embedding. — Using the stochastic derivative, we have a particular stochastic embedding procedure.

***Definition 4.3* (Stochastization).** — *Let $O \in \mathcal{A}[d/dt]$ be a differential operator, of the form*

$$O = a_0(\bullet, t) + a_1(\bullet, t)\frac{d}{dt} + \cdots + a_n(\bullet, t)\frac{d^n}{dt^n},\ a_i \in \mathcal{A},\ = 0, \ldots, n, \tag{4.8}$$

where $n \in \mathbb{N}$ is given.

The stochastic embedding of O with respect to the stochastic extension $\mathcal{D}_\mu$ is an element O_{stoc} of $C^1(I)[\mathcal{D}_\sigma]$ defined by

$$O_{\text{stoc}} = a_0(\bullet, t) + a_1(\bullet, t)\mathcal{D} + \cdots + a_n(\bullet, t)\mathcal{D}^n,\ a_i \in C^1(I),\ i = 0, \ldots, n. \tag{4.9}$$

We denote by $\mathcal{S}$ the operator associating to an operator O of the form 4.8 the operator O_{stoc}. As a consequence, we will frequently use the notation $\mathcal{S}(O)$ for O_{stoc}.

In some occasions, in particular for the Euler-Lagrange equation, we will need to consider differential operators in a non-standard form. Precisely, we need to consider operators like

$$B_a = \frac{d}{dt} \circ a(\bullet, t). \tag{4.10}$$

This notation means that B_a acts on a given function as

$$B_a \cdot x = \frac{d}{dt} \left(a(x(t), t) \right). \tag{4.11}$$

The basic idea is to define the stochastic embedding of B_a as follow:

Definition 4.4. — *The stochastic embedding of the basic brick B_a is given by*

$$\mathcal{B}_a = \mathcal{D} \circ a(\bullet, t). \tag{4.12}$$

However, classical properties of the differential calculus allow us to write B_a equivalently as

$$B_a \cdot x = a'(x) \frac{dx}{dt}. \tag{4.13}$$

The stochastic embedding of this new form of B_a is given by

$$\mathbb{B}_a.X = a'(X)\mathcal{D}X. \tag{4.14}$$

The main problem is that in general, we do not have

$$\mathcal{B}_a = \mathbb{B}_a, \tag{4.15}$$

as in the classical case.

This reflects the fact that $\mathcal{S}$ acts on operators of a given form and not on operators as an abstract element of a given algebra. In particular, this is not a mapping.

Nevertheless, there exists a class of functions a such that equation (4.15) is valid:

Lemma 4.1. — *Equation (4.15) is satisfied on the set Λ_d with constant diffusion if a is an harmonic function.*

Proof. — This follows easily from corollary 2.2. □

In the sequel we study some basic properties of this embedding procedure on differential equations.

4.2. First examples

4.2.1. First order differential equations. — Let us consider a first order differential equation

$$\frac{dx}{dt} = f(x,t), \qquad 1-(ODE)$$

where $x \in \mathbb{R}$ and $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a given function. The stochastic embedding of (1-ODE) leads to

$$\mathcal{D}X = F(X,t), \qquad 1-(SODE)$$

where F is real valued.

The reality of F imposes important constraints on solutions of 1-(SODE). Indeed, we must have

$$DX = D_*X,$$

so that X belongs to the class of Nelson-differentiable processes.

In our general philosophy, ordinary differential equations are only coarse approximations to reality which must include stochastic behaviour in its foundation. A stochastic perturbation of a first order differential equation is then highly non-trivial. Indeed, we must consider SODE's of the form

$$\mathcal{D}X = F(X,t) + \epsilon G(X,t),$$

where $G(X,t)$ is now complex valued. As a consequence, we allow solutions to leave the Nelson-differentiable class.

4.2.2. Second order differential equations. — Let us consider a second order differential equation

$$\frac{d^2x}{dt^2} + a(x)\frac{dx}{dt} + b(x) = 0, \qquad (2-(ODE)$$

where $x \in \mathbb{R}$, and $a, b : \mathbb{R} \to \mathbb{R}$ are given functions. The stochastic embedding of $(2-(ODE))$ leads to

$$\mathcal{D}^2X + a(X)\mathcal{D}X + b(X) = 0.$$

In this case, contrary to what happens for first order differential equations, we have no reality condition which constrains our stochastic process.

In order to study such kind of equations, one can try to reduce it to a first order equation, using standard ideas. We denote by $Y = \mathcal{D}X$, then the second order equation

is equivalent to the following system of first order stochastic differential equations:

$$\left\{ \begin{array}{rcl} \mathcal{D}X & = & Y, \\ \mathcal{D}Y & = & -a(X)Y - b(X). \end{array} \right. \tag{4.16}$$

One must be careful to take $Y \in \mathcal{C}^1_{\mathbb{C}}(I)$ as Y is a priori a complex stochastic process. This remark is of importance since if we apply the stochastic embedding procedure[(2)] to the classical system of first order differential equations

$$\left\{ \begin{array}{rcl} \dfrac{dx}{dt} & = & y, \\ \dfrac{dy}{dt} & = & -a(x)y - b(x), \end{array} \right. \tag{4.17}$$

by saying that we apply separately the stochastic embedding on each differential equations, we obtain the stochastic equation (4.16) but with $Y \in \mathcal{C}^1(I)$, which imposes strong constraints on the solutions of our equations.

This example proves that the stochastic embedding procedure is not so easy to define if one wants to deal with systems of differential equations. We will return on this problem concerning the stochastic embedding of Hamiltonian systems.

[(2)] Note that we have not defined the stochastic embedding procedure on systems of differential equations.

CHAPTER 5

REVERSIBLE STOCHASTIC EMBEDDING

5.1. Reversible stochastic derivative

In our construction of the stochastic derivative, we have imposed some constraints as for example the gluing to the classical derivative on differentiable deterministic processes. We have moreover kept some properties of the classical derivative such as linearity. However, we have not conserved more important properties of the classical derivative which are used in the study of classical differential equations. For example, let us consider

$$\frac{d^2x}{dt^2} = f(x), \tag{E}$$

which is the basic equation of Newton's mechanics. An important property of this kind of equations is its *reversibility*:

Let $t \to x(t)$ be a solution of (E). We denote by $\tilde{x}(t) = x(-t)$. Then, we have

$$\frac{d^2\tilde{x}}{dt^2} = \frac{d}{dt}(-\frac{dx}{dt}(-t)) = \frac{d^2x}{dt^2}(-t) = f(x(-t)) = f(\tilde{x}(t)),$$

proving that the reversed solution $\tilde{x}(t)$ is again a solution of the same equation. In this case, we say that the differential equation is *reversible*.

The reversibility argument used the following important property:

$$\frac{d}{dt}(x(-t)) = -\frac{dx}{dt}(-t). \tag{R}$$

The natural way to introduce a notion of *reversibility* is then to look for the stochastic differential equation satisfied by $\tilde{X}(t) = X(-t) \in \mathcal{C}^1(I)$ the reversed processes. However, in general, we do not have access to $D\tilde{X}$ or $D_*\tilde{X}$. As a consequence, a definition using this characterization is not effective. In the following, we follow a different strategy.

A convenient way to characterize the reversibility of a given differential equation, described by a differential operator

$$O = \sum_i a_i \frac{d^i}{dt^i} \in \mathbb{R}[d/dt] \tag{5.1}$$

is to prove that this operator is invariant under the substitution

$$r : \mathbb{R}[d/dt] \longrightarrow \mathbb{R}[d/dt] \tag{5.2}$$

which is $\mathbb{R}$ linear and defined by

$$r(d/dt) = -d/dt. \tag{5.3}$$

We then introduce in our setting, the following analogous substitution:

Definition 5.1. — *The reversibility operator* $R : \mathbb{C}[D, D_*] \to \mathbb{C}[D, D_*]$ *is a* $\mathbb{C}$ *morphism defined by*

$$R(D) = -D_*, \quad R(D_*) = -D. \tag{5.4}$$

We have the following immediate consequence of the definition:

Lemma 5.1. — *The reversibility operator is an involution of* $\mathbb{C}[D, D_*]$.

This operator acts non trivially on our stochastic derivative. Precisely, we have:

Lemma 5.2. —

$$R(\mathcal{D}) = -\overline{\mathcal{D}}. \tag{5.5}$$

The complex nature of the stochastic derivative induces new phenomenon which are different from the classical case. For example, we have

$$R(\mathcal{D}^2) = \overline{\mathcal{D}}^2, \tag{5.6}$$

contrary to what happens for r.

We now define our notion of a reversible stochastic equation.

Definition 5.2. — *[Reversibility] Let* $O \in \mathbb{R}[D, D_*]$*, then the stochastic equation* $O \cdot X = 0$ *is reversible if and only if* $R(O) \cdot X = 0$.

A natural problem is the following:

Reversibility problem: *Find an operator such that the stochastic embedding of a reversible equation is again a reversible equation in the sense of definition 5.2.*

Let us consider the family of stochastic derivatives $\mathcal{D}_\mu$, $\mu = 0, \pm 1$. Without assuming a particular form for the underlying equation, the preservation of the reversible character reduces to prove that the operator δ which is chosen satisfies

$$R(\delta) = -\delta. \tag{5.7}$$

In the family of stochastic derivatives $\mathcal{D}_\mu$, $\mu = 0, \pm 1$, only one case is possible:

Lemma 5.3. — *A reversibility of a differential equation is always preserved under a stochastic embedding if and only if this embedding is associated to the stochastic derivative* $\mathcal{D}_0$.

Proof. — Essentially this follows from equation (5.5). If we want to preserve reversibility then the operator $\mathcal{D}_\mu$ must satisfied $R(\mathcal{D}_\mu) = -\mathcal{D}_\mu$. This is only possible if $\mathcal{D}_\mu$ is real, i.e. $\mu = 0$. □

It must be pointed out that the operator

$$\mathcal{D}_0 = \frac{D + D_*}{2},$$

has been obtained by different authors using the following argument:

If we use only D (or D_*) then, we give a special importance to the future (or past) of the process, which has no physical justification. As a consequence, one must construct an operator which combines these two quantities in a more or less symmetric way. The simplest combination is a linear one $aD + bD_*$ with equal coefficients $a = b$. The gluing to the classical derivative leads to $a = b = 1/2$.

The problem with this construction is that this argument is used on diffusion processes, where D and D_* are not *free*. As a consequence, working with D is the same (even if the connection with D_* is not trivial) than working with D_*. We can not really justify then the use of $\mathcal{D}_0$. It must be pointed out that E. Nelson [**53**] does not use $\mathcal{D}_0$ in his derivation of the Schrödinger equation, but simply D.

Here, this operator is obtained by specialization of $\mathcal{D}_\mu$, which form is imposed by our construction (linearity, gluing to the classical derivative, reconstruction property). The reconstruction property imposes that $\mu \neq 0$ unless we work with diffusion processes.

Imposing a new constraint on the reversibility on this operator leads us to $\mu = 0$. The operator $\mathcal{D}_0$ is of course defined on $\mathcal{C}^1(I)$, but in order to satisfy the whole constraints of

our construction, we must restrict its domain to diffusion processes.

We can of course find reversible equations without using $\mathcal{D}_0$ but $\mathcal{D}_\mu$. We keep the notations and conventions of chapter 4. We first define the action of R on a given operator of the form

$$\mathcal{O} = \sum_{i=0}^{n} a_i(\bullet, t)(-1)^i \overline{\mathcal{D}}^i. \tag{5.8}$$

Definition 5.3. — *The action of R on (5.8) is denoted $R(\mathcal{O})$ and defined by*

$$R(\mathcal{O}) = \sum_{i=0}^{n} a_i(\bullet, t)\mathcal{D}^i. \tag{5.9}$$

The definition 5.2 of a reversible equation can then be extended to cover operators of the form 5.8.

Using this definition, we can prove that the stochastic equation

$$\mathcal{D}_\mu^2 X = -\nabla U(X), \tag{E}$$

is reversible.

Indeed, we have:

Lemma 5.4. — *Equation (E) is reversible.*

Proof. — We have

$$\begin{aligned} R(\mathcal{D}_\mu^2 X + \nabla U(X)) &= \overline{\mathcal{D}}^2 X + \nabla U(X), \\ &= \overline{\mathcal{D}_\mu^2 X} + \nabla U(X). \end{aligned} \tag{5.10}$$

As U is real valued and X are real stochastic processes, we deduce from (E) that

$$\overline{\mathcal{D}_\mu^2 X} = -\overline{\nabla U(X)} = -\nabla U(X). \tag{5.11}$$

We deduce that

$$R(\mathcal{D}_\mu^2 X + \nabla U(X)) = 0, \tag{5.12}$$

which concludes the proof. □

5.2. Iterates

There exists a fundamental difference between $\mathcal{D}_0$ and $\mathcal{D}_\mu$, $\mu \neq 0$. The operator $\mathcal{D}_0$ send real stochastic processes to real stochastic processes in the contrary of $\mathcal{D}_\mu$, $\mu \neq 0$,

which leads to complex stochastic processes. As a consequence, the n-ième iterates of $\mathcal{D}_0$ is simply defined by

$$\mathcal{D}_0^n = \mathcal{D}_0 \circ \cdots \circ \mathcal{D}_0, \tag{5.13}$$

without problem, where a special extension of $\mathcal{D}_\mu$, $\mu \neq 0$ to complex stochastic processes must be discussed.

5.3. Reversible stochastic embedding

Using $\mathcal{D}_0$, we can define a stochastic embedding which conserves the fundamental property of reversibility of a given equation. We keep notations from chapter 4.

***Definition 5.4* (Reversible stochastization).** — *Let $O \in \mathcal{A}[d/dt]$ be a differential operator, of the form*

$$O = a_0(\bullet, t) + a_1(\bullet, t)\frac{d}{dt} + \cdots + a_n(\bullet, t)\frac{d^n}{dt^n},\ a_i \in \mathcal{A},\ = 0, \ldots, n,$$

where $n \in \mathbb{N}$ is given.

