

Functional Ito calculus and functional Kolmogorov equations*

Rama CONT

Lecture Notes of the
BARCELONA SUMMER SCHOOL ON STOCHASTIC ANALYSIS
Centre de Recerca Matemàtica, July 2012.

Abstract

The *Functional Ito calculus* is a non-anticipative functional calculus which extends the Ito calculus to path-dependent functionals of stochastic processes. We present an overview of the foundations and key results of this calculus, first using a pathwise approach, based on a notion of directional derivative introduced by Dupire, then using a probabilistic approach, which leads to a weak, Sobolev-type, functional calculus for square-integrable functionals of semimartingales, without any smoothness assumption on the functionals. The latter construction is shown to have connections with the Malliavin calculus and leads to computable martingale representation formulae. Finally, we show how the Functional Ito Calculus may be used to extend the connections between diffusion processes and parabolic equations to a non-Markovian, path-dependent setting: this leads to a new class of 'path-dependent' partial differential equations, Functional Kolmogorov equations, of which we study the key properties.

Keywords: stochastic calculus, functional Ito calculus, Malliavin calculus, change of variable formula, functional calculus, martingale, Kolmogorov equations, path-dependent PDE.

Published as:

Functional Itô calculus and functional Kolmogorov equations, in: V. Bally, L Caramellino, R Cont: *Stochastic Integration by Parts and Functional Itô Calculus*, Advanced Courses in Mathematics CRM Barcelona, pages 123–216.

*These notes are based on a series of lectures given at various summer schools: Tokyo (Tokyo Metropolitan University, July 2010), Munich (Ludwig Maximilians University, Oct 2010), the Spring School "Stochastic Models in Finance and Insurance" (Jena, 2011), the International Congress for Industrial and Applied Mathematics (Vancouver 2011), the Barcelona Summer School on Stochastic Analysis (Barcelona, July 2012). I am grateful to Hans Engelbert, Jean Jacod, Hans Föllmer, Yuri Kabanov, Arman Khaledian, Shigeo Kusuoka, Bernt Oksendal, Candia Riga, Josep Vives, Frederic Utzet and especially the late Paul Malliavin for helpful comments and discussions.

Contents

4	Overview	125
4.1	Functional Ito calculus	125
4.2	Martingale representation formulas	126
4.3	Functional Kolmogorov equations and path-dependent PDEs . .	127
4.4	Outline	127
5	Non-anticipative functional calculus	130
5.1	Non-anticipative functionals	130
5.2	Horizontal and vertical derivatives	133
5.2.1	Horizontal derivative	134
5.2.2	Vertical derivative	135
5.2.3	Regular functionals	137
5.3	Pathwise integration and functional change of variable formula .	140
5.3.1	Pathwise quadratic variation	141
5.3.2	Functional change of variable formula	144
5.3.3	Pathwise integration for paths of finite quadratic variation	147
5.4	Functionals defined on continuous paths	150
5.5	Application to functionals of stochastic processes	155
6	The Functional Ito formula	157
6.1	Semimartingales and quadratic variation	157
6.2	The Functional Ito formula	159
6.3	Functionals with dependence on quadratic variation	161
7	Weak functional calculus for square-integrable processes	166
7.1	Vertical derivative of an adapted process	167
7.2	Martingale representation formula	170
7.3	Weak derivative for square-integrable functionals	171
7.4	Relation with the Malliavin derivative	174
7.5	Extension to semimartingales	177
7.6	Changing the reference martingale	182
7.7	Application to Forward-Backward SDEs	182
8	Functional Kolmogorov equations (with D. Fournié)	185
8.1	Functional Kolmogorov equations and harmonic functionals . . .	186
8.1.1	SDEs with path-dependent coefficients	186
8.1.2	Local martingales and harmonic functionals	188
8.1.3	Sub-solutions and super-solutions	190
8.1.4	Comparison principle and uniqueness	191
8.1.5	Feynman-Kac formula for path-dependent functionals . .	192
8.2	FBSDEs and semilinear functional PDEs	193
8.3	Non-Markovian control and path-dependent HJB equations . . .	195
8.4	Weak solutions	198

4 Overview

4.1 Functional Ito calculus

Many questions in stochastic analysis and its applications in statistics of processes, physics or mathematical finance involve the study of *path-dependent functionals* of stochastic processes and there has been a sustained interest in developing an analytical framework for the systematic study of such path-dependent functionals.

When the underlying stochastic process is the Wiener process, the Malliavin calculus [4, 55, 52, 56, 68, 74] has proven to be a powerful tool for investigating various properties of Wiener functionals. The Malliavin calculus, which is as a weak functional calculus on Wiener space, leads to differential representations of Wiener functionals in terms of *anticipative* processes [5, 37, 56]. However, the interpretation and computability of such anticipative quantities poses some challenges, especially in applications such as mathematical physics or optimal control where causality, or non-anticipativeness, is a key constraint.

In a recent insightful work, motivated by applications in mathematical finance, Bruno Dupire [21] proposed a method for defining a non-anticipative calculus which extends the Ito calculus to path-dependent functionals of stochastic processes. The idea can be intuitively explained by first considering the variations of a functional along a *piecewise constant* path. Any (right-continuous) piecewise constant path, represented as

$$\omega(t) = \sum_{k=1}^n x_k 1_{[t_k, t_{k+1}[},$$

is simply a finite sequence of 'horizontal' and 'vertical' moves, so the variation of a (time-dependent) functional $F(t, \omega)$ along such a path ω is composed of

- 'horizontal increments': variations of $F(t, \omega)$ between each time point t_i and the next, and
- 'vertical increments': variations of $F(t, \omega)$ at each discontinuity point of ω .

If one can control the behavior of F under these two types of path perturbations then one can reconstitute its variations along any piecewise-constant path. Under additional continuity assumptions, this control can be extended to any cadlag path using a density argument.

This intuition was formalized by Dupire [21] by introducing directional derivatives corresponding to infinitesimal versions of these variations: given a (time-dependent) functional $F : [0, T] \times D([0, T], \mathbb{R}) \rightarrow \mathbb{R}$ defined on the space $D([0, T], \mathbb{R})$ of right-continuous paths with left limits, Dupire introduced a directional derivative which quantifies the sensitivity of the functional to a shift in the future portion of the underlying path $\omega \in D([0, T], \mathbb{R})$:

$$\nabla_{\omega} F(t, \omega) = \lim_{\epsilon \rightarrow 0} \frac{F(t, \omega + \epsilon 1_{[t, T]}) - F(t, \omega)}{\epsilon},$$

as well as a time-derivative corresponding to the sensitivity of F to a small 'horizontal extension' of the path:

$$\mathcal{D}F(t, \omega) = \lim_{h \rightarrow 0^+} \frac{F(t+h, \omega(t \wedge \cdot)) - F(t, \omega(t \wedge \cdot))}{h}.$$

Since any cadlag path may be approximated, in supremum norm, by piecewise constant paths, this suggests that one may control the functional F on the entire space $D([0, T], \mathbb{R})$ if F is twice differentiable in the above sense and $F, \nabla_\omega F, \nabla_\omega^2 F$ are continuous in supremum norm; under these assumptions, one can then obtain a change of variable formula for $F(X)$ for any Ito process X .

As this brief description already suggests, the essence of this approach is pathwise. While Dupire's original presentation [21] uses probabilistic arguments and Ito calculus, one can in fact do entirely without such arguments and derive these results in a purely analytical framework without any reference to probability. This task, undertaken in [9, 8] and developed in [6], leads to a pathwise functional calculus for *non-anticipative functionals* which clearly identifies the set of paths to which the calculus is applicable. The pathwise nature of all quantities involved makes this approach quite intuitive and appealing for applications, especially in finance [13] and optimal control, where all quantities involved need to make sense pathwise.

However, once a probability measure is introduced on the space of paths, under which the canonical process is a semimartingale, one can go much further: the introduction of a reference measure allows to consider quantities which are defined almost-everywhere and construct a **weak functional calculus** for stochastic processes defined on the canonical filtration. Unlike the pathwise theory, this construction, developed in [7, 11], is applicable to all square-integrable functionals **without any regularity condition**. This calculus can be seen as a non-anticipative analog of the Malliavin calculus.

The Functional Ito calculus has led to various applications in the study of path-dependent functionals of stochastic processes. Here we focus on two particular directions: martingale representation formulas and functional ('path-dependent') Kolmogorov equations [10].

4.2 Martingale representation formulas

The representation of martingales as stochastic integrals is an important result in stochastic analysis with many applications in control theory and mathematical finance. One of the challenges in this regard has been to obtain explicit versions of such representations, which may then be used to compute or simulate such martingale representations. The well-known Clark-Haussmann-Ocone formula [56, 55], which expresses the martingale representation theorem in terms of Malliavin derivative is one such tool and has inspired various algorithms for the simulation of such representations [33].

One of the applications of the Functional Ito calculus is to derive explicit, computable versions of such martingale representation formulas, without resorting to the anticipative quantities such as the Malliavin derivative. This

approach, developed in [11, 7], leads to simple algorithms for computing martingale representations which have straightforward interpretations in terms of sensitivity analysis [12].

4.3 Functional Kolmogorov equations and path-dependent PDEs

One important application of the Ito calculus has been to characterize the deep link between Markov processes and partial differential equations of parabolic type [2]. A pillar of this approach is the analytical characterization of a Markov process by Kolmogorov's backward and forward equations [46]. These equations have led to many developments in the theory of Markov processes and stochastic control theory including the theory of controlled Markov processes and their links with viscosity solutions of PDEs [29].

The functional Ito calculus provides a natural setting for extending many of these results to more general, non-Markovian semimartingales, leading to a new class of partial differential equations on path space—*functional Kolmogorov equations*—which have only started to be explored [10, 14, 22]. This class of PDEs on the space of continuous functions is distinct from the infinite dimensional Kolmogorov equations studied in the literature [15]. Functional Kolmogorov equations have many properties in common with their finite dimensional counterparts and lead to new Feynman-Kac formulas for path-dependent functionals of semimartingales [10]. We will explore this topic in Section 8. Extensions of these connections to the fully nonlinear case and their connection to non-Markovian stochastic control and forward-backward stochastic differential equations (FBSDEs) currently constitute an active research topic [10, 14, 23, 22, 59].

4.4 Outline

These notes, based on lectures given at the Barcelona Summer School on Stochastic Analysis (2012), constitute an introduction to the foundations and applications of the Functional Ito calculus.

- We first develop a **pathwise calculus for non-anticipative functionals** possessing some directional derivatives, by combining Dupire's idea with insights from the early work of Hans Föllmer [31]. This construction is purely analytical (i.e. non-probabilistic) and applicable to functionals of paths with finite quadratic variation. Applied to functionals of a semimartingale, it yields a functional extension of the Ito formula applicable to functionals which are continuous in the supremum norm and admit certain directional derivatives. This construction and its various extensions, which are based on [8, 9, 21] are described in Sections 5 and 6. As a by-product (!) we obtain a method for constructing pathwise integrals with respect to paths of infinite variation but finite quadratic variation, for a class of integrands which may be described as 'vertical 1-forms';

the connection between this pathwise integral and 'rough path' theory is described in Section 5.3.3.