The reversible stochastic embedding of O is an element O_{rev} of $\mathcal{C}^1(I)[\mathcal{D}_0]$ defined by

$$O_{\text{rev}} = a_0(\bullet, t) + a_1(\bullet, t)\mathcal{D}_0 + \cdots + a_n(\bullet, t)\mathcal{D}_0^n,\ a_i \in \mathcal{C}^1(I),\ i = 0, \ldots, n. \tag{5.14}$$

A differential equation (E) is defined by a differential operator $O \in \mathcal{A}[d/dt]$, i.e. an equation of the form

$$O \cdot x = 0, \tag{E}$$

where x is a function.

Using stochastization, the reversible stochastic analogue of (E) is defined by

$$O_{\text{rev}} \cdot X = 0, \tag{RSE}$$

where X is a stochastic process.

5.4. Reversible versus general stochastic embedding

The reversible stochastic embedding leads to very different results than the general stochastic embedding. We can already see this difference on first order differential equations. Let us consider

$$\frac{dx}{dt} = f(x),$$

where $x \in \mathbb{R}$ and f is a real valued function. The reversible stochastic embedding gives

$$\mathcal{D}_0 X = f(X).$$

Contrary to what happens for the stochastic embedding, this equation does not impose for the solution to be a Nelson differentiable processes.

5.5. Stochastic mechanics and the Stochastization procedure

5.5.1. The Stochastic Newton Equation. — The stochastized version of the classical system:

$$\begin{aligned} \dot{x}(t) &= v(t) \\ \dot{v}(t) &= K(x(t)) \end{aligned} \tag{5.15}$$

is given by:

$$\begin{aligned} \mathcal{D}X(t) &= V(t) \\ \mathcal{D}V(t) &= K(X(t)) \end{aligned} \tag{5.16}$$

where $V \in \mathcal{C}^1_{\mathbb{C}}(I)$ and K is a force: $K(x) = -\nabla U(x)$ and U a potential.

We can give at least two different kind of solutions of this equation, and so two relevant models.

In the first one, the component X is the position in the Ornstein-Uhlenbeck theory of Brownian Motion and is not submitted to a random noise. The system writes:

$$\begin{cases} dX(t) = V(t)dt \\ dV(t) = -\alpha V(t)dt + K(X(t))dt + \sigma dW(t) \\ X(0) = X_0,\ V(0) = V_0, \end{cases} \tag{5.17}$$

We have noticed in a previous section that, at an equilibrium (i.e. X has a stationary density) and if e^{-U} is integrable, then:

$$\mathcal{D}X(t) = V(t), \tag{5.18}$$

$$\mathcal{D}^2X(t) = K(X(t)) - i\alpha V(t). \tag{5.19}$$

Therefore (X, V) solves the Newton stochastized system (5.16) if and only if $\alpha = 0$. Moreover we note in this particuliar case that X is a Nelson-differentiable process.

The second one is described by

$$dX(t) = b(t, X(t))dt + \sigma dW(t), \tag{5.20}$$

where the function b must be determined. In this case, we proved that the density $p_t(x)$ of a solution X of (5.16) writes $p_t(x) = \Psi(t,x)\overline{\Psi}(t,x)$ where Ψ solves the Schrödinger equation: $i\sigma^2\partial_t\Psi + \frac{\sigma^4}{2}\partial_{xx}\Psi = U\Psi$. In this case, X is driven by a Brownian motion and is not Nelson-differentiable.

PART III

STOCHASTIC EMBEDDING OF LAGRANGIAN AND HAMILTONIAN SYSTEMS

CHAPTER 6

STOCHASTIC LAGRANGIAN SYSTEMS

Most of classical mechanics can be formulated using Lagrangian formalism ([**5**],[**2**]). Lagrangian mechanics contains important problems, like the n-body problem. Using our framework, we study Lagrangian dynamical systems under stochastic perturbations[(1)].

Our approach is first to embed classical Lagrangian systems, in particular the associated Euler-Lagrange equation (EL) in order to obtain an idea of what kind of equation govern stochastic Lagrangian systems. We then develop a stochastic calculus of variations. We obtain an analogue of the *least-action principle*[(2)] which gives a second stochastic Euler-Lagrange equation, denoted by (SEL) in the sequel. We then prove the following surprising result, called the coherence lemma: we have $\delta(EL) = (SEL)$.

The principal interest of Lagrangian systems is that the action of a group of symmetries leads to first integrals of motion, i.e. functions which are constants on solutions of the equations of motion. The celebrated theorem of E. Noether gives a precise relation between symmetries and first integrals. We prove a stochastic analogue of E. Nother theorem.

Finally, we prove that the stochastic embedding of Newton's Lagrangian systems lead to a non linear Schrödinger's equation for a given wave function whose modulus is equal to the probability density of the underlying stochastic process.

6.1. Reminder about Lagrangian systems

We refer to [**5**] for more details, as well as [**2**].

---

(1) For the n-body problem, which is usually used to study the long term behavior of the solar system [**47**], this problem is of crucial importance. Indeed, the n-body problem is only an approximation of the real problem, and even if some numerical simulations take into account relativistic effects [**40**], this is not sufficient [**50**].

(2) In our case, the word *least-action* is misleading and a better terminology is *stationary* (see below).

Lagrangian systems play a central role in dynamical systems and physics, in particular for mechanical systems. A Lagrangian system is defined by a Lagrangian function, commonly denoted by L, and depending on three variables: x, v, and t which belongs in the sequel to $\mathbb{R}$. As Lagrangian systems come from mechanics, the letter x stands for position, the letter v for speed and the letter t for time. In what follows, we consider a special type of Lagrangian function called admissible in the following.

Definition 6.1. — *An admissible Lagrangian function is a function L such that:*

i) *The function $L(x, v, t)$ is defined on $\mathbb{R}^d \times \mathbb{C}^d \times \mathbb{R}$, holomorphic in the second variable and real for $v \in \mathbb{R}$.*

ii) *L is autonomous, i.e. L does not depend on time.*

Condition i) is fundamental. This condition is necessary in order to apply the stochastization procedure (see below). The fact that we only consider autonomous Lagrangian function is due to technical difficulties in order to take into account backward and forward filtrations in the computation of the stochastic Euler-Lagrange equation (see below).

Remark 6.1. — *In applications, admissible Lagrangian functions L are analytic extensions to the complex domain of real analytic Lagrangian functions. For example, the classical Newtonian Lagrangian $L(x, v) = (1/2)v^2 - U(x)$, defined on an open*[3] *subset of $\mathbb{R} \times \mathbb{R}$, with an analytic potential is an admissible Lagrangian function.*

A Lagrangian function L being given, the equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) = \frac{\partial L}{\partial x}. \qquad (EL)$$

is called the *Euler-Lagrange equations*.

An important property of the Euler-Lagrange equation is that it derives from a *variational principle*, namely the *least action principle* (see [**5**],p.59). Precisely, a curve $\gamma : t \mapsto x(t)$ is an *extremal*[4] of the functional

$$J_{a,b}(\gamma) = \int_a^b L(x(t), \dot{x}(t), t)dt,$$

on the space of curves passing through the points $x(a) = x_a$ and $x(b) = x_b$, if and only if it satisfies the Euler-Lagrange equation along the curve $x(t)$.

[3] This Lagrangian function is not always defined on $\mathbb{R} \times \mathbb{R}$. An example is given by Newton's potential $U(x) = 1/x$, $x \in \mathbb{R}^*$.

[4] We refer to [**5**], chapter 3, §.12 for an introduction to the *calculus of variations*.

6.2. Stochastic Euler-Lagrange equations

We now apply our stochastic procedure $\mathcal{S}$ to an admissible Lagrangian.

Lemma 6.1. — *Let $L(x,v) : \mathbb{R}^d \times \mathbb{C}^d \to \mathbb{C}$ be an admissible Lagrangian function. The stochastic Euler-Lagrange equation obtained from (EL) by the stochastic procedure is given by*

$$\mathcal{D}_\mu \left(\frac{\partial L}{\partial v}(X(t), \mathcal{D}_\mu X(t) \right) = \frac{\partial L}{\partial x}(X(t), \mathcal{D}_\mu X(t)). \qquad \mathcal{S}(EL)$$

Proof. — The Euler-Lagrange equation associated to $L(x,v)$ can be seen as the following differential operator

$$O_{EL} = \frac{d}{dt} \circ \frac{\partial L}{\partial v} - \frac{\partial L}{\partial x},$$

acting on $(x(t), \dot{x}(t))$. The embedding of O_{EL} gives

$$\mathcal{O}_{EL} = \mathcal{D}_\mu \circ \frac{\partial L}{\partial v} - \frac{\partial L}{\partial x}.$$

As O_{EL} acts on $(x(t), \dot{x}(t))$, the operator $\mathcal{O}_{EL}$ acts on $(X(t), \mathcal{D}_\mu X(t))$. This concludes the proof. □

The free parameter $\mu \in \{-1, 0, 1\}$ can be fixed depending on the nature of the extension used.

It must be pointed out that there exist crucial differences between all these extensions due to the fact that $\mathcal{D}_\mu$ is complex valued for $\mu = \pm 1$ and real for $\mu = 0$. Indeed, let us consider the following admissible Lagrangian function:

$$L(x,v) = \frac{1}{2}v^2 - U(x),$$

where U is a smooth real valued function. Then, equation $\mathcal{S}$(EL) gives

$$\mathcal{D}_\mu V = U(X),$$

where $V = \mathcal{D}_\mu X$. When $\mu = \pm 1$, this equation imposes strong constraints on X due to the real nature of $U(X)$, namely that $\mathcal{D}_\mu^2 X \in \mathcal{N}^1(I)$.

On the contrary, when $\mu = 0$, i.e. in the *reversible case*, these *intrinsic* conditions disappear.

6.3. The coherence problem

Up to now, the stochastic embedding procedure can be viewed as a formal manipulation of differential equations. Moreover, as most classical manipulations on equations do not

commute with the stochastic embedding, this procedure is not canonical [5]. In order to rigidify this construction and to make precise the role of this stochastic embedding procedure, we study the following problem, called *the coherence problem*:

We know that the Euler-Lagrange equations are obtained via a least-action principle on a functional. The main problem is the existence of a stochastic analogue of this least-action principle, that we can call a *stochastic least action principle*, compatible with the stochastic embedding procedure.

$$
\begin{array}{ccc}
L(x(t),\dot{x}(t)) & \xrightarrow{\mathcal{S}} & L(X(t),\mathcal{D}X(t)) \\
\text{Least action principle}\Big\downarrow & & \Big\downarrow\text{Stochastic least action principle ?} \\
(\text{EL}) & \xrightarrow[\mathcal{S}]{} & (\text{SEL})
\end{array}
\tag{6.1}
$$

In the next chapter, we develop the necessary tools to answer to this problem, i.e. a *stochastic calculus of variations*. Note that due to the fact that the stochastic Lagrangian as well as the stochastic Euler-Lagrange equation are fixed, this problem is far from being trivial. The main result of the next chapter is the Lagrangian coherence lemma which says precisely that the stochastic Euler-Lagrange equation obtained via the stochastic embedding procedure coincide with the characterization of extremals for the functional associated to the stochastic Lagrangian function using the stochastic calculus of variations. As a consequence, we obtain a rigid picture involving the stochastic embedding procedure and a first principle via the stochastic least action principle.

This picture will be then extended in another chapter when dealing with the Hamiltonian part of this theory.

[5] We return to this problem in our discussion of a stochastic symplectic geometry which can be used to bypass this kind of problem.

CHAPTER 7

STOCHASTIC CALCULUS OF VARIATIONS

The embedding procedure allows us to associate a stochastic Euler-Lagrange equation to a stochastic Lagrangian function. A basic question is then the existence of an analogue of the *least action principle*. In this section, we develop a stochastic calculus of variations for our Lagrangian function following a previous work of K. Yasue [**71**]. Our main result, called the coherence lemma, states that the stochastic Euler-Lagrange equation can be obtained as an application of a stochastic least action principle. Moreover, this derivation is consistent with the stochastic embedding procedure.

7.1. Functional and L-adapted process

In the sequel we denote by I a given open interval (a, b), $a < b$.

We first define the stochastic analogue of the classical functional.

Definition 7.1. — *Let L be an admissible Lagrangian function. The functional associated to L is defined by*

$$J_{a,b}(X) = E\left[\int_a^b L(X(t), \mathcal{D}_\mu X(t))dt\right], \tag{7.1}$$

for all $X \in \mathcal{C}^1(I)$.

In what follows, we need a special notion introduced by Yasue [**71**], and called L-adaptation:

Definition 7.2. — *Let $X \in \mathcal{C}^1(I)$ be a stochastic process. We denote by $\mathcal{P}$ and $\mathcal{F}$ the past and the future of X. Let L be an admissible Lagrangian function. A process $X \in \mathcal{C}^1(I)$ is called L-adapted if:*

i) $\dfrac{\partial L}{\partial v}(X(t), \mathcal{D}_\mu X(t))$ *is adapted to $\mathcal{P}$ and $\mathcal{F}$.*

ii) $\dfrac{\partial L}{\partial v}(X(t), \mathcal{D}_\mu X(t)) \in \mathcal{C}^1(I)$.

Diffusion processes are L-adapted.

7.2. Space of variations

Calculus of variations is concerned with the behaviour of functionals under *variations* of the underlying functional space, i.e. objects of the form $\gamma + h$, where γ belongs to the functional space and h is a given functional space of variations. A special care must be taken in our case to define what is the class of *variations* we are considering. In general, this problem is not really pointed out as both variations and curves can be taken in the same functional space (see [**5**],p.56,footnote 26). We introduce the following terminology:

Definition 7.3. — *Let P be a subspace of $\mathcal{C}^1(I)$ and $X \in \mathcal{C}^1(I)$. A P-variation of X is a stochastic process of the form $X + Z$, where $Z \in P$.*

In the sequel, we consider two subspaces of variations: $\mathcal{N}^1(I)$ and $\mathcal{C}^1(I)$.

The choice of $\mathcal{C}^1(I)$ is natural. However, doing this we can obtain stochastic processes with completely different behaviour than X[(1)].

What is the specific property of $X \in \mathcal{C}^1(I)$ that we want to keep ?

If we refer to the construction of the stochastic derivative, then a main point is the existence of an *imaginary part* in $\mathcal{D}_\mu X$[(2)]. This property is related to the non-differentiability of the underlying stochastic process. We are then lead to search for variations Z which conserve this imaginary part. As a consequence, we must consider Nelson difference processes introduced in the previous part[(3)], and denoted by $\mathcal{N}^1(I)$.

7.3. Differentiable functional and stationary processes

We now define our notion of *differentiable functional*. Let P be a subspace of $\mathcal{C}^1(I)$.

(1) Of course, this is not the case in the classical case: one consider $x \in C^\infty(I)$ and $z \in C^\infty(I)$ such that $x + h \in C^\infty(I)$ is very similar to x. For example, we don't choose $z \in C^0(I)$ which leads to radically new behaviour of $x + z$ with respect to x.

(2) Of course, as long as $\mu = \pm 1$. This is of importance since we will be able to choose a more general variations space in this case.

(3) An analogous problem is considered in [**14**], where a non differentiable variational principle is defined.

Definition 7.4. — *Let L be an admissible Lagrangian function and $J_{a,b}$ the associated functional. The functional $J_{a,b}$ is called P-differentiable at an L-adapted process $X \in \mathcal{C}^1(I)$ if*

$$J_{a,b}(X+Z) - J_{a,b}(X) = dJ_{a,b}(X,Z) + R(X,Z), \tag{7.2}$$

where $dJ_{a,b}(X,Z)$ is a linear functional of $Z \in P$ and $R(X,Z) = o(\| Z \|)$.

The stochastic analogue of a stationary point is then defined by:

Definition 7.5. — *A P-stationary process for the functional $J_{a,b}$ is a stochastic process $X \in \mathcal{C}^1(I)$ such that $dJ(X,Z) = 0$ for all $Z \in P$.*

7.3.1. The $P = \mathcal{C}^1(I)$ case. — Our main result is:

Lemma 7.1. — *The functional J_{ab} defined by (7.1) is $\mathcal{C}^1(I)$-differentiable at any L-adapted process $X \in \mathcal{C}^1(I)$, and for all $Z \in \mathcal{C}^1(I)$, the differential is given by:*

(7.3)
$$\begin{aligned} dJ_{ab}(X,Z) &= E\left[\int_a^b \left[\frac{\partial L}{\partial x}(X(u),\mathcal{D}_\mu X(u)) - \mathcal{D}_{-\mu}\left(\frac{\partial L}{\partial v}(X(u),\mathcal{D}_\mu X(u))\right)\right] Z(u)du\right] \\ &\quad + g(Z,\partial_v L)(b) - g(Z,\partial_v L)(a), \end{aligned}$$

where

$$g(Z,\partial_v L)(s) = E\left[Z(u)\partial_v L(X(u),\mathcal{D}_\mu X(u))\right]. \tag{7.4}$$

Proof. — Let X and Z be two L-adapted processes. The Taylor expansion of L gives:

$$\begin{aligned} L(X+Z,\mathcal{D}_\mu(X+Z)) - L(X,\mathcal{D}_\mu(X)) &= \partial_x L(X,\mathcal{D}_\mu(X))Z \\ &\quad + \partial_v L(X,\mathcal{D}_\mu(X))\mathcal{D}_\mu(Z) \\ &\quad + o(\|Z\|), \end{aligned} \tag{7.5}$$

which yields (7.6) by integration and (3.9). □

7.3.2. The $P = \mathcal{N}^1(I)$ case. — Our main result is:

Lemma 7.2. — *The functional J_{ab} defined by (7.1) is $\mathcal{N}^1(I)$-differentiable at any L-adapted process $X \in \mathcal{C}^1(I)$, and for all $Z \in \mathcal{N}^1(I)$ the differential is given by:*

$$\begin{aligned} dJ_{ab}(X,Z) &= E\left[\int_a^b (\partial_x L - \mathcal{D}_\mu \partial_v L)(X(u),\mathcal{D}_\mu X(u))Z(u)du\right] \\ &\quad + g(Z,\partial_v L)(b) - g(Z,\partial_v L)(a), \end{aligned} \tag{7.6}$$

where

$$g(Z,\partial_v L)(s) = E\left[Z(u)\partial_v L(X(u),\mathcal{D}_\mu X(u))\right]. \tag{7.7}$$

Proof. — Let X a $L-$adapted process and H a Nelson-differentiable process. The Taylor expansion of L gives

$$\begin{array}{rcl} L(X+H,\mathcal{D}_\mu(X+H))-L(X,\mathcal{D}_\mu(X)) & = & \partial_x L(X,\mathcal{D}_\mu(X))H \\ & & +\partial_V L(X,\mathcal{D}_\mu(X))\mathcal{D}_\mu(H) \\ & & +o(\|H\|), \end{array} \tag{7.8}$$

which yields (7.6) by integration and (3.18). □

7.4. A technical lemma

The classical derivation of the least action principle used a well known result about *bump functions* (see [**5**],p.57). In the stochastic framework, we will need the following result:

Lemma 7.3. — *Let $Y \in \mathcal{P}_{\mathbb{C}}$ be a complex stochastic process. If Y satisfies*

$$\int_0^1 E\left[Y(u)\mathcal{D}_\mu Z(u)\right]\,du = 0, \tag{7.9}$$

for all $Z \in \mathcal{N}^1([0,1])$ then Y is a constant process.

Proof. — We denote $Y = Y_1 + iY_2$, where $Y_i \in \mathcal{P}_{\mathbb{R}}$ and $\mathcal{D}_\mu Z = A$, where $A \in \mathcal{P}_{\mathbb{R}}$. The equation (7.9) is equivalent to

$$\begin{array}{rcl} \int_0^1 E\left[Y_1(u)A(u)\right]\,du & = & 0, \\ \int_0^1 E\left[Y_2(u)A(u)\right]\,du & = & 0, \end{array} \tag{7.10}$$

for all $A \in \mathcal{P}_{\mathbb{R}}$ such that there exists $Z \in \mathcal{C}^1([0,1])$ satisfying $\mathcal{D}_\mu Z = A$.

Let Z_{Y_1} be the process defined by

$$Z_{Y_1}(u) = \int_0^u Y_1(s)ds - u\int_0^1 Y_1(s)ds. \tag{7.11}$$

We have $Z_{Y_1} \in \mathcal{N}^1(I)$ with $Z(0) = Z(1) = 0$. Indeed, we have

$$\mathcal{D}_\mu Z(u) = Y_1(u) - \int_0^1 Y_1(s)\,ds. \tag{7.12}$$

As a consequence, we have in our notations $B = 0$ and the first equation of (7.10) reduces to

$$\int_0^1 E[Y_1(u)A(u)]du = E\left[\int_0^1 \left(Y_1(u) - \int_0^1 Y_1(s)ds\right)^2 du\right]. \tag{7.13}$$

We deduce that Y_1 is a constant process, that is for all $u \in I$, $Y_1(u) = C$ a.s., where C is a random variable.

The same argument with the second equation of (7.10) and Z_{Y_2} concludes the proof of the lemma. □

7.5. Least action principles

As for the computation of the differential of functionals, we must consider two cases: $P = \mathcal{C}^1(I)$ and $P = \mathcal{N}^1(I)$.

7.5.1. The $P = \mathcal{C}^1(I)$ case. — The main result of this section is the following analogue of the least-action principle for Lagrangian mechanics.

***Theorem 7.1* (Global Least action principle).** — *A necessary and sufficient condition for an L-adapted process to be a $\mathcal{C}^1(I)$-stationary process of the functional J_{ab} with fixed end points $X(a) := X_a \in H$ et $X(b) := X_b \in H$ is that it satisfies*

$$\frac{\partial L}{\partial x}(X(t), \mathcal{D}_\mu X(t)) - \mathcal{D}_{-\mu}\left[\frac{\partial L}{\partial v}(X(t), \mathcal{D}_\mu X(t))\right] = 0. \tag{7.14}$$

We call this equation the Global Stochastic Euler-Lagrange equation (GSEL).

We have conserved the terminology of least-action principle even if we have no notion of *extremals* for our complex valued functional.

Proof. — We denote by $I =]0,1[$. Let $X \in \mathcal{C}^1(I)$ be a solution of

$$(\partial_x \mathcal{L} - \mathcal{D}_\mu \partial_v L)(X(u), \mathcal{D}_\mu X(u)) = 0, \tag{7.15}$$

then X is a $\mathcal{N}^1(I)$-stationary process for the functional J_I.

Conversely, let X is a $\mathcal{C}^1(I)$-stationary process for the functional J_I, i.e. $dJ_I(X, Z) = 0$. Writing

$$(\partial_x \mathcal{L} - \mathcal{D}_\mu \partial_v L)(X(u), \mathcal{D}_\mu X(u)) = \mathcal{D}_\mu Y(u),$$

where

$$Y(u) = \int_0^u \partial_x L(X(s), \mathcal{D}_\mu X(s))ds - \partial_v L(X(u), \mathcal{D}_\mu X(u)), \tag{7.16}$$

we obtain for any $Z \in \mathcal{C}^1(I)$ with $Z(0) = Z(1) = 0$:

$$\begin{aligned} dJ_I(X, Z) &= E\left[\int_0^1 \mathcal{D}_\mu Y(u) Z(u) du\right] \\ &= \int_0^1 E[\mathcal{D}_\mu Y(u) Z(u)] du. \end{aligned} \tag{7.17}$$

Using the $\mathcal{C}^1(I)$-product rule (see equation 3.9), we obtain

$$dJ_I(X, Z) = -\int_0^1 E[Y(u)\mathcal{D}_\mu Z(u)] du. \tag{7.18}$$

Using lemma 7.3 we obtain that Y is a constant process.

Hence, we have $\mathcal{D}_\mu Y(u) = 0$ and

$$(\partial_x \mathcal{L} - \mathcal{D}_\mu \partial_v L)(X(u), \mathcal{D}_\mu X(u)) = 0, \tag{7.19}$$

which concludes the proof. □

7.5.2. The $P = \mathcal{N}^1(I)$ case. — Our main result is:

***Theorem 7.2* (least action principle).** — *A necessary and sufficient condition for an L-adapted process to be a $\mathcal{N}^1(I)$-stationary process of the functional J_{ab} with fixed end points $X(a) := X_a \in H$ et $X(b) := X_b \in H$ is that it satisfies*

$$\frac{\partial L}{\partial x}(X(t), \mathcal{D}_\mu X(t)) - \mathcal{D}_\mu \left[\frac{\partial L}{\partial v}(X(t), \mathcal{D}_\mu X(t)) \right] = 0. \tag{7.20}$$

We call this equation the weak stochastic Euler-Lagrange equation (SEL).

Proof. — We denote by $I =]0,1[$. Let $X \in \mathcal{C}^1(I)$ be a solution of

$$(\partial_x \mathcal{L} - \mathcal{D}_\mu \partial_v L)(X(u), \mathcal{D}_\mu X(u)) = 0, \tag{7.21}$$

then X is a $\mathcal{N}^1(I)$-stationary process for the functional J_I.

Conversely, let X is a $\mathcal{N}^1(I)$-stationary process for the functional J_I, i.e. $dJ_I(X, Z) = 0$. Writing

$$(\partial_x \mathcal{L} - \mathcal{D}_\mu \partial_v L)(X(u), \mathcal{D}_\mu X(u)) = \mathcal{D}_\mu Y(u),$$

where

$$Y(u) = \int_0^u \partial_x L(X(s), \mathcal{D}_\mu X(s)) ds - \partial_v L(X(u), \mathcal{D}_\mu X(u)), \tag{7.22}$$

we obtain for any $Z \in \mathcal{N}^1(I)$ with $Z(0) = Z(1) = 0$:

$$\begin{aligned} dJ_I(X, Z) &= E\left[\int_0^1 \mathcal{D}_\mu Y(u) Z(u) du \right] \\ &= \int_0^1 E[\mathcal{D}_\mu Y(u) Z(u)] du. \end{aligned} \tag{7.23}$$

Using the $\mathcal{N}^1(I)$-product rule (see equation 3.18), we obtain

$$dJ_I(X, Z) = -\int_0^1 E[Y(u) \mathcal{D}_\mu Z(u)] du. \tag{7.24}$$

Using lemma 7.3, we deduce that Y is a constant process, that is for all $u \in I$, $Y(u) = C$ a.s. where C is a random variable.

Hence, we obtain $\mathcal{D}_\mu Y(u) = 0$ and

$$(\partial_x \mathcal{L} - \mathcal{D}_\mu \partial_v L)(X(u), \mathcal{D}_\mu X(u)) = 0, \tag{7.25}$$

which concludes the proof. □

7.6. The coherence lemma

It is not clear that the stochastic Euler-lagrange equation obtained by the stochastization procedure and the $\mathcal{N}^1(I)$ or $\mathcal{C}^1(I)$ least-action principle coincide. One easily sees that this is not the case for $P = \mathcal{C}^1(I)$. In the contrary, we have the following lemma, called the *coherence lemma*, which ensure that for $P = \mathcal{N}^1(I)$ we obtain the same equations.

***Lemma 7.4* (coherence lemma).** — *The following diagram commutes :*

$$\begin{array}{ccc} L(x(t), x'(t)) & \xrightarrow{\mathcal{S}} & L(X(t), \mathcal{D}X(t)) \\ \text{Least action principle}\downarrow & & \downarrow\text{Stochastic Least action principle} \\ (EL) & \xrightarrow[\mathcal{S}]{} & (SEL) \end{array} \tag{7.26}$$

Proof. — This is an immediate consequence of the previous results. □

Remark 7.1. — *When $\mu = 0$, i.e. in the reversible case, the previous lemmas and theorems are true under $\mathcal{C}^1(I)$ variations. Note that when $\mu = 0$, our stochastic derivatives coincides with the Misawa-Yasue* [**52**] *canonical formalism for stochastic mechanics.*

CHAPTER 8

THE STOCHASTIC NOETHER THEOREM

A natural question arising from the stochastization procedure of classical dynamical systems, in particular, Lagrangian systems, is to understand what remains from classical first integrals of motion. First integrals play a central role in many problems like the n-body problem. In this section, we obtain a stochastic analogue of the Noether theorem. We then defined the notion of first integrals for stochastic dynamical systems. We also discuss the consequences of the existence of first integrals in the context of chaotic dynamical systems.