- In Section 7 we extend this pathwise calculus to a '**weak**' **functional calculus** applicable to square-integrable adapted functionals with **no regularity condition** on the path-dependence. This construction uses the probabilistic structure of the Ito integral to construct an extension of Dupire's derivative operator to all square-integrable semimartingales and introduce Sobolev spaces of non-anticipative functionals to which weak versions of the functional Ito formula applies. The resulting operator is a weak functional derivative which may be regarded as a *non-anticipative* counterpart of the Malliavin derivative (Section 7.4). This construction, which extends the applicability of the Functional Ito calculus to a large class of functionals, is based on [11]. The relation with the Malliavin derivative is described in Section 7.4. One of the applications of this construction is to obtain explicit and computable integral representation formulas for martingales (Section 7.2 and Theorem 7.8).
- Section 8 uses the Functional Ito calculus to introduce *Functional Kolmogorov equations*, a new class of partial differential equations on the space of continuous functions which extend the classical backward Kolmogorov PDE to processes with path-dependent characteristics. We first present some key properties of classical solutions for this class of equations, and their relation with FBSDEs with path-dependent coefficients (Section 8.2) and non-Markovian stochastic control problems (Section 8.3). Finally, in Section 8.4 we introduce a notion of weak solution for the functional Kolmogorov equation and characterize square-integrable martingales as weak solutions of the functional Kolmogorov equation.

Notations

In the sequel we denote by

- S_d^+ the set of symmetric positive $d \times d$ real-valued matrices,
- $\langle A, B \rangle = \text{tr}({}^t A.B)$ the Hilbert-Schmidt scalar product of two real $d \times d$ matrices,
- $D([0, T], \mathbb{R}^d)$ the space of functions on $[0, T]$ with values in \mathbb{R}^d which are right continuous functions with left limits (cadlag), and
- $C^0([0, T], \mathbb{R}^d)$ the space of continuous functions on $[0, T]$ with values in \mathbb{R}^d .

Both spaces are equipped with the supremum norm, denoted $\|\cdot\|_\infty$.

We further denote

- $C^k(\mathbb{R}^d)$ the space of k -times continuously differentiable real-valued functions on \mathbb{R}^d ,
- $H^1([0, T], \mathbb{R})$ the Sobolev space of real-valued absolutely continuous functions on $[0, T]$ whose Radon-Nikodym derivative with respect to the Lebesgue measure is square-integrable.

For a path $\omega \in D([0, T], \mathbb{R}^d)$, we denote by

- $\omega(t-) = \lim_{s \rightarrow t, s < t} \omega(s)$ its left limit at t ,
- $\Delta\omega(t) = \omega(t) - \omega(t-)$ its discontinuity at t ,
- $\|\omega\|_\infty = \sup\{|\omega(t)|, t \in [0, T]\}$
- $\omega(t) \in \mathbb{R}^d$ the value of ω at t ,
- $\omega_t = \omega(t \wedge \cdot)$ the path stopped at t , and
- $\omega_{t-} = \omega \cdot 1_{[0, t[} + \omega(t-) \cdot 1_{[t, T]}$.

Note that $\omega_{t-} \in D([0, T], \mathbb{R}^d)$ is cadlag and should *not* be confused with the caglad path $u \mapsto \omega(u-)$.

For a cadlag stochastic process X we similarly denote

- $X(t)$ its value,
- $X_t = (X(u \wedge t), 0 \leq u \leq T)$ the process stopped at t , and
- $X_{t-}(u) = X(u) \cdot 1_{[0, t[}(u) + X(t-) \cdot 1_{[t, T]}(u)$.

For general definitions and concepts related to stochastic processes, we refer to the treatises by Dellacherie & Meyer [19] and Protter [62].

5 Non-anticipative functional calculus

The focus of these lectures is to define a calculus which can be used to describe the variations of interesting classes of functionals of a given reference stochastic process X . In order to cover interesting examples of processes, we allow X to have right-continuous paths with left limits, i.e. its paths lie in the space $D([0, T], \mathbb{R}^d)$ of cadlag paths. In order to include the important case of Brownian diffusion and diffusion processes, we allow these paths to have infinite variation. It is then known that the results of Newtonian calculus and Riemann-Stieltjes integration do not apply to the paths of such processes. Itô's stochastic calculus [40, 41, 19, 53, 62] provides a way out by limiting the class of integrands to *non-anticipative*, or adapted processes; this concept plays an important role in what follows.

Although the classical framework of Itô calculus is developed in a probabilistic setting, Föllmer [31] Föllmer, Hans was the first to point out that many of the ingredients at the core of this theory—in particular the Itô formula— may in fact be constructed pathwise. Föllmer [31] further identified the concept of *finite quadratic variation* as the relevant property of the path needed to derive the Itô formula.¹

In this first part, we combine the insights from Föllmer [31] with the ideas of Dupire [21] to construct a pathwise functional calculus for non-anticipative functionals defined on the space $\Omega = D([0, T], \mathbb{R}^d)$ of cadlag paths. We first introduce the notion of *non-anticipative*, or causal, functional (Section 5.1) and show how these functionals naturally arise in the representation of processes adapted to the natural filtration of a given reference process. We then introduce, following Dupire [21], the directional derivatives which form the basis of this calculus: the horizontal and the vertical derivatives, introduced in Section 5.2. The introduction of these particular directional derivatives, unnatural at first sight, is justified a posteriori by the functional change of variable formula (Section 5.3), which shows that the horizontal and vertical derivatives are precisely the quantities needed to describe the variations of a functional along a cadlag path with finite quadratic variation.

The results in this section are entirely 'pathwise' and do not make use of any probability measure. In Section 5.5, we identify important classes of stochastic processes for which these results apply almost surely.

5.1 Non-anticipative functionals

Let X be the canonical process on $\Omega = D([0, T], \mathbb{R}^d)$, and $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t \in [0, T]}$ be the filtration generated by X .

A process Y on $(\Omega, \mathcal{F}_T^0)$ adapted to \mathbb{F}^0 may be represented as a family of functionals $Y(t, \cdot) : \Omega \mapsto \mathbb{R}$ with the property that $Y(t, \cdot)$ only depends on the path stopped at t :

$$Y(t, \omega) = Y(t, \omega(\cdot \wedge t)),$$

¹This viewpoint is nicely developed in Dieter Sondermann's lecture notes [67].

so one can represent Y as

$$Y(t, \omega) = F(t, \omega_t) \quad \text{for some functional } F : [0, T] \times D([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$$

where $F(t, \cdot)$ only needs to be defined on the set of paths stopped at t .

This motivates us to view adapted processes as functionals on the space of *stopped paths*: a stopped path is an equivalent class in $[0, T] \times D([0, T], \mathbb{R}^d)$ for the following equivalence relation:

$$(t, \omega) \sim (t', \omega') \iff (t = t' \quad \text{and} \quad \omega_t = \omega_{t'}) \quad (1)$$

where $\omega_t = \omega(t \wedge \cdot)$.

The space of stopped paths can be defined as the quotient of $[0, T] \times D([0, T], \mathbb{R}^d)$ by the equivalence relation (1):

$$\Lambda_T^d = \{(t, \omega_t), (t, \omega) \in [0, T] \times D([0, T], \mathbb{R}^d)\} = [0, T] \times D([0, T], \mathbb{R}^d) / \sim.$$

We endow this set with a metric space structure by defining the distance:

$$\begin{aligned} d_\infty((t, \omega), (t', \omega')) &= \sup_{u \in [0, T]} |\omega(u \wedge t) - \omega'(u \wedge t')| + |t - t'| \\ &= \|\omega_t - \omega'_{t'}\|_\infty + |t - t'| \end{aligned} \quad (2)$$

(Λ_T^d, d_∞) is then a complete metric space and the set of continuous stopped paths

$$\mathcal{W}_T^d = \{(t, \omega) \in \Lambda_T^d, \omega \in C^0([0, T], \mathbb{R}^d)\}$$

is a closed subset of (Λ_T^d, d_∞) .

When the context is clear we will drop the superscript d and denote these spaces Λ_T, \mathcal{W}_T .

We now define a non-anticipative functional [8], as a measurable map on the space (Λ_T, d_∞) of stopped paths:

Definition 5.1 (Non-anticipative (causal) functional). A *non-anticipative functional* on $D([0, T], \mathbb{R}^d)$ is a measurable map $F : (\Lambda_T, d_\infty) \rightarrow \mathbb{R}$ on the space (Λ_T, d_∞) of stopped paths.

This notion of causality is natural when dealing with physical phenomenon as well as in control theory [30].

A non-anticipative functional may also be seen as a family $F = (F_t)_{t \in [0, T]}$ of \mathcal{F}_t^0 -measurable maps $F_t : (D([0, t], \mathbb{R}^d), \|\cdot\|_\infty) \rightarrow \mathbb{R}$. Definition 5.1 amounts to requiring joint measurability of these maps in (t, ω) .

One can alternatively represent F as a map

$$F : \bigcup_{t \in [0, T]} D([0, t], \mathbb{R}^d) \rightarrow \mathbb{R}$$

on the vector bundle $\bigcup_{t \in [0, T]} D([0, t], \mathbb{R}^d)$. This is the original point of view developed in [9, 8, 21] but leads to slightly more complicated notations. We

will follow here the definition 5.1 which has the advantage of dealing with paths defined on a fixed interval and alleviating notations.

Any progressively-measurable process Y on the filtered canonical space $(\Omega, (\mathcal{F}_t^0)_{t \in [0, T]})$ may in fact be represented [19, Vol. I] by such a non-anticipative functional F :

$$Y(t) = F(t, X(t \wedge \cdot)) = F(t, X_t). \quad (3)$$

We will write: $Y = F(X)$. Conversely, any non-anticipative functional F applied to X yields a progressively-measurable process $Y(t) = F(t, X_t)$ adapted to the filtration \mathcal{F}_t^0 .

We now define the notion of *predictable* functional as a non-anticipative functional whose value depends only on the past, but not the present value, of the path. Recall the notation:

$$\omega_{t-} = \omega 1_{[0, t[} + \omega(t-)1_{[t, T]}.$$

Definition 5.2 (Predictable functional). A non-anticipative functional $F : (\Lambda_T, d_\infty) \rightarrow \mathbb{R}$ is called *predictable* if

$$\forall (t, \omega) \in \Lambda_T, \quad F(t, \omega) = F(t, \omega_{t-}).$$

This terminology is justified by the following property: if X is a cadlag, \mathcal{F}_t -adapted process and F is a predictable functional then $Y(t) = F(t, X_t)$ defines an \mathcal{F}_t -predictable process.

Having equipped the space of stopped paths with the metric d_∞ , we can now define various notions of continuity for non-anticipative functionals.

Definition 5.3 (Joint continuity in (t, ω)). A continuous non-anticipative functional is a continuous map $F : (\Lambda_T, d_\infty) \rightarrow \mathbb{R}$: $\forall (t, \omega) \in \Lambda_T$,

$$\forall \epsilon > 0, \exists \eta > 0, \quad \forall (t', \omega') \in \Lambda_T, d_\infty((t, \omega), (t', \omega')) < \eta \Rightarrow |F(t, \omega) - F(t', \omega')| < \epsilon.$$

The set of continuous non-anticipative functionals is denoted $\mathbb{C}^{0,0}(\Lambda_T)$.

A non-anticipative functional F is said to be *continuous at fixed times* if for all $t \in [0, T[$, the map

$$F(t, \cdot) : (D([0, T], \mathbb{R}^d), \|\cdot\|_\infty) \rightarrow \mathbb{R}$$

is continuous.

The following notion, which we will use most, distinguishes the time variable and is more adapted to probabilistic applications:

Definition 5.4 (Left-continuous non-anticipative functionals). Define $\mathbb{C}_l^{0,0}(\Lambda_T)$ as the set of non-anticipative functionals F which are continuous at fixed times and which satisfy

$$\forall (t, \omega) \in \Lambda_T, \quad \forall \epsilon > 0, \exists \eta > 0, \quad \forall (t', \omega') \in \Lambda_T, \\ [t' < t \text{ and } d_\infty((t, \omega), (t', \omega')) < \eta] \Rightarrow |F(t, \omega) - F(t', \omega')| < \epsilon.$$

The image of a left-continuous path by a left-continuous functional is again left-continuous: $\forall F \in \mathbb{C}_l^{0,0}(\Lambda_T), \forall \omega \in D([0, T], \mathbb{R}^d), t \mapsto F(t, \omega_{t-})$ is left-continuous.