8.1. Tangent vector to a stochastic process

Let $X \in \mathcal{C}^1(I)$ be a stochastic process. We define the analogue of a tangent vector to X at point t.

Definition 8.1. — *Let $X \in \mathcal{C}^1(I)$, $I \subset \mathbb{R}$. The tangent vector to X at point t is the random variable $\mathcal{D}X(t)$.*

Remark 8.1. — *Of course, in order to define stochastic Lagrangian systems in an intrinsic way, one must define the stochastic analogue of the tangent bundle to a smooth manifold. In our case, it is not clear what is the adequate geometric object underlying stochastic Lagrangian dynamics. For example, we can think of multidimensional Brownian surfaces (*[**23**],§*.16.4). All these questions will be developed in a forthcoming paper* [**17**].

8.2. Canonical tangent map

In the sequel, we will need the following mapping called the *canonical tangent map*:

Definition 8.2. — *For all $X \in \mathcal{C}^1(I)$, we define the canonical tangent map as*

$$T : \begin{array}{ccc} \mathcal{C}^1(I) & \longrightarrow & \mathcal{C}^1(I) \times \mathcal{P}_{\mathbb{C}}, \\ X & \longmapsto & (X, \mathcal{D}X). \end{array} \tag{8.1}$$

The mapping T will be used in the following section to define the analogue of the linear tangent map for a stochastic suspension of a one parameter group of diffeomorphisms.

8.3. Stochastic suspension of one parameter family of diffeomorphisms

We begin by introducing a useful notion of *stochastic suspension* of a diffeomorphism.

Definition 8.3. — *Let $\phi : \mathbb{R}^n \to \mathbb{R}^n$ be a diffeomorphism. The stochastic suspension of ϕ is the mapping $\Phi : \mathcal{P} \to \mathcal{P}$ defined by*

$$\forall X \in \mathcal{P},\ \Phi(X)_t(\omega) = \phi(X_t(\omega)). \tag{8.2}$$

In what follows, we will frequently use the same notation for the suspension of a given diffeomorphism and the diffeomorphism.

Remark 8.2. — *It seems strange that we have not defined directly the notion of diffeomorphism on a subset E of the stochastic processes, i.e. mapping $\Phi : E \to E$ which are Fréchet differentiable with an inverse which is also Fréchet differentiable. However, these objects do not always exist.*

Using the stochastic suspension, we are able to define the notion of stochastic suspension for a one-parameter group of diffeomorphisms.

Definition 8.4. — *A one-parameter group of transformations $\Phi_s : A \to A$, $s \in \mathbb{R}$, where $A \subset \mathcal{P}$, is called a ϕ-suspension group acting on A if there exist a one parameter group of diffeomorphisms $\phi_s : \mathbb{R}^n \to \mathbb{R}^n$, $s \in \mathbb{R}$, such that for all $s \in \mathbb{R}$, we have:*

i) Φ_s is the stochastic suspension of ϕ_s,
ii) for all $X \in A$, $\Phi_s(X) \in A$.

This notion of suspension group comes from our framework. It relies on the fact that we want to understand how symmetries of the underlying Lagrangian systems are transported via the stochastic embedding. The non-trivial condition on the stochastic suspension of a one-parameter group of diffeomorphisms acting on A comes from condition ii). However, imposing some conditions on the underlying one parameter group, we can obtain a stochastic one parameter group which acts on the set E of good diffusion processes.

Precisely, let us introduce the following class of one-parameter groups:

Lemma 8.1. — *An admissible one parameter group of diffeomorphisms $\Phi = \{\phi_s\}_{s\in\mathbb{R}}$ is a one parameter group of C^2-diffeomorphisms on $\mathbb{R}^n$ such that*

$$(s,x) \mapsto \frac{\partial}{\partial x}\phi_s(x) \text{ is } C^2. \tag{8.3}$$

The main property of admissible one parameter groups is the fact that they are well-behaved on the set of good diffusions.

Lemma 8.2. — *Let $\Phi = (\phi_s)_{s\in\mathbb{R}}$ be a stochastic suspension of an admissible one parameter group of diffeomorphisms. Then, for all $X \in E$, we have for all $t \in I$, and all $s \in \mathbb{R}$:*

i) The mapping $s \mapsto D_\mu \Phi_s X(t) \in C^1(\mathbb{R})$, (a.s.),

ii) We have $\dfrac{\partial}{\partial s}[\mathcal{D}_\mu(\phi_s(X))] = \mathcal{D}_\mu\left[\dfrac{\partial \phi_s(X)}{\partial s}\right]$ $(a.s.)$.

This lemma is trivial in the classical case where X is a smooth function and $\mathcal{D}_\mu$ is the classical derivative with respect to time. Indeed, it reduces to the Schwarz lemma. However, this inequality plays an essential role in the derivation of the classical Noether's theorem (see [**5**],p.89).

Proof. — According to (2.4),

$$\mathcal{D}_\mu \phi_s(X)(t) = \mathcal{D}_\mu X(t) \cdot \frac{\partial_x \phi_s}{\partial x} X(t) + i\mu \frac{\sigma(t,X_t)^2}{2} \frac{\partial_x^2 \phi_s}{\partial x^2} X(t) \quad (a.s.).$$

So :

$$\frac{\partial}{\partial s}\mathcal{D}_\mu \phi_s(X)(t) = \mathcal{D}_\mu X(t) \cdot \frac{\partial}{\partial s}\frac{\partial_x \phi_s}{\partial x} X(t) + i\mu \frac{\sigma(t,X_t)^2}{2} \frac{\partial}{\partial s}\frac{\partial_x^2 \phi_s}{\partial x^2} X(t) \quad (a.s.).$$

Since $(s,x) \mapsto \phi_s(x)$ is C^2, we have $\dfrac{\partial}{\partial s}\dfrac{\partial}{\partial x}\phi_s(x) = \dfrac{\partial}{\partial x}\dfrac{\partial}{\partial s}\phi_s(x)$ by the Schwarz lemma. In the same way,

$$\frac{\partial}{\partial s}\frac{\partial^2}{\partial x^2}\phi_s(x) = \frac{\partial}{\partial x}\frac{\partial}{\partial s}\frac{\partial}{\partial x}\phi_s(x)$$

because $(s,x) \mapsto \frac{\partial}{\partial x}\phi_s(x)$ is C^2.
Therefore :

$$\frac{\partial}{\partial s}\frac{\partial^2}{\partial x^2}\phi_s(x) = \frac{\partial^2}{\partial x^2}\frac{\partial}{\partial s}\phi_s(x).$$

Applying (2.4) to $\dfrac{\partial}{\partial s}\phi_s$, we can conclude that :

$$\frac{\partial}{\partial s}[\mathcal{D}_\mu(\phi_s(X))] = \mathcal{D}_\mu\left[\frac{\partial \phi_s(X)}{\partial s}\right] \quad (a.s.).$$

□

It must be pointed out that every extension of this lemma will lead to a substantial improvement of the following stochastic Noether theorem.

8.4. Linear tangent map

Let $X \in \mathcal{C}^1(I)$ and $\phi : \mathbb{R}^n \to \mathbb{R}^n$ be a diffeomorphism. The image of X under the stochastic suspension of ϕ, denoted by Φ, induces a natural map for tangent vectors denoted by Φ_*, called the *linear tangent map*, and defined as in classical differential geometry by:

Definition 8.5. — *Let Φ be a stochastic suspension of a diffeomorphism. The linear tangent map associated to Φ, and denoted by Φ_*, is defined for all $X \in \mathcal{C}^1(I)$ by*

$$\Phi_*(X) = T(\Phi(X)) = (\Phi(X), \mathcal{D}(\Phi(X))). \tag{8.4}$$

All the quantities are well defined as diffeomorphisms send $\mathcal{C}^1(I)$ on $\mathcal{C}^1(I)$.

8.5. Invariance

We then obtain the following notion of *invariance* under a one parameter group of diffeomorphisms.

Definition 8.6. — *Let $\Phi = \{\phi_s\}_{s\in\mathbb{R}}$ be a one-parameter group of diffeomorphisms and let L be a functional $L : \mathcal{C}^1(I) \to \mathcal{C}^1_{\mathbb{C}}(I)$. The functional L is invariant under the one-parameter group of diffeomorphisms Φ if*

$$L(\phi_* X) = L(X), \text{ for all } \phi \in \Phi.$$

As a consequence, if L is invariant under Φ, we have

$$L(\phi_s(X), \mathcal{D}(\phi_s(X))) = L(X, \mathcal{D}X),$$

for all $s \in \mathbb{R}$ and $X \in \mathcal{C}^1(I)$.

Remark 8.3. — *We note that this notion of invariance under a one parameter group of diffeomorphisms does not coincide with the same notion as defined by K. Yasue ([**71**], p.332, formula (3.1)) which in our notation is given by:*

$$L(\phi_s(X), \phi_s(\mathcal{D}X)) = L(X, \mathcal{D}X), \text{ for all } s \in \mathbb{R} \text{ and } X \in \mathcal{C}^1(I).$$

*In fact, K. Yasue definition of invariance does not reduce to the classical notion (see for example [**5**],p.88) for differentiable deterministic stochastic processes.*

*Moreover, Yasue's definition is not coherent with the invariance notion used in his proof of the stochastic Noether's theorem ([**71**],theorem 4,p.332). See the comment below.*

8.6. The stochastic Noether's theorem

Noether's theorem has already been generalized a great number of times and covers sometimes different statements [**32**]. Here, we follow V.I. Arnold's ([**5**],p.88) presentation of the Noether theorem for Lagrangian systems. We correct a previous work of K. Yasue ([**71**], Theorem 4,p. 332-333).

Theorem 8.1. — *Let $J_{a,b}$ be a functional on $\mathcal{C}^1(I)$ given by*

$$J_{a,b}(X) = E\left[\int_a^b L(X(t), \mathcal{D}X(t))dt\right].$$

with L invariant under the one-parameter group $\Phi = \{\phi_s\}_{s\in\mathbb{R}}$.

Let $X \in \mathcal{C}^1(I)$ be a $\mathcal{C}^1(I)$-stationary point of $J_{a,b}$ with fixed end points condition

$$X(a) = X_a, \text{ and } X(b) = X_b.$$

Then, we have

$$\frac{d}{dt}E\left[\mathrm{grad}_v L\,\frac{\partial Y}{\partial s}\bigg|_{s=0}\right] = 0,$$

where

$$Y_s = \Phi_s(X). \tag{8.5}$$

Proof. — Let $Y(s,t) = \phi_s X(t)$ for $s \in \mathbb{R}$ and $a \leqslant t \leqslant b$.

As L is invariant under $\Phi = \{\phi_s\}_{s\in\mathbb{R}}$, we have

$$\frac{\partial}{\partial s}L(Y(s,t), \mathcal{D}_\mu Y(s,t)) = 0 \quad (a.s.).$$

As $Y(.,t)$ and $\mathcal{D}_\mu Y(.,t)$ belong to $C^1(\mathbb{R})$ for all $t \in [a,b]$ by definition 8.4, iii), we obtain

$$\mathrm{grad}_x L \cdot \frac{\partial Y}{\partial s} + \mathrm{grad}_v L \frac{\partial \mathcal{D}_\mu Y}{\partial s} = 0 \quad (a.s.). \tag{8.6}$$

Using (Lemma 8.2,ii), this equation is equivalent to

$$\mathrm{grad}_x L \cdot \frac{\partial Y}{\partial s} + \mathrm{grad}_v L \mathcal{D}_\mu \left(\frac{\partial Y}{\partial s}\right) = 0 \quad (a.s.). \tag{8.7}$$

As $X = Y\,|_{s=0}$ is a stationary process for $J_{a,b}$, we have

$$\mathrm{grad}_x L = \mathcal{D}_{-\mu}\mathrm{grad}_v L. \tag{8.8}$$

As a consequence, we deduce that

$$\left([\mathcal{D}_\mu \mathrm{grad}_v L] \cdot \frac{\partial Y}{\partial s} + \mathrm{grad}_v L \mathcal{D}_\mu \left(\frac{\partial Y}{\partial s}\right)\right)\bigg|_{s=0} = 0 \quad (a.s.).$$

Taking the absolute expectation, we obtain

$$E\left[\left([\mathcal{D}_\mu \mathrm{grad}_v L] \cdot \frac{\partial Y}{\partial s} + \mathrm{grad}_v L \mathcal{D}_\mu \left(\frac{\partial Y}{\partial s}\right)\right)\bigg|_{s=0}\right] = 0. \tag{8.9}$$

Using the product rule, we obtain

$$\frac{d}{dt}E\left[\mathrm{grad}_v L\,\frac{\partial Y}{\partial s}\bigg|_{s=0}\right]=0,$$

which concludes the proof. □

8.7. Stochastic first integrals

The previous theorem leads us to the introduction of the notion of *first integral* for stochastic Lagrangian systems[(1)].

8.7.1. Reminder about first integrals. — Let X be a C^k vector field or $\mathbb{R}^n$, $k \geqslant 1$ (k could be ∞ or ω, i.e. analytic). We denote by $\phi_x(t)$ the solution of the associated differential equation, such that $\phi_x(0) = x$ and by S the set of all these solutions.

A first integral of X is a real valued function $f : \mathbb{R}^n \to \mathbb{R}$ such that for all $\phi_x(t) \in S$, we have

$$f(\phi_x(t)) = c_x, \tag{8.10}$$

where c_x is a constant.

We have not imposed any kind of regularity on the function f, so that f can be just C^0. In this case, the existence of a first integral does not impose many constraint on the dynamics.

If f is at least C^1, then we can characterize first integrals by the following constraint:

$$X \cdot f = 0. \tag{8.11}$$

8.7.2. Stochastic first integrals. — The previous paragraph leads us to searching for an analogue of the classical notion of first integrals as a functional defined on the set of solutions of a given stochastic Euler-Lagrange equation[(2)] and real valued. Looking for the stochastic Noether theorem, we choose the following definition:

Definition 8.7. — *Let L be an admissible Lagrangian system. A functional $I : \mathcal{C}^1(I) \to \mathbb{R}$ is a first integral for the Euler-Lagrange equation associated to L if*

$$\frac{d}{dt}\left[I(X)\right]=0, \tag{8.12}$$

for all X satisfying the Euler-Lagrange equation.

(1) Of course, one can extend this definition to general stochastic dynamical systems.
(2) Of course, this definition will extend to arbitrary stochastic dynamical systems.

We can now interpret the stochastic Noether theorem in term of first integrals, i.e. the fact that the invariance of the Lagrangian L under of a one parameter group of diffeomorphisms $\Phi = (\phi_s)_{s\in\mathbb{R}}$ induces the existence of a first integral for the associated Euler-Lagrange equation, defined by

$$I(X) = E\left[\mathrm{grad}_v L \left.\frac{\partial \phi_s X(t)}{\partial s}\right|_{s=0}\right]. \tag{8.13}$$

8.8. Examples

8.8.1. Translations. — We follow the first example given by V.I. Arnold ([**5**],p.89) for Noether theorem. Let L be the Lagrangian defined by

$$L(X,V) = \frac{V^2}{2} - U(X), \text{ where } X \in \mathbb{R}^3, \tag{8.14}$$

$V = (V_1, V_2, V_3) \in \mathbb{C}^3$, $V^2 := V_1^2 + V_2^2 + V_3^2$ and U is taken to be invariant under the one parameter group of translations:

$$\phi_s(x) = x + se_1, \tag{8.15}$$

where $\{e_1, e_2, e_3\}$ is the canonical basis of $\mathbb{R}^3$.

Then, by the Stochastic Noether's theorem, the quantity

$$E[\mathcal{D}X_1] \tag{8.16}$$

is a first integral since $\partial_V L = V$ and $\partial_s \phi_s(X_1(\omega)) = e_1$.

8.8.2. Rotations. — We keep the notations of the previous paragraph. We consider the Lagrangian of the two-body problem in $\mathbb{R}^3$, i.e.

$$L(X,V) = q(V) - \frac{1}{|X|} \text{ where } q(V) = \frac{V^2}{2}, \tag{8.17}$$

where $|\ .\ |$ denotes the classical norm on $\mathbb{R}^3$ defined for all $X \in \mathbb{R}^3$, $X = (X_1, X_2, X_3)$ by $|\ X\ |^2 = X_1^2 + X_2^2 + X_3^2$.

We already know that the classical Lagrangian L is invariant under rotations when $X \in \mathbb{R}^3$ and $V \in \mathbb{R}^3$. Here, we must prove that the same is true for the extended object, i.e. for L defined over $\mathbb{R}^3 \setminus \{0\} \times \mathbb{C}^3$. This extension, as long as it is defined, is canonical. Indeed, we define $q(z)$ for $z \in \mathbb{C}^3$ as

$$q(z) = \frac{1}{2}(z_1^2 + z_2^2 + z_3^2), \quad z = (z_1, z_2, z_3) \in \mathbb{C}^3. \tag{8.18}$$

Note that our problem is not to discuss an analytic extension of the real valued kinetic energy but only to look for the same function on $\mathbb{C}^3$ simply replacing real variables by complex one. As long as the new object is well defined this procedure is canonical, which

is not the case if we search for an analytic extension of q over $\mathbb{C}^3$ which reduces to q on $\mathbb{R}^3$.

Our main result is then that this group of symmetry is preserved under stochastization, which is in fact a general phenomenon that will be discuss elsewhere.

Lemma 8.3. — *The lagrangian L defined over $\mathbb{R}^3 \setminus \{0\} \times \mathbb{C}^3$ is invariant under rotations $\phi_{\theta,k}$ around the e_k axis by the angle θ, $k = 1, 2, 3$.*

The proof is based on the two following facts:

– As $\phi_{\theta,k}$ is a linear map whose matrix coefficients do not depend on t, we have

$$\mathcal{D}_\mu \left[\phi_{\theta,k}(X)\right] = \phi_{\theta,k} \left[\mathcal{D}_\mu X\right], \tag{8.19}$$

where $\phi_{\theta,k}$ is trivially extended to $\mathbb{C}^3$.

– A simple calculation gives

$$\forall\ z \in \mathbb{C}^3, \quad q(\phi_{\theta,k}(z)) = q(z). \tag{8.20}$$

We easily deduce the $\phi_{\theta,k}$ invariance of L, i.e. that

$$L(\phi_{\theta,k}X, \mathcal{D}(\phi_{\theta,k}X)) = L(X, \mathcal{D}X). \tag{8.21}$$

We now compute: $\partial_\theta \phi_{\theta,k}(X)|_{\theta=0} = e_k \wedge X$ and

$$\partial_V L(X, \mathcal{D}X) \cdot \partial_\theta \phi_{\theta,k}(X)|_{\theta=0} = (X \wedge \mathcal{D}X)_k.$$

Therefore the expectation of the "complex angular momentum" $X \wedge \mathcal{D}X$ is a conserved vector ($\wedge$ is extended in a natural way to complex vectors).

8.9. About first integrals and chaotic systems

In this section, we discuss some consequences of the stochastic Noether's theorem in the context of chaotic dynamical systems. The study of deterministic chaotic dynamical systems is difficult.

Here again, we return to the classical n-body problem, $n \geqslant 3$. In this case, in particular for large n, the dynamics of the system is very complicated and only numerical results give a global picture of the phase space. Despite the existence of a chaotic behaviour, there exist several well known first integrals of the system.

These integrals are used as constraints on the dynamics and can give interesting results, as for example J. Laskar's [**41**] approach to the Titus-Bode law for the repartition of the

planets in the solar systems and extra-solar systems.

Using our approach, we can go further by claiming that such kind of integrals continue to exist even if one consider a more general class of perturbations including stochasticity. We note that this result is fundamental as long as one wants to relate numerical computations on the n-body problem with the real dynamical behaviour of the solar systems, and in this particular example, the dynamics of the protoplanetary nebulae.

CHAPTER 9

NATURAL LAGRANGIAN SYSTEMS AND THE SCHRÖDINGER EQUATION

In this section, we explore in details the stochastization procedure for natural Lagrangian systems. In particular, by introducing a suitable analogue of the action functional, we prove that the stochastic Euler-Lagrange equation leads to a non-linear Schrödinger equation, depending on a free parameter related to a normalization constraint. For a suitable choice of this parameter we then obtain the classical linear Schrödinger equation.

9.1. Natural Lagrangian systems

In ([**5**],p.84), V.I. Arnold introduces the following notion of *natural Lagrangian systems*:

Definition 9.1. — *A Lagrangian system is called natural if the Lagrangian function is equal to the difference between kinetic and potential energy:*

$$L(x, v) = T(v) - U(x).$$

As an example, we have the natural Lagrangian function associated to *Newtonian mechanics*:

$$L(x, v) = \frac{1}{2}v^2 - U(x),$$

where U is of class C^∞.

9.2. Schrödinger equations

9.2.1. Some notations and a reminder of the Nelson wave function. — We recall that Λ_d is the space of "good" diffusion processes. Let Λ_d^g be the subspace of Λ_d whose elements have a smooth gradient drift. We then set:

$$\mathcal{S} = \{X \in \Lambda_d \mid \mathcal{D}^2 X(t) = -\nabla U(X(t))\}.$$

For a diffusion X in Λ_d with drift b and density function $p_t(x)$, we set:

$$\Theta = (\mathbb{R}^+ \times \mathbb{R}^d) \setminus \{(t,x), \mid p_t(x) = 0\}. \tag{9.1}$$

If $X \in \Lambda_d^g$ then there exist real valued functions R and S smooth on Θ such that

$$\mathcal{D}X(t) = \left(b - \frac{\sigma^2}{2}\nabla \log(p_t) + i\frac{\sigma^2}{2}\nabla \log(p_t)\right)(X(t)) = (\nabla S + i\nabla R)(X(t)), \tag{9.2}$$

since b is a gradient. Obviously:

$$R(t,x) = \frac{\sigma^2}{2}\log(p_t(x)). \tag{9.3}$$

In this case, we introduce the function:

$$\Psi(t,x) = e^{\dfrac{(R+iS)(t,x)}{K}} \tag{9.4}$$

(where K is a positive constant) called the wave function.

The wave function has the same form than that of Nelson one (see [**53**]). We then set $A = S - iR$. So $\Psi = e^{\frac{iA}{K}}$ and $\nabla A(t, X(t)) = \overline{\mathcal{D}}X(t)$. For a suitable K, Nelson shows that if X satisfies its stochastized Newton equation (which is the real part of ours) then Ψ satisfies a Schrödinger equation. We show, by using our operator $\mathcal{D}$, the same kind of result in the next section.

9.2.2. Schrödinger equations as necessary conditions. —

Theorem 9.1. — *If $X \in \mathcal{S} \cap \Lambda_d^g$, then the wave function (9.4) satisfies the following non-linear Schrödinger equation on the set Θ:*

$$iK\partial_t\Psi + \frac{K(K-\sigma^2)}{2}\frac{(\partial_x\Psi)^2}{\Psi} + \frac{\sigma^2}{2}\Delta\Psi = U\Psi, \tag{9.5}$$

Proof. — As U is a real valued function, $X \in \mathcal{S}$ implies

$$\overline{\mathcal{D}}^2 X(t) = -\nabla U(X(t)).$$

The definition of Ψ implies that on Θ

$$\nabla A = -iK\frac{\nabla\Psi}{\Psi}.$$

Since $\nabla A(t, X(t)) = \overline{(\mathcal{D}X)(t)}$, we obtain

$$iK\overline{\mathcal{D}}\frac{\partial_x\Psi}{\Psi}(t, X(t)) = \nabla U(t, X(t)).$$

Therefore, considering the k-th component of the last equation and using lemma 2.4, we deduce

$$iK\left(\partial_t\frac{\partial_k\Psi}{\Psi} + \overline{\mathcal{D}}X(t)\cdot\nabla\frac{\partial_k\Psi}{\Psi} - i\frac{\sigma^2}{2}\Delta\frac{\partial_k\Psi}{\Psi}\right)(t, X(t)) = \partial_k U(X(t)).$$

Now $\overline{\mathcal{D}}X(t) = -iK\dfrac{\nabla\Psi}{\Psi}(t, X(t))$. Thus, by Schwarz lemma, we obtain

$$\overline{\mathcal{D}}X(t)\cdot\nabla\frac{\partial_k\Psi}{\Psi} = -iK\sum_{j=1}^{d}\frac{\partial_j\Psi}{\Psi}\partial_j\frac{\partial_k\Psi}{\Psi} = -\frac{iK}{2}\partial_k\sum_{j=1}^{d}\left(\frac{\partial_j\Psi}{\Psi}\right)^2,$$

and

$$\Delta\frac{\partial_k\Psi}{\Psi} = \sum_{j=1}^{d}\partial_j^2\frac{\partial_k\Psi}{\Psi} = \partial_k\sum_{j=1}^{d}\partial_j\left(\frac{\partial_j\Psi}{\Psi}\right) = \partial_k\sum_{j=1}^{d}\frac{\partial_j^2\Psi}{\Psi} - \left(\frac{\partial_j\Psi}{\Psi}\right)^2.$$

Therefore

$$iK\partial_k\left(\frac{\partial_t\Psi}{\Psi} + i\frac{\sigma^2 - K}{2}\partial_k\sum_{j=1}^{d}\left(\frac{\partial_j\Psi}{\Psi}\right)^2 - i\frac{\sigma^2}{2}\frac{\Delta\Psi}{\Psi}\right)(t, X(t)) = \partial_k U(X(t)).$$

By adding an appropriate function of t in S, we can arrange the constant in x of integration in equation to be zero, and formula (9.5) follows as claimed. □

In order to recover the classical linear Schrödinger equation, we must choose the normalization constant K. The main point is that in this case, we obtain a clear relation between the modulus of the wave function and the density of the underlying diffusion process. Precisely, we have:

Corollary 9.1. — *We keep the notations and assumptions of theorem (9.2.3). We assume that*

$$K = \sigma^2.$$

Then the wave functional Ψ satisfies the linear Schrödinger equation

$$i\sigma^2\partial_t\Psi + \frac{\sigma^4}{2}\Delta\Psi = U\Psi, \tag{9.6}$$

Moreover, if $p_t(x)$ is the density of the process $X(t)$ at point x, then we have

$$(\Psi\overline{\Psi})(t, x) = p_t(x).$$

Proof. — $K = \sigma^2$ kills the non-linearity in equation (9.5) and furthermore

$$\log(\Psi\overline{\Psi}) = \frac{2}{K}R = \frac{2}{\sigma^2}R = \log(p).$$

which concludes the proof. □

9.2.3. Remarks and questions. —

– Obviously $\Lambda_1 \subset \Lambda_1^g$ since b is continuous.

– A natural question is to know if the converse of the corollary of () is true. More precisely, if Ψ satisfies a linear Schrödinger equation, can we construct a process X which belongs to $\mathcal{S} \cap \Lambda_d^g$ and whose density is such that $p_t(x) = |\Psi(t, x)|^2$?

R. Carmona tackled the problematic of the so-called Nelson processes and proved in [**11**] under some conditions the existence of a process X with gradient drift related to Ψ and whose density is such that $p_t(x) = |\Psi(t,x)|^2$. However we do not know if this process belongs to our space of good diffusions processes (which may turn to be a little restrictive class in this case), but we can prove formally, *i.e.* even so assuming that the formulae of the stochastized derivative to a function of the process holds, that X satisfies the Newton stochastized equation. Therefore, this leads one to question the extension of the derivative operator and the way it acts on a large class of processes. This problem will be treated in a forthcoming paper (See [**18**]).