Let U be a cadlag \mathcal{F}_t^0 -adapted process. A non-anticipative functional F applied to U generates a process adapted to the natural filtration \mathcal{F}_t^U of U :

$$Y(t) = F(t, \{U(s), 0 \leq s \leq t\}) = F(t, U_t). \quad (4)$$

The following result is shown in [8, Proposition 1]:

Proposition 5.5. *Let $F \in \mathbb{C}_l^{0,0}(\Lambda_T)$ and U be a cadlag \mathcal{F}_t^0 -adapted process. Then*

- $Y(t) = F(t, U_{t-})$ is a left-continuous \mathcal{F}_t^0 -adapted process.
- $Z : [0, T] \times \Omega \rightarrow \mathbb{R}$ is an optional process.
- If $F \in \mathbb{C}^{0,0}(\Lambda_T)$ then $Z(t) = F(t, U_t)$ is a cadlag process, continuous at all continuity points of U .

We also introduce a notion of 'local boundedness' for functionals: we call a functional F "boundedness preserving" if it is bounded on each bounded set of paths:

Definition 5.6 (Boundedness-preserving functionals). Define $\mathbb{B}(\Lambda_T)$ as the set of non-anticipative functionals $F : \Lambda_T \rightarrow \mathbb{R}$ such that for any compact $K \subset \mathbb{R}^d$ and $t_0 < T$,

$$\exists C_{K,t_0} > 0, \forall t \leq t_0, \forall \omega \in D([0, T], \mathbb{R}^d), \omega([0, t]) \subset K \Rightarrow |F(t, \omega)| \leq C_{K,t_0}. \quad (5)$$

5.2 Horizontal and vertical derivatives

To understand the key idea behind this pathwise calculus, consider first the case of a non-anticipative functional F applied to a piecewise-constant path

$$\omega = \sum_{k=1}^n x_k 1_{[t_k, t_{k+1}[} \in D([0, T], \mathbb{R}^d).$$

Any such piecewise-constant path ω is obtained by a finite sequence of operations consisting of

- "horizontal stretching" of the path from t_k to t_{k+1} , followed by
- the addition of a jump at each discontinuity point.

In terms of the stopped path (t, ω) , these two operations correspond to

- incrementing the first component: $(t_k, \omega_{t_k}) \rightarrow (t_{k+1}, \omega_{t_k})$

- shifting the path by $(x_{k+1} - x_k)1_{[t_{k+1}, T]}$:

$$\omega_{t_{k+1}} = \omega_{t_k} + (x_{k+1} - x_k)1_{[t_{k+1}, T]}$$

The variation of a non-anticipative functional along ω can also be decomposed into the corresponding 'horizontal' and 'vertical' increments:

$$\begin{aligned} F(t_{k+1}, \omega_{t_{k+1}}) - F(t_k, \omega_{t_k}) &= \\ \underbrace{F(t_{k+1}, \omega_{t_{k+1}}) - F(t_{k+1}, \omega_{t_k})}_{\text{vertical increment}} &+ \underbrace{F(t_{k+1}, \omega_{t_k}) - F(t_k, \omega_{t_k})}_{\text{horizontal increment}} \end{aligned} \quad (6)$$

Thus, if one can control the behavior of F under these two types of path perturbations, then one can compute the variations of F along any piecewise constant path ω . If, furthermore, these operations may be controlled with respect to the supremum norm, then one can use a density argument to extend this construction to all cadlag paths.

Dupire [21] formalized this idea by introducing directional derivatives corresponding to infinitesimal versions of these variations: the *horizontal* and *vertical* derivatives, which we now define.

5.2.1 Horizontal derivative

Let us introduce the notion of *horizontal extension* of a stopped path (t, ω_t) to $[0, t+h]$: this is simply the stopped path $(t+h, \omega_t)$. Denoting $\omega_t = \omega(t \wedge \cdot)$ recall that for any non-anticipative functional,

$$\forall (t, \omega) \in [0, T] \times D([0, T], \mathbb{R}^d), \quad F(t, \omega) = F(t, \omega_t). \quad (7)$$

Definition 5.7 (Horizontal derivative). A non-anticipative functional $F : \Lambda_T \rightarrow \mathbb{R}$ is said to be horizontally differentiable at $(t, \omega) \in \Lambda_T$ if the limit

$$\mathcal{D}F(t, \omega) = \lim_{h \rightarrow 0^+} \frac{F(t+h, \omega_t) - F(t, \omega_t)}{h} \quad \text{exists.} \quad (8)$$

We will call $\mathcal{D}F(t, \omega)$ the horizontal derivative $\mathcal{D}F$ of F at (t, ω) .

Importantly, note that given the non-anticipative nature of F , the first term in the numerator in (8) depends on $\omega_t = \omega(t \wedge \cdot)$, not ω_{t+h} . If F is horizontally differentiable at all $(t, \omega) \in \Lambda_T$, then the map $\mathcal{D}F : (t, \omega) \rightarrow \mathcal{D}F(t, \omega)$ defines a non-anticipative functional which is \mathcal{F}_t^0 -measurable, without any assumption on the right-continuity of the filtration.

If $F(t, \omega) = f(t, \omega(t))$ with $f \in C^{1,1}([0, T] \times \mathbb{R}^d)$ then $\mathcal{D}F(t, \omega) = \partial_t f(t, \omega(t))$ is simply the partial (right) derivative in t : the horizontal derivative is thus an extension of the notion of 'partial derivative in time' for non-anticipative functionals.

5.2.2 Vertical derivative

We now define the *Dupire derivative* or *vertical derivative* of a non-anticipative functional [21, 8]: this derivative captures the sensitivity of a functional to a 'vertical' perturbation'

$$\omega_t^e = \omega_t + e1_{[t,T]} \quad (9)$$

of a stopped path (t, ω) .

Definition 5.8 (Vertical derivative [21]). A non-anticipative functional F is said to be *vertically differentiable* at $(t, \omega) \in \Lambda_T$ if the map

$$\begin{aligned} \mathbb{R}^d &\mapsto \mathbb{R} \\ e &\rightarrow F(t, \omega_t + e1_{[t,T]}) \end{aligned}$$

is differentiable at 0. Its gradient at 0 is called the *vertical derivative* of F at (t, ω) :

$$\begin{aligned} \nabla_\omega F(t, \omega) &= (\partial_i F(t, \omega), i = 1..d) \quad \text{where} \\ \partial_i F(t, \omega) &= \lim_{h \rightarrow 0} \frac{F(t, \omega_t + he_i 1_{[t,T]}) - F(t, \omega_t)}{h} \end{aligned} \quad (10)$$

If F is vertically differentiable at all $(t, \omega) \in \Lambda_T$ then $\nabla_\omega F$ is a non-anticipative functional called the *vertical derivative* of F .

For each $e \in \mathbb{R}^d$, $\nabla_\omega F(t, \omega).e$ is simply the directional derivative of $F(t, \cdot)$ in the direction $1_{[t,T]}e$. A similar notion of functional derivative was introduced by Fliess [30] in the context of optimal control, for causal functionals on bounded variation paths.

Note that $\nabla_\omega F(t, \omega)$ is 'local' in time: $\nabla_\omega F(t, \omega)$ only depends on the partial map $F(t, \cdot)$. However, even if $\omega \in C^0([0, T], \mathbb{R}^d)$ to compute $\nabla_\omega F(t, \omega)$ we need to compute F outside $C^0([0, T], \mathbb{R}^d)$. Also, all terms in (10) only depend on ω through $\omega_t = \omega(t \wedge \cdot)$ so, if F is vertically differentiable, then

$$\nabla_\omega F : (t, \omega) \rightarrow \nabla_\omega F(t, \omega)$$

defines a non-anticipative functional. This is due to the fact that the perturbations involved only affect the future portion of the path, in contrast, for example, with the Fréchet derivative, which involves perturbing in all directions. One may repeat this operation on $\nabla_\omega F$ and define $\nabla_\omega^2 F, \nabla_\omega^k F, \dots$. For instance, $\nabla_\omega^2 F(t, \omega)$ is defined as the gradient (if it exists) at 0 of the map

$$e \in \mathbb{R}^d \mapsto \nabla_\omega F(t, \omega + e 1_{[t,T]}) \quad (11)$$

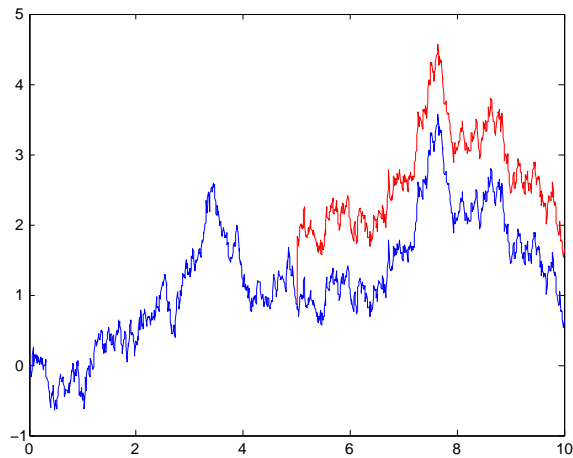


Figure 1: The **vertical** perturbation (t, ω_t^e) of the stopped path $(t, \omega) \in \Lambda_T$ in the direction $e \in \mathbb{R}^d$ is obtained by shifting the *future portion of the path* by e : $\omega_t^e = \omega_t + e1_{[t,T]}$.

5.2.3 Regular functionals

A special role is played by non-anticipative functionals which are horizontally differentiable, twice vertically differentiable and whose derivatives are left-continuous in the sense of Definition 5.4 and boundedness-preserving (in the sense of Definition 5.6):

Definition 5.9 ($\mathbb{C}_b^{1,2}$ functionals). Define $\mathbb{C}_b^{1,2}(\Lambda_T)$ as the set of left-continuous functionals $F \in \mathbb{C}_l^{0,0}(\Lambda_T)$ such that

- F admits a horizontal derivative $\mathcal{D}F(t, \omega)$ for all $(t, \omega) \in \Lambda_T$ and $\mathcal{D}F(t, \cdot) : (D([0, T], \mathbb{R}^d, \|\cdot\|_\infty) \rightarrow \mathbb{R}$ is continuous for each $t \in [0, T[$.
- $\nabla_\omega F, \nabla_\omega^2 F \in \mathbb{C}_l^{0,0}(\Lambda_T)$.
- $\mathcal{D}F, \nabla_\omega F, \nabla_\omega^2 F \in \mathbb{B}(\Lambda_T)$ (see Definition 5.6).

Similarly, we can define the class $\mathbb{C}_b^{1,k}(\Lambda_T)$.

Note that this definition only involves *certain* directional derivatives and is therefore much weaker than requiring (even first order) Fréchet or even Gâteaux-differentiability: it is easy to construct examples of $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$ for which even the first-order Fréchet derivative does not exist.

Remark 5.10. In the proofs of the key results below, one needs either left- or right-continuity of the derivatives but not both. The pathwise statements of this chapter and the results on functionals of *continuous* semimartingales hold in either case. For functionals to cadlag semimartingales, however, left- and right continuity are not interchangeable.

We now give some fundamental examples of classes of smooth functionals.