– The fact that a process X satisfies the stochastized Newton equation of Nelson implies $(D^2 - D_*^2)X = 0$ (for the potential U is real). This is a general fact for diffusion with gradient drift. Indeed, we can prove:

Lemma 9.1. — *Let $X \in \Lambda_d$, b its drift and p its density function. Let G_i be the i-th column of the matrix $(G_{ij}) := (\partial_j b_i - \partial_i b_j)$. Then $(D^2 - D_*^2)X = 0$ if and only if for all $t > 0$, $\mathrm{div}(p_t G_i) = 0$.*

Thus, if $X \in \Lambda_d^g$ it is clear that $(D^2 - D_*^2)X = 0$ since the form $\sum b_k \partial_k$ is closed and so $G = 0$. An interesting question is then to know if the converse is true. So we may wonder ourselves if $\mathcal{S} \subset \Lambda_d^g$.

The difficulty relies on the fact that p and b are related via the Fokker-Planck equation, so the condition $\mathrm{div}(p_t G_i) = 0$ may not be the good formulation. However, one could use the work of S. Roelly and M. Thieullen in [**61**] who use an integration by parts via Malliavin Calculus to characterize gradient diffusion, in order to give a positive or negative answer to our question.

– A basic notion in mechanics is that of *action* (see [**5**],p.60). The action associated to a Lagrangian system is in general obtained via the action functional. In our framework, a natural definition for such an action functional is given by:

Definition 9.2. — *Let $\mathcal{A}$ be the functional defined on $[a,b] \times \mathcal{C}^1([a,b])$ by:*

$$\forall t \in [a,b],\ \forall X \in \mathcal{C}^1(I),\ \mathcal{A}(t,X) = E\left[\int_a^t L(X_s, (\mathcal{D}X)_s) ds \,|X_t\right]. \tag{9.7}$$

This functional is called the action functional.

Using this action functional, we have some freedom to define the corresponding "action". The natural one is defined by

$$A_X(t,x) = E\left[\int_a^t L(X_s,(\mathcal{D}X)_s)ds\,|X_t = x\right]. \tag{9.8}$$

Usually, the wave function associated to A_X an denoted by $\tilde{\psi}$ is then defined as

$$\tilde{\psi}_X(t,x) = \exp^{iA_X(t,x)}. \tag{9.9}$$

However, it is not at all clear that such kind of function satisfies the *gradient condition*, i.e. that

$$\nabla A(t,X(t)) = \overline{\mathcal{D}}X(t), \tag{9.10}$$

which is fundamental in our derivation of the Schrödinger equation.

However, the condition 9.10 is equivalent to prove that the real part of $\mathcal{D}X$ is a gradient, which is not at all trivial in dimension greater than two.

9.3. About quantum mechanics

Even if we look for dynamical systems, our work can be used in the context of the so-called *Stochastic mechanics*, developed by Nelson [**53**]. The basic idea is to reexpress quantum mechanics in terms of random trajectories. We refer to [**12**] for a review.

The stochastic embedding theory can be seen as a quantization procedure, i.e. a formal way to go from classical to quantum mechanics. This approach is already different from Nelson's approach, which do not define a rigid procedure to associate to a given equation a stochastic analogue. Moreover, the acceleration defined by Nelson as

$$a(X) = \frac{DD_*(X) + D_*D(X)}{2}, \tag{9.11}$$

is only a particular choice. Many authors have tried to justify this form ([**59**],[**60**]) or to try another one. In our context, the form of the acceleration is fixed and corresponds, as in the usual case, to the second (stochastic) derivative of X. As a consequence, stochastic embeddings can be used to provide a conceptual framework to stochastic mechanics. We refer to [**59**] where a complex valued velocity for a stochastic process is introduced corresponding to the stochastic derivative of X.

However, stochastic mechanics as well as its variants have many drawbacks with respect to the initial wish to describe quantum mechanical behaviours. We refer to [**55**] and [**12**] for details. This is the reason why we will not develop further this topic.

CHAPTER 10

STOCHASTIC HAMILTONIAN SYSTEMS

In this part, we introduce the stochastic pendant of Hamiltonian systems for classical Lagrangian systems. The strategy is first to define the stochastic analogue of the classical momentum. We then define a stochastic Hamiltonian. However, this Hamiltonian is not obtained by the classical stochastic embedding procedure. This is due to the fact that the momentum process is complex valued. As a consequence, we must modify the procedure in order to obtain a coherent picture between the classical formalism and the stochastic one. This leads us to define the stochastic Hamiltonian embedding procedure which reflects in fact the non trivial character of the underlying stochastic symplectic geometry to develop. Having the stochastic Hamiltonian we prove a Hamilton least action principle using our stochastic calculus of variations. We then obtain an analogue of the Lagrangian coherence lemma in this case up to the fact that the underlying stochastic embedding procedure is now the Hamiltonian one.

10.1. Reminder about Hamiltonian systems

We denote by I an open interval (a, b), $a < b$.

Let $L : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ be a convex Lagrangian. The Lagrangian functional over $C^1(\mathbb{R})$ is defined by

$$L : \begin{array}{ccc} C^1(\mathbb{R}) & \longrightarrow & C^1(\mathbb{R}), \\ x & \longmapsto & L(x, \dot{x}, t). \end{array} \tag{10.1}$$

We can associate to L a Hamiltonian function using the *Legendre transformation* ([**5**],p.65). From the functional side, this induces a change of point of view, as the functional is not seen as acting on $x(t)$, which is the so-called *configuration space* of classical mechanics, but on $(x(t), \dot{x}(t))$ which is associated to the *phase-space*. This dichotomy between position and velocities has of course many consequences, one of them being that the system is more symmetric (the symplectic structure).

Definition 10.1. — *Let $L(x,v)$ be an admissible Lagrangian system. For all $x \in C^1$, we denote by*

$$p(x) = \frac{\partial L}{\partial v}(x, \dot{x}), \tag{10.2}$$

the momentum variable.

We now introduce an important class of Lagrangian systems.

Definition 10.2. — *Let $L(x,v)$ be an admissible lagrangian system. The Lagrangian L is said to possess the Legendre property if there exists a function $f : \mathbb{R}^d \to \mathbb{R}^d$, called the Legendre transform, such that*

$$\dot{x} = f(x,p), \tag{10.3}$$

for all $x \in C^1$.

Most classical examples in mechanics possess the Legendre property. This follows from the convexity of L in the second variable (see [**5**],p.61-62).

We can introduce the fundamental object of this section:

Definition 10.3. — *Let L be an admissible Lagrangian system which possesses the Legendre property. The Hamiltonian function associated to L is defined by*

$$H(p,x) = pf(x,p) - L(x, f(x,p)), \tag{10.4}$$

where f is the Legendre transform.

The Hamiltonian function plays a fundamental role in classical mechanics. We introduce the stochastic analogue in the next section.

10.2. The momentum process

A natural stochastic analogue of the momentum variable is defined as follow:

Definition 10.4. — *Let $L(x,v)$ be an admissible Lagrangian system. For all $X \in \mathcal{C}^1(I)$, we define the stochastic process $P(t)$, called the canonical momentum process, by*

$$P(t) = \frac{\partial L}{\partial v}(X(t), \mathcal{D}X(t)). \tag{10.5}$$

This definition can be made more natural using the embedding ι defined from $C^0(I)$ on $\mathcal{P}_{\det}$ and the linear tangent map introduced in chapter 8. Indeed, the momentum process can be viewed as a functional on $X \in \mathcal{C}^1(I)$, $P : \mathcal{C}^1(I) \to \mathcal{P}_{\mathbb{C}}$ defined by (10.5). We have for all $X \in \mathcal{P}^1_{\det} = \iota(C^1(I))$,

$$P(X) = \iota(p(x)), \tag{10.6}$$

where $x \in C^1(I)$ is such that $X = \iota(x)$. As by definition, we have

$$\iota(p(x)) = p(\iota(x)) = p(X). \tag{10.7}$$

As p keeps a sense for $X \in \mathcal{C}^1(I)$, we extend formula (10.7) to $\mathcal{C}^1(I)$ leading to definition 10.4.

If we assume that the Lagrangian possesses the Legendre property, then there exists a Legendre transform f such that for all $x \in C^1$, $\dot{x} = f(x, p)$. We can ask if such a property is conserved for the momentum process. We have:

Lemma 10.1. — *Let $L(x, v)$ be an admissible Lagrangian system possessing the Legendre property. Let f be the Legendre transform associated to L. We have*

$$\mathcal{D}X(t) = f(X, P), \tag{10.8}$$

for all $X \in \mathcal{C}^1(I)$.

We can now define the stochastic Hamiltonian associated to L:

Definition 10.5. — *Let $L(x, v)$ be an admissible Lagrangian system possessing the Legendre property. The stochastic Hamiltonian system associated to L is defined by*

$$H : \begin{array}{ccc} \mathcal{P}_{\mathbb{C}} \times \mathcal{C}^1(I) & \longrightarrow & \mathcal{P}_{\mathbb{C}} \\ (P, X) & \longmapsto & Pf(X, P) - L(X, f(X, P)). \end{array} \tag{10.9}$$

10.3. The Hamiltonian stochastic embedding

As in the previous chapter, we want to use the stochastic embedding procedure to associate a natural stochastic analogue of the Hamiltonian equations. However, we must be careful with such a procedure, as already discussed in chapter 4, §.4.2.2. Indeed, the embedding procedure does not allow us to fix the notion of embedding for *systems* of differential equations. Moreover, we must keep in mind that the principal idea behind the Hamiltonian formalism is to work not in the *configuration space*, i.e. the space of positions, but in the *phase space*, i.e. the space of positions and momenta. As the stochastic speed is by definition complex, this induces a particular choice for the embedding procedure in the case of Hamiltonian differential equations.

Definition 10.6. — *Let $F : \mathbb{R}^d \times \mathbb{C}^d \mapsto \mathbb{C}$ be a holomorphic function, real valued on real arguments. This function defines a real valued functional over $C^1(I) \times C^1(I)$, for I a given open interval of $\mathbb{R}$. The Hamiltonian embedding of the functional F is the functional denoted by F_S, defined on $\mathcal{C}^1(I) \times \mathcal{P}_{\mathbb{C}}(I)$ by H, i.e.*

$$F_S(X, P)(t) = F(X(t), P(t)). \tag{10.10}$$

We denote by S_H the procedure associating the stochastic functional F_S to F. This procedure reduces to change the functional spaces for F from $C^1(I)\times C^1(I)$ to $\mathcal{C}^1(I)\times\mathcal{P}_{\mathbb{C}}$.

The main property of the Hamiltonian stochastic embedding procedure (and in fact it can be used as a definition) is to lead to a coherent definition with respect to the momentum process. Precisely, we have:

***Lemma 10.2* (Legendre coherence lemma).** — *Let $L(x,v)$ be an admissible Lagrangian system possessing the Legendre property. The following diagram commutes*

$$\begin{array}{ccc} (x,p) & \xrightarrow{H} & H(x,p) \\ {\scriptstyle S_H}\downarrow & & \downarrow{\scriptstyle S_H} \\ (X,P) & \xrightarrow[H_S]{} & H(X,P) \end{array} \tag{10.11}$$

The proof follows essentially from the fact that the stochastic Hamiltonian embedding of the functional H, denoted by H_S coincide with the definition 10.5 of the stochastic Hamiltonian system associated to H via the Legendre transform and the definition of the momentum process.

10.4. The Hamiltonian least action principle

Using the stochastic Hamiltonian function, we can use the stochastic calculus of variations in order to obtain the set of equations which characterize the stationary processes of the following functional:

$$I_{a,b}(X,P) = E\left[\int_a^b (P(t)\mathcal{D}X - H(X(t),P(t)))\,dt\right], \tag{10.12}$$

defined on $\mathcal{C}^1(I)\times\mathcal{P}_{\mathbb{C}}$.

In order to apply our stochastic calculus of variations, we restrict our attention to I on $\mathcal{C}^1(I)\times\mathcal{C}^1(I)$. The fundamental result of this section is the following:

Theorem 10.1. — *A necessary and sufficient condition for an L-adapted process (X,P) to be $\mathcal{N}^1(I)$-stationary process of the functional $I_{a,b}$ with fixed end points $(X(a),P(a)) = (X_a,P_a)\in H$, $(X(b),P(b)) = (X_b,P_b)\in H$ is that it satisfies the stochastic Hamiltonian equations*

$$\begin{aligned} \mathcal{D}X &= \frac{\partial H}{\partial P}(X(t),P(t)), \\ \mathcal{D}P &= -\frac{\partial H}{\partial X}(X(t),P(t)). \end{aligned} \tag{10.13}$$

Proof. — We must use the weak least action principle using the process $Z = (X, P) \in \mathcal{C}^1(I) \times \mathcal{P}_{\mathbb{C}}$ and the Lagrangian denoted by $\mathcal{L}$ defined on $\mathbb{R}^d \times \mathbb{C}^d \times \mathbb{C}^d \times \mathbb{C}^d$ by

$$\mathcal{L}(x, p, v, w) = pv - H(x, p). \tag{10.14}$$

As $\mathcal{L}(x, p, v, w) = L(x, v)$ formally via the Legendre transform, and L is assumed to be admissible, we deduce that $\mathcal{L}$ is again admissible.

Let δZ be a $\mathcal{N}^1(I)$ variation of the form $Z + \delta Z = (X + X_1, P + P_1)$, where X_1 and P_1 are $\mathcal{N}^1$ processes.

The Euler-Lagrange equation associated to $\mathcal{L}$ is given by

$$\begin{array}{l} \dfrac{\partial \mathcal{L}}{\partial x}(Z(t), \mathcal{D}Z(t)) - \mathcal{D}_\mu \left[\dfrac{\partial \mathcal{L}}{\partial v}(Z(t), \mathcal{D}Z(t)) \right] = 0, \\ \dfrac{\partial \mathcal{L}}{\partial p}(Z(t), \mathcal{D}Z(t)) - \mathcal{D}_\mu \left[\dfrac{\partial \mathcal{L}}{\partial w}(Z(t), \mathcal{D}Z(t)) \right] = 0. \end{array} \tag{10.15}$$

An easy computation leads to

$$\begin{array}{rcl} -\dfrac{\partial H}{\partial x}(Z(t), \mathcal{D}Z(t)) - \mathcal{D}_\mu P(t) & = & 0, \\ \mathcal{D}X(t) - \dfrac{\partial H}{\partial p}(Z(t), \mathcal{D}Z(t)) & = & 0. \end{array} \tag{10.16}$$

This concludes the proof. □

Remark 10.1. — *In this proof we do not need a uniform assumption on the set of variations as the Lagrangian does not depend on the variable w. In fact, we can assume a variation in the direction P which belongs to $\mathcal{C}^1(I)$.*

10.5. The Hamiltonian coherence lemma

In this section, we derive the Hamiltonian analogue of the Lagrangian coherence lemma.