Example 5.11 (Cylindrical non-anticipative functionals). For $g \in C^0(\mathbb{R}^{n \times d})$, $h \in C^k(\mathbb{R}^d)$ with $h(0) = 0$, let

$$F(t, \omega) = h(\omega(t) - \omega(t_n -)) \mathbf{1}_{t \geq t_n} g(\omega(t_1 -), \omega(t_2 -), \dots, \omega(t_n -)).$$

Then $F \in \mathbb{C}_b^{1,k}(\Lambda_T)$ and

$$\mathcal{D}F(t, \omega) = 0, \quad \text{and} \quad \forall j = 1..k,$$

$$\nabla_\omega^j F(t, \omega) = \nabla^j h(\omega(t) - \omega(t_n -)) \mathbf{1}_{t \geq t_n} g(\omega(t_1 -), \omega(t_2 -), \dots, \omega(t_n -)).$$

We shall denote by $\mathbb{S}(\pi_n, \Lambda_T)$ the space of cylindrical non-anticipative functionals piecewise constant along π_n , i.e. of the form

$$\phi(t, \omega) = \sum_{\pi_n} g_k(\omega(t_0^n -), \omega(t_1^n -), \dots, \omega(t_k^n -)) \mathbf{1}_{[t_k^n, t_{k+1}^n[}(t)$$

where $g_k \in C^0(\mathbb{R}^{(k+1) \times d})$ and by

$$\mathbb{S}(\pi, \Lambda_T) = \bigcup_{n \geq 1} \mathbb{S}(\pi_n, \Lambda_T) \tag{12}$$

the space of all 'simple' cylindrical non-anticipative functionals.

Example 5.12 (Integral functionals). Let $g \in C^0(\mathbb{R}^d)$ and $\rho : \mathbb{R}_+ \rightarrow \mathbb{R}$ be bounded and measurable. Define

$$F(t, \omega) = \int_0^t g(\omega(u))\rho(u)du. \quad (13)$$

Then $F \in \mathbb{C}_b^{1,\infty}(\Lambda_T)$, with

$$\mathcal{D}F(t, \omega) = g(\omega(t))\rho(t) \quad \nabla_\omega^j F(t, \omega) = 0.$$

Integral functionals of the type (13) are thus 'purely horizontal' (i.e., they have zero vertical derivative) while cylindrical functionals are 'purely vertical' (they have zero horizontal derivative). We will see, in Section 5.3, that any smooth functional may in fact be decomposed into horizontal and vertical components.

Another important class of functionals are conditional expectation operators. We now give an example of smoothness result for such functionals [12]:

Example 5.13 (Weak Euler-Maruyama scheme). Let $\sigma : (\Lambda_T, d_\infty) \rightarrow \mathbb{R}^{d \times d}$ be a Lipschitz map and W a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$. Then the path-dependent SDE

$$X(t) = X(0) + \int_0^t \sigma(u, X_u)dW(u) \quad (14)$$

has a unique \mathcal{F}_t^W -adapted strong solution. The (piecewise-constant) Euler-Maruyama approximation for (14) then defines a non-anticipative functional ${}_nX$, given by the recursion

$${}_nX(t_{j+1}, \omega) = {}_nX(t_j, \omega) + \sigma(t_j, {}_nX_{t_j}(\omega)) \cdot (\omega(t_{j+1}-) - \omega(t_j-)). \quad (15)$$

For a Lipschitz functional $g : (D([0, T], \mathbb{R}^d), \|\cdot\|_\infty) \rightarrow \mathbb{R}$, consider the 'weak Euler approximation'

$$F_n(t, \omega) = \mathbb{E} [g({}_nX_T(W_T)) | \mathcal{F}_t^W] (\omega). \quad (16)$$

for the conditional expectation $\mathbb{E} [g(X_T) | \mathcal{F}_t^W]$, computed by initializing the scheme on $[0, t]$ with ω and then iterating (15) with the increments of the Wiener process between t and T . Then $F_n \in \mathbb{C}_b^{1,\infty}(\mathcal{W}_T)$ [12].

This last example implies that a large class of functionals defined as conditional expectations may be approximated in L^p norm by smooth functionals [12]. We will revisit this important point in Section 7.

If $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$, then for any $(t, \omega) \in \Lambda_T$ the map

$$g_{(t,\omega)} : e \in \mathbb{R}^d \mapsto F(t, \omega + e1_{[t,T]})$$

is twice continuously differentiable on a neighborhood of the origin and

$$\nabla_{\omega} F(t, \omega) = \nabla g_{(t, \omega)}(0), \quad \text{and} \quad \nabla_{\omega}^2 F(t, \omega) = \nabla^2 g_{(t, \omega)}(0).$$

A second-order Taylor expansion of the map $g_{(t, \omega)}$ at the origin yields that any $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$ admits a second-order Taylor expansion with respect to a vertical perturbation:

$$\begin{aligned} & \forall (t, \omega) \in \Lambda_T, \forall e \in \mathbb{R}^d, \\ F(t, \omega + e1_{[t, T]}) &= F(t, \omega) + \nabla_{\omega} F(t, \omega) \cdot e + \langle e, \nabla_{\omega}^2 F(t, \omega) \cdot e \rangle + o(\|e\|^2) \end{aligned}$$

Schwarz's theorem applied to $g_{(t, \omega)}$ then entails that $\nabla_{\omega}^2 F(t, \omega)$ is a symmetric matrix. As is well known, the assumption of continuity of the derivatives cannot be relaxed in Schwarz's theorem, so one can easily construct counterexamples where $\nabla_{\omega}^2 F(t, \omega)$ exists but is not symmetric by removing the continuity requirement on $\nabla_{\omega}^2 F$.

However, unlike the usual partial derivatives in finite dimensions, the horizontal and vertical derivative do not commute: in general

$$\mathcal{D}(\nabla_{\omega} F) \neq \nabla_{\omega}(\mathcal{D}F).$$

This stems from the fact that the elementary operations of 'horizontal extension' and 'vertical perturbation' defined above do not commute: a horizontal extension of the stopped path (t, ω) to $t+h$ followed by a vertical perturbation yields the path $\omega_t + e1_{[t+h, T]}$ while a vertical perturbation at t followed by a horizontal extension to $t+h$ yields

$$\omega_t + e1_{[t, T]} \neq \omega_t + e1_{[t+h, T]}.$$

Note that these two paths have the same value at $t+h$, so only functionals which are truly path-dependent (as opposed to functions of the path at a single point in time) will be affected by this lack of commutativity. Thus, the 'functional Lie bracket'

$$[\mathcal{D}, \nabla_{\omega}]F = \mathcal{D}(\nabla_{\omega} F) - \nabla_{\omega}(\mathcal{D}F)$$

may be used to quantify the 'path-dependency' of F .

Example 5.14. Let F be the integral functional given in Example 5.12, with $g \in C^1(\mathbb{R}^d)$. Then, $\nabla_{\omega} F(t, \omega) = 0$ so $\mathcal{D}(\nabla_{\omega} F) = 0$. However, $\mathcal{D}F(t, \omega) = g(\omega(t))$ so

$$\nabla_{\omega} \mathcal{D}F(t, \omega) = \nabla g(\omega(t)) \neq \mathcal{D}(\nabla_{\omega} F)(t, \omega) = 0.$$

Locally regular functionals Many examples of functionals, especially those involving exit times, may fail to be globally smooth but their derivatives may still be well behaved except at certain stopping times. The following is a prototypical example of a functional involving exit times:

Example 5.15 (A functional involving exit times). Let W be real Brownian motion, $b > 0$, $M(t) = \sup_{0 \leq s \leq t} W(s)$. Consider the \mathcal{F}_t^W -adapted martingale:

$$Y(t) = E[1_{M(T) \geq b} | \mathcal{F}_t^W]. \quad (17)$$

Then Y has the functional representation $Y(t) = F(t, W_t)$ where

$$F(t, \omega) = 1_{\sup_{0 \leq s \leq t} \omega(s) \geq b} + 1_{\sup_{0 \leq s \leq t} \omega(s) < b} \left[2 - 2N \left(\frac{b - \omega(t)}{\sqrt{T - t}} \right) \right] \quad (18)$$

where N is the $N(0, 1)$ distribution. $F \notin \mathbb{C}_t^{0,0}$: a path ω_t where $\omega(t) < b$ but $\sup_{0 \leq s \leq t} \omega(s) = b$ can be approximated in sup norm by paths where $\sup_{0 \leq s \leq t} \omega(s) < b$. However, one can easily check that $\nabla_\omega F$, $\nabla_\omega^2 F$ and $\mathcal{D}F$ exist almost everywhere.

Recall that a stopping time (or non-anticipative random time) on $(\Omega, (\mathcal{F}_t^0)_{t \in [0, T]})$ is a measurable map $\tau : \Omega \rightarrow [0, \infty)$ such that

$$\forall t \geq 0, \{\omega \in \Omega, \tau(\omega) \leq t\} \in \mathcal{F}_t^0.$$

In the example above, regularity holds except on the graph of a certain stopping time. This motivates the following definition:

Definition 5.16 ($\mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$). $F \in \mathbb{C}_b^{0,0}(\Lambda_T)$ is said to be **locally regular** if there exists an increasing sequence $(\tau_k)_{k \geq 0}$ of stopping times with $\tau_0 = 0$, $\tau_k \uparrow \infty$ and $F^k \in \mathbb{C}_b^{1,2}(\Lambda_T)$ such that

$$F(t, \omega) = \sum_{k \geq 0} F^k(t, \omega) 1_{[\tau_k(\omega), \tau_{k+1}(\omega)]}(t)$$

Note that $\mathbb{C}_b^{1,2}(\Lambda_T) \subset \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$ but the notion of local regularity allows discontinuities or explosions at the times described by $(\tau_k, k \geq 1)$.

Revisiting Example 5.15, we can see that Definition 5.16 applies: recall that

$$F(t, \omega) = 1_{\sup_{0 \leq s \leq t} \omega(s) \geq b} + 1_{\sup_{0 \leq s \leq t} \omega(s) < b} \left[2 - 2N \left(\frac{b - \omega(t)}{\sqrt{T - t}} \right) \right]. \quad (19)$$

If we define

$$\begin{aligned} \tau_1(\omega) &= \inf\{t \geq 0 | \omega(t) = b\} \wedge T, & F^0(t, \omega) &= 2 - 2N \left(\frac{b - \omega(t)}{\sqrt{T - t}} \right), \\ \tau_i(\omega) &= T + i - 2, & \text{for } i \geq 2; & F^i(t, \omega) = 1, i \geq 1, \end{aligned}$$

then $F^i \in \mathbb{C}_b^{1,2}(\Lambda_T)$, so $F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$.

5.3 Pathwise integration and functional change of variable formula

In his seminal paper *Calcul d'Itô sans probabilités* [31], Hans Föllmer proposed a non-probabilistic version of the Itô formula: Föllmer showed that if a cadlag

(right continuous with left limits) function x has finite quadratic variation along some sequence $\pi_n = (0 = t_0^n < t_1^n \dots < t_{k(n)}^n = T)$ of partitions of $[0, T]$ with step size decreasing to zero, then for $f \in C^2(\mathbb{R}^d)$ one can define the pathwise integral

$$\int_0^T \nabla f(x(t)) d^\pi x = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \nabla f(x(t_i^n)) \cdot (x(t_{i+1}^n) - x(t_i^n)) \quad (20)$$

as a limit of Riemann sums along the sequence $\pi = (\pi_n)_{n \geq 1}$ of partitions and obtain a change of variable formula for this integral [31]. We now revisit the approach of Föllmer and show how it may be combined with Dupire's directional derivatives to obtain a pathwise change of variable formula for functionals in $\mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$.

5.3.1 Quadratic variation of a path along a sequence of partitions

We first define the notion of quadratic variation of a path along a sequence of partitions. Our presentation is different from Föllmer [31] but can be shown to be mathematically equivalent.

Throughout this section we denote by $\pi = (\pi_n)_{n \geq 1}$ a sequence of partitions of $[0, T]$ into intervals:

$$\pi_n = (0 = t_0^n < t_1^n \dots < t_{k(n)}^n = T).$$

$|\pi_n| = \sup\{|t_{i+1}^n - t_i^n|, i = 1..k(n)\}$ will denote the mesh size of the partition.

As an example, one can keep in mind the *dyadic partition*, for which $t_i^n = iT/2^n, i = 0..k(n) = 2^n$, and $|\pi_n| = 2^{-n}$. This is an example of a *nested* sequence of partitions: for $n \geq m$, every interval $[t_i^n, t_{i+1}^n]$ of the partition π_n is included in one of the intervals of π_m . Unless specified, we will assume that the sequence π_n is a nested sequence of partitions.

Definition 5.17 (Quadratic variation of a path along a sequence of partitions). Let $\pi_n = (0 = t_0^n < t_1^n \dots < t_{k(n)}^n = T)$ be a sequence of partitions of $[0, T]$ with step size decreasing to zero. A càdlàg path $x \in D([0, T], \mathbb{R})$ is said to have finite quadratic variation along the sequence of partitions $(\pi_n)_{n \geq 1}$ if for any $t \in [0, T]$ the limit

$$[x](t) := \lim_{n \rightarrow \infty} \sum_{t_{i+1}^n \leq t} (x(t_{i+1}^n) - x(t_i^n))^2 < \infty \quad (21)$$

exists and the increasing function $[x]$ has Lebesgue decomposition

$$[x]_\pi(t) = [x]_\pi^c(t) + \sum_{0 < s \leq t} |\Delta x(s)|^2$$

where $[x]_\pi^c$ is a continuous, increasing function.

The increasing function $[x] : [0, T] \rightarrow \mathbb{R}_+$ defined by (21) is then called the *quadratic variation of the path x along the sequence of partitions $\pi = (\pi_n)_{n \geq 1}$* and $[x]_\pi^c$ is the continuous quadratic variation of x along π .

Note that the sequence of sums in (21) need not be a monotone sequence and its convergence is far from obvious in general.

In general the quadratic variation of a path x along a sequence of partitions π depends on the choice of the sequence π , as the following example shows.

Example 5.18. Let $\omega \in C^0([0, 1], \mathbb{R})$ be an *arbitrary* continuous function. lectutees

This example shows that ‘having finite quadratic variation along *some* sequence of partitions’ is not an interesting property and that the notion of quadratic variation along a sequence of partitions depends on the chosen partition. Definition 5.19 becomes non-trivial only if one *fixes* the partition beforehand. In the sequel, we **fix** a sequence $\pi = (\pi_n, n \geq 1)$ of partitions with $|\pi_n| \rightarrow 0$ and all limits will be considered along the *same* sequence π , thus enabling us to drop the subscript in $[x]_\pi$ whenever the context is clear.

The notion of quadratic variation along a sequence of partitions is different from the p-variation of the path ω for $p = 2$: the p-variation involves taking a supremum over *all* partitions, not necessarily formed of intervals, whereas (21) is a limit taken along a given sequence $(\pi_n)_{n \geq 1}$. In general $[x]_\pi$ given by (21) is smaller than the 2-variation and there are many situations where the 2-variation is infinite while the limit in (21) is finite. This is in fact the case for instance for typical paths of Brownian motion, which have finite quadratic variation along any sequence of partitions with mesh size $o(1/\log n)$ [20] but have infinite p-variation almost surely for $p \leq 2$ [73]!

The extension of this notion to vector-valued paths is somewhat subtle [31]:

Definition 5.19. A d-dimensional path $x = (x^1, \dots, x^d) \in D([0, T], \mathbb{R}^d)$ is said to have finite quadratic variation along $\pi = (\pi_n)_{n \geq 1}$ if $x^i \in Q^\pi([0, T], \mathbb{R})$ and $x^i + x^j \in Q^\pi([0, T], \mathbb{R})$ for all $i, j = 1..d$ and $t \in [0, T]$, we have

$$\sum_{t_k^n \in \pi_n, t_k^n \leq t} (x^i(t_{k+1}^n) - x^i(t_k^n)) \cdot (x^j(t_{k+1}^n) - x^j(t_k^n)) \xrightarrow{n \rightarrow \infty} [x]_{ij}(t) = \frac{[x^i + x^j](t) - [x^i](t) - [x^j](t)}{2}.$$

The matrix-valued function $[x] : [0, T] \rightarrow S_d^+$ whose elements are given by

$$[x]_{ij}(t) = \frac{[x^i + x^j](t) - [x^i](t) - [x^j](t)}{2}$$

is called the *quadratic covariation* of the path x : for any $t \in [0, T]$,

$$\sum_{t_i^n \in \pi_n, t_n \leq t} (x(t_{i+1}^n) - x(t_i^n)) \cdot (x(t_{i+1}^n) - x(t_i^n)) \xrightarrow{n \rightarrow \infty} [x](t) \in S_d^+ \quad (22)$$

and $[x]$ is increasing in the sense of the order on positive symmetric matrices: for $h \geq 0$, $[x](t+h) - [x](t) \in S_d^+$.

We denote $Q^\pi([0, T], \mathbb{R}^d)$ the set of \mathbb{R}^d -valued cadlag paths with finite quadratic variation with respect to the partition $\pi = (\pi_n)_{n \geq 1}$.

Remark 5.20. Note that Definition 5.19 requires that $x^i + x^j \in Q^\pi([0, T], \mathbb{R})$: this does not necessarily follow from requiring $x^i, x^j \in Q^\pi([0, T], \mathbb{R})$. Indeed, denoting $\delta x^i = x^i(t_{k+1}) - x^i(t_k)$, we have

$$|\delta x^i + \delta x^j|^2 = |\delta x^i|^2 + |\delta x^j|^2 + 2\delta x^i \delta x^j$$

and the cross-product may be positive, negative, or have an oscillating sign which may prevent convergence of the series $\sum_{\pi_n} \delta x^i \delta x^j$ (for counterexamples, see [64] Schied, Alexander). However, if x^i, x^j are differentiable functions of the *same* path ω i.e. $x^i = f_i(\omega)$ with $f_i \in C^1(\mathbb{R}^d, \mathbb{R})$ then

$$\delta x^i \delta x^j = f'_i(\omega(t_k^n)) f'_j(\omega(t_k^n)) |\delta \omega|^2 + o(|\delta \omega|^2)$$

so $\sum_{\pi_n} \delta x^i \delta x^j$ converges and

$$\lim_{n \rightarrow \infty} \sum_{\pi_n} \delta x^i \delta x^j = [x^i, x^j]$$

is well-defined. This remark is connected to the notion of 'controlled rough path' introduced by Gubinelli [35], see Remark 5.25 below.

For any $x \in Q^\pi([0, T], \mathbb{R}^d)$, since $[x] : [0, T] \rightarrow S_d^+$ is increasing (in the sense of the order on S_d^+), we can define the Riemann-Stieltjes integral $\int_0^T f d[x]$ for any $f \in C_b^0([0, T])$. A key property of $Q^\pi([0, T], \mathbb{R}^d)$ is the following:

Proposition 5.21 (Uniform convergence of quadratic Riemann sums).

$$\forall \omega \in Q^\pi([0, T], \mathbb{R}^d), \forall h \in C_b^0([0, T], \mathbb{R}^{d \times d}), \forall t \in [0, T],$$

$$\sum_{t_i^n \in \pi_n, t_i^n \leq t} \text{tr} (h(t_i^n) (\omega(t_{i+1}^n) - \omega(t_i^n))^t (\omega(t_{i+1}^n) - \omega(t_i^n))) \xrightarrow{n \rightarrow \infty} \int_0^t \langle h, d[\omega] \rangle,$$

where we use the notation $\langle A, B \rangle = \text{tr}(A \cdot B)$ for $A, B \in \mathbb{R}^{d \times d}$.

Furthermore, if $\omega \in C^0([0, T], \mathbb{R}^d)$, the convergence is uniform in $t \in [0, T]$.

Proof. It suffices to show this property for $d = 1$. Let $\omega \in Q^\pi([0, T], \mathbb{R})$, $h \in D([0, T], \mathbb{R})$. Then the integral $\int_0^t h d[\omega]$ may be defined as a limit of Riemann sums:

$$\int_0^t h d[\omega] = \lim_{n \rightarrow \infty} \sum_{\pi_n} h(t_i^n) ([\omega](t_{i+1}^n) - [\omega](t_i^n)).$$

Using the definition of $[\omega]$, the result is thus true for $h : [0, T] \rightarrow \mathbb{R}$ of the form $h = \sum_{\pi_n} a_k 1_{[t_k^n, t_{k+1}^n[}$. Consider now $h \in C_b^0([0, T], \mathbb{R})$ and define the piecewise constant approximations

$$h^n = \sum_{\pi_n} h(t_k^n) 1_{[t_k^n, t_{k+1}^n[}$$

Then h^n converges uniformly to h : $\|h - h^n\|_\infty \rightarrow 0$ as $n \rightarrow \infty$, and for each $n \geq 1$,

$$\sum_{t_i^k \in \pi_k, t_i^k \leq t} h^n(t_i^k) (\omega(t_{i+1}^k) - \omega(t_i^k))^2 \xrightarrow{k \rightarrow \infty} \int_0^t h^n d[\omega].$$

Since h is bounded on $[0, T]$ this sequence is dominated; we can then conclude using a diagonal convergence argument. \square

Proposition 5.21 implies weak convergence on $[0, T]$ of the discrete measures

$$\xi^n = \sum_{i=0}^{k(n)-1} (\omega(t_{i+1}^n) - \omega(t_i^n))^2 \delta_{t_i^n} \xrightarrow{n \rightarrow \infty} \xi = d[\omega] \quad (23)$$

where δ_t is the Dirac measure (point mass) at t .²

5.3.2 Functional change of variable formula

Consider now a path $\omega \in Q^\pi([0, T], \mathbb{R}^d)$ with finite quadratic variation along π . Since ω has at most a countable set of jump times, we may always assume that the partition 'exhausts' the jump times in the sense that

$$\sup_{t \in [0, T] - \pi_n} |\omega(t) - \omega(t-)| \xrightarrow{n \rightarrow \infty} 0. \quad (24)$$

Then the piecewise-constant approximation

$$\omega^n(t) = \sum_{i=0}^{k(n)-1} \omega(t_{i+1}-) 1_{[t_i, t_{i+1}[}(t) + \omega(T) 1_{\{T\}}(t) \quad (25)$$

converges uniformly to ω :

$$\sup_{t \in [0, T]} \|\omega^n(t) - \omega(t)\| \xrightarrow{n \rightarrow \infty} 0. \quad (26)$$

Note that with the notation (25), $\omega^n(t_i^n-) = \omega(t_i^n-)$ but $\omega^n(t_i^n) = \omega(t_{i+1}^n-)$. If we define

$$\omega^{n, \Delta\omega(t_i^n)} = \omega^n + \Delta\omega(t_i^n) 1_{[t_i^n, T]}, \quad \text{then} \quad \omega_{t_i^n-}^{n, \Delta\omega(t_i^n)}(t_i^n) = \omega(t_i^n).$$

By decomposing the variations of the functional into vertical and horizontal increments along the partition π_n , we obtain the following pathwise change of variable formula for $\mathbb{C}^{1,2}(\Lambda_T)$ functionals, derived in [8]:

²In fact this weak convergence property was used by Föllmer [31] Föllmer, Hans as the definition of pathwise quadratic variation. We use the more natural definition (Def. 5.17) of which Proposition 5.21 is a consequence.