Lemma 10.3 **(The Hamiltonian cohrence lemma)**. — *Let $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be an admissible Hamiltonian system. Then, the following diagram commutes*

$$\begin{array}{ccc} H(x(t), p(t)) & \xrightarrow{S_H} & H(X(t), P(t)) \\ \text{Least action principle} \downarrow & & \downarrow \text{Stochastic least action principle} \\ (HE) & \xrightarrow[S_H]{} & (SHE) \end{array} \tag{10.17}$$

The main point is that this result is not valid if one replaces the Hamiltonian stochastic embedding by the natural stochastic embedding that we have used up to now. We can keep the classical embedding procedure only when dealing with real valued versions of the stochastic derivative. For example, if one deals with the reversible stochastic embedding procedure, we obtain a unified stochastic embedding procedure for both Lagrangian

an Hamiltonian systems. We think however that as well as the complex nature of the stochastic derivative has a fundamental influence on the form of the stochastic Lagrangian equations, i.e. that we obtain the Nelson acceleration, the fact to move from S to S_H reflects a basic properties of the underlying stochastic symplectic geometry we must take into account this complex character of the speed. This problem will be studied in another paper.

CHAPTER 11

CONCLUSION AND PERSPECTIVES

This part aims at discussing possible developments and applications of the stochastic embedding procedure.

11.1. Mathematical developments

11.1.1. Stochastic symplectic geometry. — The Hamiltonian formalism developed in the last part suggest the introduction of what can be called a stochastic symplectic geometry. An interesting construction of symplectic structures on Hilbert spaces is given in [**34**].

The main point here is to construct an analogue of the geometrical structure which puts in evidence the very particular symmetries of the Lagrangian equations in classical mechanics. There exists already many attempt to construct a given notion of symplectic geometry or at least a given geometry for stochastic processes, but they are as far as we know of a different nature. We refer to the book of Elworthy, LeJan and Li [**44**] for an overview. These geometries are only associated to stochastic processes and translate into data of geometrical nature properties of the underlying stochastic processes (like the Riemannian or sub-Riemannian structure associated to Brownian motions and diffusions).

A recent work of J-C. Zambrini and P. Lescot ([**37**] and [**38**]) deals specifically with symplectic geometry and a notion of integrability by quadratures.

For a discussion of integrability in our context see section 11.1.2.

11.1.2. PDE's and the stochastic embedding. — The stochastic embedding of Lagrangian systems over diffusion processes lead to a PDE governing the density of the

solutions of the stochastic Euler-Lagrange equation. Moreover, we have defined a stochastic Hamiltonian system naturally associated to the Lagrangian. However, some classical PDEs, as for example the Schrödinger equation, possess an Hamiltonian formulation. This remark, which goes back to the work of Zakharov V.E. and Faddeev D. [**72**] is now an important subject in PDEs known as Hamiltonian PDEs (see for example [**34**]). As a consequence, we have the following situation:

$$\begin{array}{ccc} & & H_S \\ & & \downarrow \\ PDE & \longrightarrow & H \end{array} \tag{11.1}$$

Of course the relation between the PDE and H_S is not of the same nature as the relation with H.

In the sequel, we list a number of problems and questions which naturally arise from the previous diagram:

– There exists a notion of completely integrable Hamiltonian PDE (see [**34**]). What about out stochastic Hamiltonian systems ?

Assuming that we have a good notion of integrability for H_S, we have the following questions:

– Are there any relations between the integrability of H and H_S?
– Is there a stochastic analogue of the Arnold-Liouville theorem?
– Is there a special set of "coordinates" similar to the action/angle variables?

We note that there already exists such a notion for Hamiltonian PDEs (see [**72**]).

– Is there a notion of integrability by "quadratures"?

In that respect, we think about Lax work [**36**] on the integrability of PDEs.

11.2. Applications

11.2.1. Long term behaviour of chaotic Lagrangian systems. — The dynamical behaviour of unstable or chaotic dynamical systems is far from being understood, unless we restrict to a very particular class of systems like hyperbolic systems or weak version of hyperbolicity. This question arises naturally for small perturbations of Hamiltonian systems for which there exists a large family of results dealing with this problem, as for

example the KAM (Kolmogorov-Arnold-Moser) theorem, Nekhoroshev theorem and special phenomena like the Arnold diffusion related to the so-called quasi-ergodic hypothesis.

Unfortunately, these results are difficult to use in concrete situations and only direct numerical simulations provide some understanding of the dynamics [**22**].

There exists of course ergodic theory which tries to look for weaker information on the dynamics than a direct qualitative approach. However, this theory leads also to very difficult problems when one tries to implement it, as for example in the case of Sinaï billiard. Moreover, there is a widely opinion in the applied community that the long term behaviour of a chaotic systems is more or less equivalent to a stochastic process. One example of such opinion is well expressed in the article of J. Laskar [**41**] in the context of the chaotic behaviour of the Solar system: "Since the characteristic time scale for the divergence of nearby orbits in the Solar system is approximately 5 Myr, the orbital evolution of the planet becomes practically unpredictable after 100 Myr. Thus in the long term, the motion of the Solar system may be described by a random process, where orbits wander erratically in a chaotic zone."

What are the arguments leading to this idea ?

The first point is that chaotic dynamical systems are in general characterized by the so-called sensitivity to initial conditions, meaning that a small error on the initial condition leads to very different solutions. Of course, one must quantify this kind of sentence, and we can do that, with more or less canonicity, by introducing Lyapounov exponents and Lyapounov time. Whatever we do, there is a non canonical data in this, which is precisely to what extent we consider that two solutions are different. This must be a matter of choice for a given system, and cannot be fixed by any mathematical tool. In the sequel, we assume that a system is sensitive to initial conditions in some region R of the phase space, and for a given metric, if for all $x_0 \in R$ and all $\epsilon > 0$, the distance at time t between a trajectory starting at x_0 and $x_0 + \epsilon$, denoted by $d(t)$ is [(1)] approximately given by

$$d(t) = \epsilon e^{t/T}, \tag{11.2}$$

[(1)] As we already stress, we can in some situations gives a precise meaning to all this point, like for example in the Smale Horseshoes, but this is far to cover the wide variety of chaotic behaviour which are studied in the applied literature.

where $T > 0$ is the so-called Lyapounov time or horizon of predictability for the system[(2)]. For an example of such an estimate, we refer to J. Laskar [**42**] where he gives numerical evidences for the chaotic behaviour of the solar system.

As a consequence, for t sufficiently large with respect to T, we have no prediction any more, or in other words, we can not assign to a given prediction a precise initial condition. We then have lost the deterministic character of the equations of motions. An idea is then to say that one musts then consider not a fixed initial condition x_0, but a given random variable representing all the possible behaviours (kind of trajectories) one is lead to after a fixed time t: for example, $\epsilon > 0$ being fixed, we consider all the intersections of trajectories starting in the disk $D(x_0, \epsilon)$ with the ball $B(x_0, \epsilon)$. We then obtain a family of directions. Assuming that we can compute an average over the family of such a quantity which obtain an averaged direction which select a given point of the ball $B(x_0, \epsilon)$. We then follow the selected trajectory during the time t, and continue again this procedure. Such a construction is reminiscent of the classical construction of the Brownian motion (see [**30**],p.66). Of course, this programme can only be carried in some specific examples. We refer to the article of Y. Sinaï [**62**] for an heuristic introduction to all these problems.

If we agree with the previous heuristic idea, one can then ask for the following: how is the underlying stochastic process governed by the dynamical system ?

We return again to the Hamiltonian/Lagrangian case. The stochastic embedding procedure answers precisely this question. The stochastic Euler-Lagrange equation is the track of the underlying Lagrangian system on stochastic processes. As a consequence, we can think that we are able to capture even the desired long term behaviour of the Lagrangian system using this procedure.

In order to support our point of view, we suggest the following strategy:

Consider a perturbation of a completely integrable Hamiltonian system $H_\epsilon(x) = h(x) + \epsilon f(x)$, with $x \in \mathbb{R}^{2n}$ for example. Let us assume that $h(x)$ leads to a particular PDE under stochastic embedding, which can be well understood and solved. The long term behaviour of the completely integrable Hamiltonian system is trivial. This not the case for the stochastic analogue. What about the long term behaviour of H_ϵ ? We think that

[(2)] In concrete systems, one must involve a macroscopic scale (see [**21**],p.17), which bound the admissible size of an error on a prediction. Here, this quantity is arbitrary replaced by e.

it is controlled by the stochastic analogue of the unperturbed Hamiltonian. This result is related to a kind of stochastic stability which we must define. However, this approach can be tested on a wide variety of examples, in particular celestial mechanical problems.

11.2.2. Celestial mechanics. — There exist many theories dealing with the problem of the formation of gravitational structures. For planetary systems this question is related to a long standing problem related to the "regular" spacing of planets in the Solar system. This problem which goes back to Kepler (1595), Kant (1755), von Wolf (1726), Lambert (1761), takes a mathematical form under the Titius (1766) formulation of the so called Titius-Bode law giving a geometric progression of the distance of the planets from the sun. We refer to the book of Nieto [**56**] for more details. Even if this empirical law fails to predict correctly the real distance for the Planet Pluto for example, its interest is that it suggests that the repartition of exoplanet orbital semi-major axes could satisfy a simple law. As a consequence, one searchs for a possible physical/dynamical theory supporting the existence of such kind of law. Moreover, the discovery of many exo-planetary systems can be used to test if the theory is based on universal phenomena and not related to our knowledge of the Solar system.

All the actual theories about the origin of the solar system presuppose the formation of a protoplanetary nebula, formed by some material (gas, dust, etc ...) with a central body (a star or a big planet). We refer to Lissauer [**43**] for more details.

Instead, we use a simplified model consisting of a large central body of mass m_0 with a large number of small bodies $(m_j)_{j=1,\dots,n}$, whose mass is assumed to be small with respect to m_0. The main problem is to understand the long term dynamics of this model.

Following the work of Albeverio S., Blanchard Ph. and R. Hoegh-Krohn ([**3**], see also [**4**]), we can modelize the motion of a given grain in the protoplanetary nebula by a stochastic process (see [**3**],p.366-367), more precisely a diffusion process. The problem is then to find what is the equation governing the dynamics of such a stochastic process. Using our stochastic embedding theory, we can use the classical formulation in order to obtain the desired equation. This question will be detailed in a forthcoming article.

The main idea behind stochastic modelisation is the following:

The motion of a given small body in a protoplanetary nebula is given by the Kepler model and a perturbation due to the large number of number of small bodies. In [**3**],

this perturbation is replaced by a white noise. As a consequence, the movement of a small body is assumed to be described by a diffusion process. It must be noted that this assumption is related to a number of arguments, one of them being that the dynamics of the underlying classical system is unstable. We then return to our previous description of the chaotic behaviour of a dynamical system. However, using the stochastic embedding theory, we can try to justify the passage from a classical motion to a stochastic one looking at the following problem:

Let $L_\epsilon = L_{\text{Kepler}} + P_\epsilon$, be the Lagrangian system describing the dynamics of our model. The Lagrangian L_{Kepler} is the classical Lagrangian of the Kepler problem, and P_ϵ is the perturbation. Using the stochastic embedding theory, we can deduce two stochastic dynamical systems, one associated to L_ϵ and denoted by S_ϵ and one associated to L_{Kepler} denoted by S_{Kepler}. If the previous strategy to replace the perturbative effect by a White noise is valid, then we must have a kind of stochastic stability between S_{Kepler} and S_ϵ. The notion of stochastic stability must be defined rigorously and be consistent with the stochastic embedding theory[(3)]. Why such a stability result is reasonable ? The main thing is that we already look in S_{Kepler} for statistical properties of the set of trajectories of stochastic (diffusion) processes under the Kepler Lagrangian. There is no reason that the statistic of this trajectories really differs when adding a small perturbation. This is of course different if one look for the underlying deterministic system. All these questions will be studied in a forthcoming paper.

11.2.3. Strange attractors. — *Strange attractors* play a fundamental role in turbulence and lead to many difficult problems. Most of the time, one is currently interested in the geometrical properties of attractors (Hausdorf dimension,...), special dynamical properties (existence of an SRB (Sibaï-Ruelle-Bowen) measure [**68**], stability under perturbations....). However, focusing on a given attractor hides the fact that most of the time we can not predict from the equation the *existence* of such an attractor. This is in particular the case for the *Lorenz attractor* or the *Henon attractor*. These attractors are obtained numerically. In some models, we can construct a geometric model from which we can prove the existence of such a structure (this is the case for the geometric Lorenz model) [**27**]. For example, S. Smale [**63**] asks for an existence proof for the Lorenz equation of the attractor. This has been done recently by W. Tucker ([**66**], [**67**]). However, no general strategy exists in order to predict such an attractor.

[(3)]It must be noted that there exists already several notion of stochastic stability in the literature, as for example Has'inskii [**29**], Kushner [**35**] and more recently Handel [**28**].

Our idea is to use the stochastic embedding theory in order to predict the existence of such an object. Let us consider the Lorenz equations. These equations are not a Lagrangian system. However, there exits a canonical embedding in a Lagrangian system (see the report of M. Audin [**7**]). This lagrangian can then be studied via the stochastic embedding procedure. The solutions are stochastic processes whose density is controlled by a PDE. As we already explain, we expect that the long term behaviour of the system is coded by this PDE. As the long term dynamics of the Lorenz system if precisely supported by the Lorenz attractor, we think that this structure can be detected in the PDE (as a stationary state for example).

We can also take this problem as a first step towards understanding the existence of coherent structures in chaotic dynamical systems. Moreover, the Lorenz attractor is widely studied and there exists a great amount of results like the existence of a unique SRB measure (see [**67**]). We can then take this example as a good system to compare classical methods of *ergodic theory* and our approach. For more problems related to the Lorenz attractor, SRB measure ..., see ([**69**],[**70**]).

NOTATIONS

d: dimension

$(\Omega, \mathcal{A}, P)$ a probability space

- Stochastic processes

- We denote by
$$dX = b(t, X)dt + \sigma(t, X)dW, \qquad (*)$$
the stochastic differential equation where b is the drift, σ the diffusion matrix and W is a d-dimensional Wiener process defined on $(\Omega, \mathcal{A}, P)$.
- We denote by $X(t)$ the solution of (*) and by $p_t(x)$ its density (when it exists) at point x.
- $\sigma(X_s, a \leqslant s \leqslant b)$: the σ-algebra generated by X between a and b
- $\mathcal{F}_t$: an increasing σ algebra
- $\mathcal{P}_t$: an decreasing σ algebra
- $E[\bullet \mid \mathcal{B}]$: the conditional expectation.
- $\| \, . \, \|$: norm on stochastic processes.

- Functional spaces

- $\mathcal{P}_{\mathbb{R}}$: real valued stochastic processes
- $\mathcal{P}_{\mathbb{C}}$: complex valued stochastic processes
- $\mathcal{P}_{\text{det}}$: the set of deterministic stochastic processes
- $\mathcal{P}^k_{\text{det}}$: the set of deterministic stochastic processes such that $X(\omega)$ is of class C^k
- Λ_d: good diffusion processes
- Λ^g_d: good diffusion processes with a gradient drift

- $L^p(\Omega)$: set of random variables which belongs to L^p
- $\mathbf{L}^2$: the set of real valued processes which are $\mathcal{P}_t$ and $\mathcal{F}_t$ adapted and such that $E\left[\int_0^1 X_t^2\,dt\right] < \infty$.
- $C^{1,2}((0,1)\times\mathbb{R}^d)$ the set of function which are C^1 in the first variable and C^2 in the second one.
- $\mathcal{N}^1$: the set of Nelson differentiable processes.

- Operators

- ∇: the gradient
- Δ: the Laplacian
- Let $f(x_1,\ldots,x_n)$ be a given function. We denote by $\partial_{x_i} f$ the partial derivative of f with respect to x_i
- Let $f(x_1,\ldots,x_n,y_1,\ldots,y_m)$ be a given function. We denote by $\partial_x f$, $x=(x_1,\ldots,x_n)$ the partial differential of f in the direction x.
- D: Nelson forward derivative
- D_*: Nelson backward derivative
- $\mathcal{D}$: the stochastic derivative
- D^n, D_*^n, $\mathcal{D}^n$: the n-th iterate of D, D_* or $\mathcal{D}$
- $\mathfrak{d}$ and $\mathfrak{d}_*$: adapted forward and backward derivative

$k \geqslant 1$

- $\mathcal{C}^k$: the set of real valued processes which are $\mathcal{P}_t$ and $\mathcal{F}_t$ adapted and such that $\mathcal{D}^i$ exists, $1 \leqslant i \leqslant k$.
- $\mathcal{C}_{\mathbb{C}}^k$: the set of complex valued processes which are $\mathcal{P}_t$ and $\mathcal{F}_t$ adapted and such that $\mathcal{D}^i$ exists, $1 \leqslant i \leqslant k$.

- $\mathrm{Re}(z)$: real part of $z \in \mathbb{C}$.
- $\mathrm{Im}(z)$: imaginary part of $z \in \mathbb{C}$.

BIBLIOGRAPHY

[1] Aase K., Oksendal B., Privault N., Uboe J., White noise generalizations of the Clark-Haussmann-Ocone theorem with application to mathematical finance, Finance Stochast. 4, 465-496 (2000).

[2] Abraham R., Marsden J.E., *Foundations of Mechanics*, 2d edition, Benjamin/Cummings Pub. Comp., 1978.

[3] Albeverio S., Blanchard Ph., Hoegh-Krohn R., A stochastic model for the orbits of planets and satellites: an interpretation of Titius-Bode law, Expositiones Mathematicae 4, 363-373 (1983).

[4] Albeverio S., Blanchard Ph., Hoegh-Krohn R., Newtonian diffusions and planets, with a remark on non-standard Dirichlet forms and polymers,

[5] Arnold V.I., *Mathematical Methods of Classical Mechanics*, 2d edition, Springer, 1989.

[6] Arnold V.I., *Chapitres supplémentaires de la théorie des équations différentielles ordinaires*, Editions Mir Moscou, Editions Librairie du Globe, Paris, 1996.

[7] Audin M., *Les systèmes hamiltoniens et leur intégrabilité*, Cours spécialisé, 8, Société Mathématiques de France et EDP Sciences, 2001.

[8] Bates S., Weinstein A., *Lectures on geometric quantization*, Berkeley Mathematics Lectures Notes 8, Amer. Math. Soc., Providence, 1997.

[9] Biernacki C, Cresson J, Chretien S, Darses S, About the Titus-Bode law, *in preparation.*

[10] Bismut J.-M., *Mécanique aléatoire*, Lecture Notes in Mathematics 866, 1981.

[11] Carmona R, Probabilistic construction of Nelson processes, Taniguchi Symp. PMMP Katata, 1985, pp. 55-81.

[12] Chung K.L., Zambrini J.C., *Introduction to random time and quantum randomness*, Monographs of the Portuguese Mathematical society, Vol. 1, World Scientific, 2003.

[13] Cresson J., Scale calculus and the Schrodinger equation, Journal of Mathematical Physics, Vol. 44, no. 12, 32.p, 2003.

[14] Cresson J., Non-differentiable variational principles, Journal of Mathematical analysis and Applications, Vol. 307, no. 1, 48-64, 2005.

[15] Cresson J., Théories de plongement des systèmes dynamiques - un programme, 21.p, 2005.

[16] Cresson J., Sur les lois de la nature [des lois de Kepler aux lois du chaos], in preparation, 2005.

[17] Cresson J., Darses S., Stochastic symplectic geometry, *in preparation*.

[18] Cresson J., Darses S., Long term behaviour of Lagrangian systems, *in preparation*.

[19] Darses S., Saussereau B., Time reversal for stochastic differential equations driven by fractional Brownian motion, *in preparation*.

[20] Doob J., The Brownian movement and stochastic equtions, Annals of Mathematics Vol. 43, No. 2, 1942.

[21] Douady A., Déterminisme et indéterminisme dans un modèle mathématique, in *Chaos et déterminisme*, A. Dahan Dalmedico, J.-L. Chabert, K. Chemla, Editions du Seuil, 1992, pp. 11-18.

[22] Dumas S., Laskar J., Global dynamics and long-time stability in Hamiltonian systems via numerical frequency analysis, Phys. Rev. Letters Vol. 70, No. 20, 2975-2979, 1993.

[23] Falconer K, *Fractal geometry; Mathematical Foundations and Applications*, John Wiley and Sons ed, 1990.

[24] Feynman R., The development of the space-time view of quantum electrodynamics, Nobel lecture, December 11, 1965.

[25] Föllmer H., Time reversal on Wiener space, in *Stochastic processes - Mathematics and Physics*, Proc. Bielefeld, 1984, Lecture Notes in Math. 1158 (1986), 119-129.

[26] Gignoux C., Silvestre-Brac B., *Mécanique de la formulation lagrangienne au chaos hamiltonien*, EDP Sciences, 2002.

[27] Guckenheimer J., Holmes P., *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Appl. Math. Sciences 42, (1983).

[28] Handel R.V., Almost global stochastic stability, arxiv/Math.PR/0411311, 2004.

[29] Has'manskii R.Z., *Stochastic stability of differential equations*, Sijthoff and Noordhoff, 1980.

[30] Karatzas I., Schreve S.E., *Brownian motion and stochastic calculus*, 2d edition, Graduate Text in Mathematics 113, Springer, 1991.

[31] Katok A., Hasselblatt B., *Introduction to the modern theory of dynamical systems*, Cambridge University Press, 1999.

[32] Kosmann-Schwarzbach Y, *Les théorèmes de Noether; Invariance et lois de conservation au XXème siècle*, Les éditions de l'École Polytechnique, 2004.

[33] Kostant B., *Quantization of unitary representations, part 1: prequantization*, Lecture Notes in Mathematics 170 (1970), 87-208.

[34] Kuksin S.B., *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its applications, 19, 2000.

[35] Kushner R.J., *Stochastic stability and control*, Academic Press, 1967.

[36] Lax P.D., Integrals of Nonlinear equations of evolution and solitary waves, Communications on pure and applied Mathematics, Vol. XXI, 467-490, 1968.

[37] Lescot P., Zambrini J-C, Isovectors for Hamilton-Jacobi-Bellman equation, formal stochastic differentials and first integrals in Euclidean Quantum Mechanics, in Seminar on Stochastic Analysis, random fields and applications IV, Ascona 2002, Progress in probability, Vol. 58, 187-202, Birkhäuser, 2004.

[38] Lescot P., Zambrini J-C, private communication, 2005.

[39] Lamb J, Roberts A, Time-reversal symmetry in dynamical systems, preprint, 1997.

[40] Laskar J, A numerical experiment on the chaotic behaviour of the solar system, Nature 338 (1989), 237-238.

[41] Laskar J., On the spacing of planetary systems, Phys. Rev. Let., 84, 15, pp. 3240-3243, 2000.

[42] Laskar J., Chaos in the solar system, 2003.

[43] Lissauer J., in *Origins, Proceedings of the ASP conference* edited by C.E. Woodward, J.M. Shull, H.A. Thronson, Astronomical Society of the Pacific, San Francisco, p. 327, 1998.

[44] Elworthy K.D., LeJan Y., Li X-M, *On the geometry of diffusion operators and stochastic flows*, Lecture Notes in Mathematics 1720, Springer, 1999.

[45] Mackey G.W., *The mathematical foundations of quantum mechanics*, W.A. Benjamin, Inc., New York, 1963.

[46] Marle C-M., De la mécanique classique à la mécanique quantique: pourquoi et comment quantifier ?, preprint

[47] Marmi S, Chaotic behaviour in the solar system [following J. Laskar], Séminaire Bourbaki, 51ème année, 1998-99, no. 854.

[48] McDonald S., Beard R., An application of stochastic calculus of variations and the stochastic Euler equation to financial economics, preprint, 2003.

[49] Millet A, Nualart D., Sanz M., Integration by parts and time reversal for diffusion processes, The Annals of Probability 1989, Vol. 17, No. 1, 208-238.

[50] Moser J., Is the solar system stable ?,

[51] Mumford D., The dawning of the age of stochasticity, in *Mathematics: Frontiers and perspectives*, V. Arnold, M. Atiyah, P. Lax, B. Mazur editors, AMS, 2000, 197-218.

[52] Misawa T., Yasue K., Canonical dynamical systems, J. Math. Phys. 28(11), 1987, 2569-2573.

[53] Nelson E., *Dynamical theories of Brownian motion*, second edition, Princeton, 2001.

[54] Nelson E., Derivation of the Schrödinger equation from Newtonian mechanics, Physical Review, Vol. 150, No. 4, 1079-1084 (1966).

[55] Nelson E., Stochastic mechanics and random fields, in Ecole d'été de Probabilité de Saint-Flour XV-XVII, edited by P.L. Hennequin, Lecture Notes in Mathematics, Springer-Verlag, 1988, Vol. 1362, pp. 428-450.

[56] Nieto M.M., *The Titius-Bode law of planetary distances: its history and theory*, Pergamon Press, Oxford, 1972.

[57] Nottale L, New formulation of stochastic mechanics. Application to chaos, in "*Chaos and diffusion in Hamiltonian systems*", Proceedings of the fourth workshop in Astronomy and Astrophysics of Chamonix (France), 7-12 February 1994, Eds. D. Benest and C. Froeschlé (Editions Frontières), pp. 173-198 (1995).

[58] Nualart D., *The Malliavin calculus and related topics*, Probability and its applications, Springer, 1995.

[59] Pavon M., A new formulation of stochastic mechanics, Physics Letters A 209 (1995), 143-149.

[60] Pavon M., Hamilton's principle in stochastic mechanics, J. Math. Phys. 36 (1995), 6774.

[61] Roelly S., Thieullen M., A characterizatin of reciprocal processes via an integration by part formula on the path space, Prob. Th. Rel. Fields 123 (2002) 97-120.

[62] Sinaï Y.G., L'aléatoire du non aléatoire, in *Chaos et déterminisme*, A. Dahan Dalmedico, J.-L. Chabert, K. Chemla, Editions du Seuil, 1992, pp. 68-87.

[63] Smale S., Mathematical problems for the Next century, Math. Intell., 1998.

[64] Thieullen M., Second order stochastic differential equations and non-Gaussian reciprocal diffusions, Proba. Theory and Rel. Fields 97, 231-257 (1993).

[65] Souriau J-M, *Structure des systèmes dynamiques*, Dunod, Paris, 1970.

[66] Tucker W., A rigorous ODE solver and Smale's 14th problem, Found. Comp. Math. 2:1, 53-117, 2002.

[67] Tucker W., The Lorenz attractor exists, C. R. Acad. Sci. Math. t. 328, Série I, 1197-1202, 1999.

[68] Viana M., Stochastic dynamics of deterministic systems, Braz. Math. Colloq. 21, IMPA, 1997.

[69] Viana M., Dynamical systems: moving into the next century, in *Mathematics unlimited*, 2001.

[70] Viana M., A probabilistic and geometric perspective, Doc. Math. Extra Volume ICM 1998.

[71] Yasue K, Stochastic calculus of variations, Journal of functional Analysis 41, 327-340 (1981).

[72] Zakharov V.E., Faddeev L.D., Korteweg-de Vries equation: a completely integrable Hamiltonian system, Functional analysis and its applications, 5:4 (1971), 280-287.

[73] Zheng W.A., Meyer P.A., Quelques résultats de mécanique stochastique, Séminaire de Probabilités XVIII, 223-243

[74] Zambrini J.-C., Yasue K., Semi-Classical quantum mechanics and stochastic calculus of variations, Annals of Physics 143, 54-83 (1982).