Theorem 5.22 (Pathwise change of variable formula for $\mathbb{C}^{1,2}$ functionals [8]).
Let $\omega \in Q^\pi([0, T], \mathbb{R}^d)$ verifying (24). Then for any $F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$ the limit

$$\int_0^T \nabla_\omega F(t, \omega_{t-}) d^\pi \omega := \lim_{n \rightarrow \infty} \sum_{i=0}^{k(n)-1} \nabla_\omega F\left(t_i^n, \omega_{t_i^n-}^{n, \Delta\omega(t_i^n)}\right) \cdot (\omega(t_{i+1}^n) - \omega(t_i^n)) \quad (27)$$

exists and

$$\begin{aligned} F(T, \omega_T) - F(0, \omega_0) &= \int_0^T \mathcal{D}F(t, \omega_t) dt + \int_0^T \frac{1}{2} \text{tr} \left({}^t \nabla_\omega^2 F(t, \omega_{t-}) d[\omega]^c(t) \right) \\ &+ \int_0^T \nabla_\omega F(t, \omega_{t-}) d^\pi \omega + \sum_{t \in]0, T]} [F(t, \omega_t) - F(t, \omega_{t-}) - \nabla_\omega F(t, \omega_{t-}) \cdot \Delta\omega(t)]. \end{aligned} \quad (28)$$

The detailed proof of Theorem 5.22 may be found in [8] under more general assumptions. Here we reproduce a simplified version of this proof, in the case where ω is continuous.

Proof. First, we note that up to localization by a sequence of stopping times, we can assume that $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$, which we shall do in the sequel. Denote $\delta\omega_i^n = \omega(t_{i+1}^n) - \omega(t_i^n)$. Since ω is continuous on $[0, T]$, it is uniformly continuous so

$$\eta_n = \sup\{|\omega(u) - \omega(t_{i+1}^n)| + |t_{i+1}^n - t_i^n|, 0 \leq i \leq k(n) - 1, u \in [t_i^n, t_{i+1}^n]\} \xrightarrow{n \rightarrow \infty} 0$$

Since $\nabla_\omega^2 F, \mathcal{D}F$ satisfy the boundedness-preserving property (5), for n sufficiently large there exists $C > 0$ such that

$$\forall t < T, \forall \omega' \in \Lambda_T, \quad d_\infty((t, \omega), (t', \omega')) < \eta_n \Rightarrow |\mathcal{D}F(t', \omega')| \leq C, |\nabla_\omega^2 F(t', \omega')| \leq C$$

For $i \leq k(n) - 1$, consider the decomposition of increments into 'horizontal' and 'vertical' terms:

$$\begin{aligned} F(t_{i+1}^n, \omega_{t_{i+1}^n-}^n) - F(t_i^n, \omega_{t_i^n-}^n) &= F(t_{i+1}^n, \omega_{t_{i+1}^n-}^n) - F(t_i^n, \omega_{t_i^n-}^n) \\ &+ F(t_i^n, \omega_{t_i^n-}^n) - F(t_i^n, \omega_{t_i^n-}^n) \end{aligned} \quad (29)$$

The first term in (29) can be written $\psi(h_i^n) - \psi(0)$ where $h_i^n = t_{i+1}^n - t_i^n$ and

$$\psi(u) = F(t_i^n + u, \omega_{t_i^n-}^n). \quad (30)$$

Since $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$, ψ is right-differentiable. Moreover by Proposition 5.5, ψ is left-continuous, so:

$$F(t_{i+1}^n, \omega_{t_{i+1}^n-}^n) - F(t_i^n, \omega_{t_i^n-}^n) = \int_0^{t_{i+1}^n - t_i^n} \mathcal{D}F(t_i^n + u, \omega_{t_i^n-}^n) du \quad (31)$$

The second term in (29) can be written $\phi(\delta\omega_i^n) - \phi(0)$, where:

$$\phi(u) = F(t_i^n, \omega_{t_i^n-}^n + u1_{[t_i^n, T]}) \quad (32)$$

Since $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$, $\phi \in C^2(\mathbb{R}^d)$ with:

$$\nabla\phi(u) = \nabla_\omega F(t_i^n, \omega_{t_i^n-}^n + u1_{[t_i^n, T]}), \quad \nabla^2\phi(u) = \nabla_\omega^2 F(t_i^n, \omega_{t_i^n-}^n + u1_{[t_i^n, T]}).$$

A second order Taylor expansion of ϕ at $u = 0$ yields

$$F(t_i^n, \omega_{t_i^n}^n) - F(t_i^n, \omega_{t_i^n-}^n) = \nabla_\omega F(t_i^n, \omega_{t_i^n-}^n) \delta\omega_i^n + \frac{1}{2} \text{tr} \left(\nabla_\omega^2 F(t_i^n, \omega_{t_i^n-}^n) \cdot {}^t \delta\omega_i^n \delta\omega_i^n \right) + r_i^n$$

where r_i^n is bounded by

$$K |\delta\omega_i^n|^2 \sup_{x \in B(0, \eta_n)} |\nabla_\omega^2 F(t_i^n, \omega_{t_i^n-}^n + x1_{[t_i^n, T]}) - \nabla_\omega^2 F(t_i^n, \omega_{t_i^n-}^n)| \quad (33)$$

Denote $i^n(t)$ the index such that $t \in [t_{i^n(t)}^n, t_{i^n(t)+1}^n[$. We now sum all the terms above from $i = 0$ to $k(n) - 1$:

- The left-hand side of (29) yields $F(T, \omega_{T-}^n) - F(0, \omega_0^n)$, which converges to $F(T, \omega_{T-}) - F(0, \omega_0)$ by left-continuity of F and this quantity equals $F(T, \omega_T) - F(0, \omega_0)$ since ω is continuous.
- The first line in the right-hand side can be written:

$$\int_0^T \mathcal{D}F(u, \omega_{t_{i^n(u)}^n}^n) du \quad (34)$$

where the integrand converges to $\mathcal{D}F(u, \omega_u)$ and is bounded by C . Hence the dominated convergence theorem applies and (34) converges to:

$$\int_0^T \mathcal{D}F(u, \omega_u) du \quad (35)$$

- The second line can be written:

$$\sum_{i=0}^{k(n)-1} \nabla_\omega F(t_i^n, \omega_{t_i^n-}^n) \cdot (\omega_{t_{i+1}^n}^n - \omega_{t_i^n}^n) + \sum_{i=0}^{k(n)-1} \frac{1}{2} \text{tr} \left(\nabla_\omega^2 F(t_i^n, \omega_{t_i^n-}^n) \cdot {}^t \delta\omega_i^n \delta\omega_i^n \right) + \sum_{i=0}^{k(n)-1} r_i^n.$$

The term $\nabla_\omega^2 F(t_i^n, \omega_{t_i^n-}^n) 1_{]t_i^n, t_{i+1}^n]}$ is bounded by C , and converges to $\nabla_\omega^2 F(t, \omega_t)$ by left-continuity of $\nabla_\omega^2 F$ and the paths of both are left-continuous by Proposition 5.5. We now use a 'diagonal lemma' for weak convergence of measures [8]:

Lemma 5.23 ([8]). *Let $(\mu_n)_{n \geq 1}$ be a sequence of Radon measures on $[0, T]$ converging weakly to a Radon measure μ with no atoms, and $(f_n)_{n \geq 1}$, f be left-continuous functions on $[0, T]$ with*

$$\forall t \in [0, T], \lim_n \int_s^t f_n d\mu_n = \int_s^t f d\mu \quad \forall t \in [0, T], \|f_n(t)\| \leq K \quad (36)$$

$$\text{then} \quad \int_s^t f_n d\mu_n \xrightarrow{n \rightarrow \infty} \int_s^t f d\mu. \quad (37)$$

Applying the above lemma to the second term in the sum, we obtain:

$$\int_0^T \frac{1}{2} \text{tr} \left({}^t \nabla_{\omega}^2 F(t_i^n, \omega_{t_i^n}^n) \cdot d\xi^n \right) \xrightarrow{n \rightarrow \infty} \int_0^T \frac{1}{2} \text{tr} \left({}^t \nabla_{\omega}^2 F(u, \omega_u) \cdot d[\omega](u) \right)$$

Using the same lemma, since $|r_i^n|$ is bounded by $\epsilon_i^n |\delta \omega_i^n|^2$ where ϵ_i^n converges to 0 and is bounded by C ,

$$\sum_{i=i^n(s)+1}^{i^n(t)-1} r_i^n \xrightarrow{n \rightarrow \infty} 0.$$

Since all other terms converge as $n \rightarrow \infty$, we conclude that the limit of the 'Riemann sums'

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{k(n)-1} \nabla_{\omega} F(t_i^n, \omega_{t_i^n}^n) (\omega(t_{i+1}^n) - \omega(t_i^n))$$

exists: this is the pathwise integral $\int \nabla_{\omega} F(t, \omega) \cdot d^{\pi} \omega$. \square

5.3.3 Pathwise integration for paths of finite quadratic variation

A by-product of Theorem (5.22) is that we can define, for $\omega \in Q^{\pi}([0, T], \mathbb{R}^d)$, the pathwise integral $\int_0^T \phi d^{\pi} \omega$ as a limit of non-anticipative Riemann sums:

$$\int_0^T \phi \cdot d^{\pi} \omega := \lim_{n \rightarrow \infty} \sum_{i=0}^{k(n)-1} \phi \left(t_i^n, \omega_{t_i^n}^n, \Delta \omega(t_i^n) \right) \cdot (\omega(t_{i+1}^n) - \omega(t_i^n)) \quad (38)$$

for any integrand of the form

$$\phi(t, \omega) = \nabla_{\omega} F(t, \omega)$$

where $F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$, without requiring that ω be of finite variation. This construction extends Föllmer's pathwise integral, defined in [31] for integrands of the form $\phi = \nabla f \circ \omega$ with $f \in C^2(\mathbb{R}^d)$, to path-dependent integrands of the form $\phi(t, \omega)$ where ϕ belongs to the space:

$$V(\Lambda_T) = \{ \nabla_{\omega} F(\cdot, \cdot), F \in \mathbb{C}_b^{1,2}(\Lambda_T^d) \} \quad (39)$$

where Λ_T^d denotes the space of \mathbb{R}^d -valued stopped cadlag paths. We call such integrands *vertical 1-forms*. Since, as noted before, the horizontal and vertical derivatives do not commute, $V(\Lambda_T^d)$ does not coincide with $\mathbb{C}_b^{1,1}(\Lambda_T^d)$.

This set of integrands has a natural vector space structure and includes as subsets the space $\mathbb{S}(\Lambda_T)$ of simple predictable cylindrical functionals as well as Föllmer's space of integrands $\{ \nabla f, f \in C^2(\mathbb{R}^d, \mathbb{R}) \}$. For $\phi = \nabla_{\omega} F \in V(\Lambda_T^d)$,

the pathwise integral $\int_0^t \phi(\cdot, \omega_{\cdot-}) \cdot d^\pi \omega$ is in fact given by

$$\begin{aligned} \int_0^T \phi(t, \omega_{t-}) \cdot d^\pi \omega &= F(T, \omega_T) - F(0, \omega_0) - \frac{1}{2} \int_0^T \langle \nabla_\omega \phi(t, \omega_{t-}), d[\omega]_\pi \rangle \\ &\quad - \int_0^T \mathcal{D}F(t, \omega_{t-}) dt - \sum_{0 \leq s \leq T} F(t, \omega_t) - F(t, \omega_{t-}) - \phi(t, \omega_{t-}) \cdot \Delta \omega(t) \end{aligned} \quad (40)$$

The following proposition, whose proof is given in [6], summarizes some key properties of this integral:

Proposition 5.24. *Let $\omega \in Q^\pi([0, T], \mathbb{R}^d)$. The pathwise integral (40) defines a map*

$$\begin{aligned} I_\omega : V(\Lambda_T^d) &\mapsto Q^\pi([0, T], \mathbb{R}^d) \\ \phi &\rightarrow \int_0^\cdot \phi(t, \omega_{t-}) \cdot d^\pi \omega(t) \end{aligned}$$

with the following properties:

1. *Pathwise isometry formula:* $\forall \phi \in \mathbb{S}(\pi, \Lambda_T^d), \forall t \in [0, T]$,

$$[I_\omega(\phi)]_\pi(t) = \left[\int_0^\cdot \phi(t, \omega_{t-}) \cdot d^\pi \omega \right]_\pi(t) = \int_0^t \langle \phi(u, \omega_{u-})^t \phi(u, \omega_{u-}), d[\omega] \rangle. \quad (41)$$

2. *Quadratic covariation formula:* for $\phi, \psi \in \mathbb{S}(\pi, \Lambda_T^d)$, the limit

$$[I_\omega(\phi), I_\omega(\psi)]_\pi(T) := \lim_{n \rightarrow \infty} \sum_{\pi_n} (I_\omega(\phi)(t_{k+1}^n) - I_\omega(\phi)(t_k^n)) (I_\omega(\psi)(t_{k+1}^n) - I_\omega(\psi)(t_k^n))$$

exists and is given by

$$[I_\omega(\phi), I_\omega(\psi)]_\pi(T) = \int_0^T \langle \psi(t, \omega_{t-})^t \phi(t, \omega_{t-}), d[\omega] \rangle. \quad (42)$$

3. *Associativity:* Let $\phi \in V(\Lambda_T^d), \psi \in V(\Lambda_T^1)$ and $x \in D([0, T], \mathbb{R})$ defined by $x(t) = \int_0^t \phi(u, \omega_{u-}) \cdot d^\pi \omega$. Then

$$\int_0^T \psi(t, x_{t-}) \cdot d^\pi x = \int_0^T \psi(t, (\int_0^t \phi(u, \omega_{u-}) \cdot d^\pi \omega)_{t-}) \phi(t, \omega_{t-}) d^\pi \omega. \quad (43)$$

This pathwise integration has interesting applications in mathematical finance [32, 13] and stochastic control [32], where integrals of such vertical 1-forms naturally appear as hedging strategies [13] or optimal control policies and pathwise interpretations of quantities are arguably necessary to interpret the results in terms of the original problem.

Unlike $Q^\pi([0, T], \mathbb{R}^d)$ itself, which does not have a vector space structure, the space

$$\mathbb{C}_b^{1,2}(\omega) = \{F(\cdot, \omega), \quad F \in \mathbb{C}_b^{1,2}(\Lambda_T^d)\} \subset Q^\pi([0, T], \mathbb{R}^d)$$

is a vector space of paths with finite quadratic variation whose properties are 'controlled' by ω , on which the quadratic variation along the sequence of partitions π is well-defined. This space, and not $Q^\pi([0, T], \mathbb{R}^d)$, is the appropriate starting point for the studying the pathwise calculus developed here.

Remark 5.25 (Relation with 'rough path' theory). Integrand of the form $\nabla_\omega F(t, \omega)$ with $F \in \mathbb{C}_b^{1,2}$ may be viewed as 'controlled rough paths' in the sense of Gubinelli [35]: their increments are 'controlled' by those of ω . However, unlike the approach of rough path theory [49, 35], the pathwise integration defined here does not resort to the use of p-variation norms on iterated integrals: convergence of Riemann sums is pointwise (and, for continuous paths, uniform in t). The reason is that the obstruction to integration posed by the Lévy area, which is the focus of rough path theory, vanishes when considering integrands which are vertical 1-forms. Fortunately, all integrands one encounters when applying the change of variable formula, as well as in applications involving optimal control, hedging,... are precisely of the form (39)! This observation simplifies the approach and, importantly, yields an integral which may be expressed as a limit of (ordinary) Riemann sums, which have an intuitive (physical, financial, etc.) interpretation, without resorting to 'rough path' integrals, whose interpretation is less intuitive.

If ω has finite variation then the Föllmer integral reduces to the Riemann-Stieltjes integral and we obtain:

Proposition 5.26. *For any $F \in \mathbb{C}_{\text{loc}}^{1,1}(\Lambda_T)$, $\omega \in BV([0, T]) \cap D([0, T], \mathbb{R}^d)$,*

$$\begin{aligned} F(T, \omega_T) - F(0, \omega_0) &= \int_0^T \mathcal{D}F(t, \omega_{t-}) du + \int_0^T \nabla_\omega F(t, \omega_{t-}) d\omega \\ &+ \sum_{t \in [0, T]} [F(t, \omega_t) - F(t, \omega_{t-}) - \nabla_\omega F(t, \omega_{t-}) \cdot \Delta\omega(t)] \end{aligned} \quad (44)$$

where the integrals are defined as limits of Riemann sums along any sequence of partitions $(\pi_n)_{n \geq 1}$ with $|\pi_n| \rightarrow 0$.

In particular, if ω is continuous with finite variation we have:
 $\forall F \in \mathbb{C}_{\text{loc}}^{1,1}(\Lambda_T)$, $\forall \omega \in BV([0, T]) \cap C^0([0, T], \mathbb{R}^d)$,

$$F(T, \omega_T) - F(0, \omega_0) = \int_0^T \mathcal{D}F(t, \omega_t) dt + \int_0^T \nabla_\omega F(t, \omega_t) d\omega. \quad (45)$$

Thus the restriction of any functional $F \in \mathbb{C}_{\text{loc}}^{1,1}(\Lambda_T)$ to $BV([0, T]) \cap C^0([0, T], \mathbb{R}^d)$ may be decomposed into 'horizontal' and 'vertical' components.

5.4 Functionals defined on continuous paths

Consider now an \mathbb{F} -adapted process $(Y(t))_{t \in [0, T]}$ given by a functional representation

$$Y(t) = F(t, X_t) \quad (46)$$

where $F \in \mathbb{C}_l^{0,0}(\Lambda_T)$ has left-continuous horizontal and vertical derivatives $\mathcal{D}F \in \mathbb{C}_l^{0,0}(\Lambda_T)$ and $\nabla_\omega F \in \mathbb{C}_l^{0,0}(\Lambda_T)$.

If the process X has continuous paths, Y only depends on the restriction of F to

$$\mathcal{W}_T = \{(t, \omega) \in \Lambda_T, \omega \in C^0([0, T], \mathbb{R}^d)\}$$

so the representation (46) is not unique. However, the definition of $\nabla_\omega F$ (Definition 5.8), which involves evaluating F on paths to which a jump perturbation has been added, seems to depend on the values taken by F *outside* \mathcal{W}_T . It is crucial to resolve this point if one is to deal with functionals of continuous processes or, more generally, processes for which the topological support of the law is not the full space $D([0, T], \mathbb{R}^d)$, otherwise the very definition of the vertical derivative becomes ambiguous.

This question is resolved by the following two theorems (Theorems 5.27 and 5.28), derived in [10].

The first result below shows that if $F \in \mathbb{C}_l^{1,1}(\Lambda_T)$ then $\nabla_\omega F(t, X_t)$ is uniquely determined by the restriction of F to continuous paths:

Theorem 5.27 ([10]). *Consider $F^1, F^2 \in \mathbb{C}_l^{1,1}(\Lambda_T)$ with left-continuous horizontal and vertical derivatives. If F^1, F^2 coincide on continuous paths:*

$$\begin{aligned} & \forall t \in [0, T[, \quad \forall \omega \in C^0([0, T], \mathbb{R}^d), \quad F^1(t, \omega_t) = F^2(t, \omega_t), \\ \text{then} \quad & \forall t \in [0, T[, \quad \forall \omega \in C^0([0, T], \mathbb{R}^d), \quad \nabla_\omega F^1(t, \omega_t) = \nabla_\omega F^2(t, \omega_t). \end{aligned}$$

Proof. Let $F = F^1 - F^2 \in \mathbb{C}_l^{1,1}(\Lambda_T)$ and $\omega \in C^0([0, T], \mathbb{R}^d)$. Then $F(t, \omega) = 0$ for all $t \leq T$. It is then obvious that $\mathcal{D}F(t, \omega)$ is also 0 on continuous paths. Assume now that there exists some $\omega \in C^0([0, T], \mathbb{R}^d)$ such that for some $1 \leq i \leq d$ and $t_0 \in [0, T)$, $\partial_i F(t_0, \omega_{t_0}) > 0$. Let $\alpha = \frac{1}{2} \partial_i F(t_0, \omega_{t_0})$. By the left-continuity of $\partial_i F$ and, using the fact that $\mathcal{D}F \in \mathbb{B}(\Lambda_T)$, there exists $\epsilon > 0$ such that for any $(t', \omega') \in \Lambda_T$,

$$[t' < t_0, \quad d_\infty((t_0, \omega), (t', \omega')) < \epsilon] \Rightarrow (\partial_i F(t', \omega') > \alpha \text{ and } |\mathcal{D}F(t', \omega')| < 1). \quad (47)$$

Choose $t < t_0$ such that $d_\infty(\omega_t, \omega_{t_0}) < \frac{\epsilon}{2}$, define $h := t_0 - t$ and define the following extension of ω_t to $[0, T]$:

$$\begin{aligned} z(u) &= \omega(u), u \leq t \\ z_j(u) &= \omega_j(t) + 1_{i=j}(u - t), t \leq u \leq T, 1 \leq j \leq d \end{aligned} \quad (48)$$

and define the following sequence of piecewise constant approximations of z_{t+h} :

$$z^n(u) = \tilde{z}^n = z(u) \quad \text{for } u \leq t$$

$$z_j^n(u) = \omega_j(t) + 1_{i=j} \frac{h}{n} \sum_{k=0}^n 1_{\frac{kh}{n} \leq u-t}, \quad \text{for } t \leq u \leq t+h, 1 \leq j \leq d \quad (49)$$

Since $\|z_{t+h} - z_{t+h}^n\|_\infty = \frac{h}{n} \rightarrow 0$,

$$|F(t+h, z_{t+h}) - F(t+h, z_{t+h}^n)| \xrightarrow{n \rightarrow +\infty} 0$$

We can now decompose $F(t+h, z_{t+h}^n) - F(t, \omega)$ as

$$F(t+h, z_{t+h}^n) - F(t, \omega) = \sum_{k=1}^n \left(F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}}^n\right) - F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n\right) \right) \\ + \sum_{k=1}^n \left(F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n\right) - F\left(t + \frac{(k-1)h}{n}, z_{t+\frac{(k-1)h}{n}}^n\right) \right) \quad (50)$$

where the first sum corresponds to jumps of z^n at times $t + \frac{kh}{n}$ and the second sum to the ‘horizontal’ variations of z^n on on $[t + \frac{(k-1)h}{n}, t + \frac{kh}{n}]$.

$$F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}}^n\right) - F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n\right) = \phi\left(\frac{h}{n}\right) - \phi(0) \quad (51)$$

$$\text{where } \phi(u) = F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n + ue_i 1_{[t+\frac{kh}{n}, T]}\right)$$

Since F is vertically differentiable, ϕ is differentiable and

$$\phi'(u) = \partial_i F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n + ue_i 1_{[t+\frac{kh}{n}, T]}\right)$$

is continuous. For $u \leq h/n$ we have

$$d_\infty \left((t, \omega_t), \left(t + \frac{kh}{n}, z_{(t+kh/n)-}^n + ue_i 1_{[t+\frac{kh}{n}, T]}\right) \right) \leq h,$$

so $\phi'(u) > \alpha$ hence

$$\sum_{k=1}^n F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}}^n\right) - F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n\right) > \alpha h.$$

On the other hand

$$F\left(t + \frac{kh}{n}, z_{t+\frac{kh}{n}-}^n\right) - F\left(t + \frac{(k-1)h}{n}, z_{t+\frac{(k-1)h}{n}}^n\right) = \psi\left(\frac{h}{n}\right) - \psi(0)$$

where

$$\psi(u) = F\left(t + \frac{(k-1)h}{n}, z_{t+\frac{(k-1)h}{n}}^n + u, z_{t+\frac{(k-1)h}{n}}^n\right)$$

so that ψ is right-differentiable on $]0, \frac{h}{n}[$ with right-derivative:

$$\psi'_r(u) = \mathcal{D}F\left(t + \frac{(k-1)h}{n} + u, z_{t + \frac{(k-1)h}{n}}^n\right)$$

Since $F \in \mathbb{C}_l^{1,1}(\Lambda_T)$, ψ is left-continuous by Proposition 5.5 so

$$\sum_{k=1}^n F\left(t + \frac{kh}{n}, z_{t + \frac{kh}{n}}^n\right) - F\left(t + \frac{(k-1)h}{n}, z_{t + \frac{(k-1)h}{n}}^n\right) = \int_0^h \mathcal{D}F(t + u, z_t^n) du$$

Noting that

$$d_\infty((t + u, z_{t+u}^n), (t + u, z_{t+u})) \leq \frac{h}{n},$$

we obtain

$$\mathcal{D}F(t + u, z_{t+u}^n) \xrightarrow{n \rightarrow +\infty} \mathcal{D}F(t + u, z_{t+u}) = 0$$

since the path of z_{t+u} is continuous. Moreover $|\mathcal{D}F(t + u, z_{t+u}^n)| \leq 1$ since $d_\infty((t + u, z_{t+u}^n), (t_0, \omega)) \leq \epsilon$, so by dominated convergence the integral converges to 0 as $n \rightarrow \infty$. Writing

$$F(t+h, z_{t+h}) - F(t, \omega) = [F(t+h, z_{t+h}) - F(t+h, z_{t+h}^n)] + [F(t+h, z_{t+h}^n) - F(t, \omega)]$$

and taking the limit on $n \rightarrow \infty$ leads to $F(t+h, z_{t+h}) - F(t, \omega) \geq \alpha h$, a contradiction. \square

The above result implies in particular that, if $\nabla_\omega F^i \in \mathbb{C}^{1,1}(\Lambda_T)$, $\mathcal{D}(\nabla_\omega F) \in \mathbb{B}(\Lambda_T)$ and $F^1(\omega) = F^2(\omega)$ for any continuous path ω , then $\nabla_\omega^2 F^1$ and $\nabla_\omega^2 F^2$ must also coincide on continuous paths. The next theorem shows that this result can be obtained under the weaker assumption that $F^i \in \mathbb{C}^{1,2}(\Lambda_T)$, using a probabilistic argument. Interestingly, while the uniqueness of the first vertical derivative (Theorem 5.27) is based on the fundamental theorem of calculus, the proof of the following theorem is based on its stochastic equivalent, the Itô formula [40, 41].

Theorem 5.28. *If $F^1, F^2 \in \mathbb{C}_b^{1,2}(\Lambda_T)$ coincide on continuous paths:*

$$\forall \omega \in C^0([0, T], \mathbb{R}^d), \quad \forall t \in [0, T], \quad F^1(t, \omega_t) = F^2(t, \omega_t), \quad (52)$$

then their second vertical derivatives also coincide on continuous paths:

$$\forall \omega \in C^0([0, T], \mathbb{R}^d), \quad \forall t \in [0, T], \quad \nabla_\omega^2 F^1(t, \omega_t) = \nabla_\omega^2 F^2(t, \omega_t).$$

Proof. Let $F = F^1 - F^2$. Assume that there exists $\omega \in C^0([0, T], \mathbb{R}^d)$ such that for some $1 \leq i \leq d$ and $t_0 \in [0, T)$ and some direction $h \in \mathbb{R}^d, \|h\| = 1$, ${}^t h \nabla_\omega^2 F(t_0, \omega_{t_0}) \cdot h > 0$, and denote $\alpha = \frac{1}{2} {}^t h \nabla_\omega^2 F(t_0, \omega_{t_0}) \cdot h$. We will show that

this leads to a contradiction. We already know that $\nabla_\omega F(t, \omega_t) = 0$ by Theorem 5.27. There exists $\eta > 0$ such that

$$\forall (t', \omega') \in \Lambda_T, \{t' \leq t_0, d_\infty((t_0, \omega), (t', \omega')) < \eta\} \Rightarrow \max(|F(t', \omega') - F(t_0, \omega_{t_0})|, |\nabla_\omega F(t', \omega')|, |\mathcal{D}F(t', \omega')|) < 1, \quad {}^t h \nabla_\omega^2 F(t', \omega') \cdot h > \alpha. \quad (53)$$

Choose $t < t_0$ such that $d_\infty(\omega_t, \omega_{t_0}) < \frac{\eta}{2}$ and denote $\epsilon = \frac{\eta}{2} \wedge (t_0 - t)$. Let W be a real Brownian motion on an (auxiliary) probability space $(\tilde{\Omega}, \mathcal{B}, \mathbb{P})$ whose generic element we will denote w , $(\mathcal{B}_s)_{s \geq 0}$ its natural filtration, and let

$$\tau = \inf\{s > 0, |W(s)| = \frac{\epsilon}{2}\}. \quad (54)$$

Define, for $t' \in [0, T]$, the 'Brownian extrapolation'

$$U_{t'}(\omega) = \bar{\omega}(t')1_{t' \leq t} + (\omega(t) + W((t' - t) \wedge \tau)h)1_{t' > t}. \quad (55)$$

For all $s < \frac{\epsilon}{2}$, we have

$$d_\infty((t + s, U_{t+s}(\omega)), (t, \omega_t)) < \epsilon \quad \mathbb{P} - a.s. \quad (56)$$

Define the following piecewise constant approximation of the stopped process W_τ :

$$W^n(s) = \sum_{i=0}^{n-1} W(i \frac{\epsilon}{2n} \wedge \tau) 1_{s \in [i \frac{\epsilon}{2n}, (i+1) \frac{\epsilon}{2n})} + W(\frac{\epsilon}{2} \wedge \tau) 1_{s = \frac{\epsilon}{2}}, \quad 0 \leq s \leq \frac{\epsilon}{2} \quad (57)$$

Denoting

$$Z(s) = F(t + s, U_{t+s}), \quad s \in [0, T - t], \quad Z^n(s) = F(t + s, U_{t+s}^n), \quad (58)$$

$$U_{t'}^n(\omega) = \bar{\omega}(t')1_{t' \leq t} + (\omega(t) + W^n((t' - t) \wedge \tau)h)1_{t' > t}, \quad (59)$$

we have the following decomposition:

$$\begin{aligned} Z(\frac{\epsilon}{2}) - Z(0) &= Z(\frac{\epsilon}{2}) - Z^n(\frac{\epsilon}{2}) + \sum_{i=1}^n \left(Z^n(i \frac{\epsilon}{2n}) - Z^n(i \frac{\epsilon}{2n} -) \right) \\ &\quad + \sum_{i=0}^{n-1} \left(Z^n((i+1) \frac{\epsilon}{2n} -) - Z^n(i \frac{\epsilon}{2n}) \right) \end{aligned} \quad (60)$$

The first term in (60) vanishes almost surely since

$$\|U_{t+\frac{\epsilon}{2}} - U_{t+\frac{\epsilon}{2}}^n\|_\infty \xrightarrow{n \rightarrow \infty} 0.$$

The second term in (60) may be expressed as

$$Z^n(i \frac{\epsilon}{2n}) - Z^n(i \frac{\epsilon}{2n} -) = \phi_i(W(i \frac{\epsilon}{2n}) - W((i-1) \frac{\epsilon}{2n})) - \phi_i(0) \quad (61)$$

where

$$\phi_i(u, \omega) = F\left(t + i\frac{\epsilon}{2n}, U_{t+i\frac{\epsilon}{2n}-}^n(\omega) + uh1_{[t+i\frac{\epsilon}{2n}, T]}\right).$$

Note that $\phi_i(u, \omega)$ is measurable with respect to $\mathcal{B}_{(i-1)\epsilon/2n}$ whereas (61) is independent with respect to $\mathcal{B}_{(i-1)\epsilon/2n}$. Let $\Omega_1 \subset \tilde{\Omega}$, $\mathbb{P}(\Omega_1) = 1$ such that W has continuous sample paths on Ω_1 . Then, on Ω_1 , $\phi_i(\cdot, \omega) \in C^2(\mathbb{R})$ and the following relations hold \mathbb{P} -almost surely:

$$\begin{aligned}\phi'_i(u, \omega) &= \nabla_\omega F\left(t + i\frac{\epsilon}{2n}, U_{t+i\frac{\epsilon}{2n}-}^n(\omega_t) + uh1_{[t+i\frac{\epsilon}{2n}, T]}\right) \cdot h \\ \phi''_i(u, \omega) &= {}^t h \nabla_\omega^2 F\left(t + i\frac{\epsilon}{2n}, U_{t+i\frac{\epsilon}{2n}-}^n(\omega_t) + uh1_{[t+i\frac{\epsilon}{2n}, T]}\right) \cdot h\end{aligned}\quad (62)$$

So, using the above arguments we can apply the Itô formula to (61) on Ω_1 . We therefore obtain, summing on i and denoting $i(s)$ the index such that $s \in [(i(s) - 1)\frac{\epsilon}{2n}, i(s)\frac{\epsilon}{2n}]$:

$$\begin{aligned}& \sum_{i=1}^n Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-) = \\ & \int_0^{\frac{\epsilon}{2}} \nabla_\omega F\left(t + i(s)\frac{\epsilon}{2n}, U_{t+i(s)\frac{\epsilon}{2n}-}^n + (W(s) - W((i(s) - 1)\frac{\epsilon}{2n}))h1_{[t+i(s)\frac{\epsilon}{2n}, T]}\right) \cdot dW(s) \\ & + \frac{1}{2} \int_0^{\frac{\epsilon}{2}} \langle h, \nabla_\omega^2 F\left(t + i(s)\frac{\epsilon}{2n}, U_{t+i(s)\frac{\epsilon}{2n}-}^n + (W(s) - W((i(s) - 1)\frac{\epsilon}{2n}))h1_{[t+i(s)\frac{\epsilon}{2n}, T]}\right) \cdot h \rangle ds\end{aligned}$$

Since the first derivative is bounded by (53), the stochastic integral is a martingale, so taking expectation leads to

$$E\left[\sum_{i=1}^n Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-)\right] \geq \alpha \frac{\epsilon}{2}. \quad (63)$$

$$Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) = \psi\left(\frac{\epsilon}{2n}\right) - \psi(0) \quad (64)$$

where

$$\psi(u) = F\left(t + i\frac{\epsilon}{2n} + u, U_{t+i\frac{\epsilon}{2n}}^n\right) \quad (65)$$

is right-differentiable with right derivative

$$\psi'(u) = \mathcal{D}F\left(t + i\frac{\epsilon}{2n} + u, U_{t+i\frac{\epsilon}{2n}}^n\right). \quad (66)$$

Since $F \in \mathcal{C}_l^{0,0}([0, T])$, ψ is left-continuous and the fundamental theorem of calculus yields

$$\sum_{i=0}^{n-1} Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) = \int_0^{\frac{\epsilon}{2}} \mathcal{D}F\left(t + s, U_{t+(i(s)-1)\frac{\epsilon}{2n}}^n\right) ds. \quad (67)$$

The integrand converges to $\mathcal{D}F(t+s, U_{t+s}) = 0$ as $n \rightarrow \infty$ since $\mathcal{D}F(t+s, \omega) = 0$ whenever ω is continuous. Since this term is also bounded, by dominated convergence

$$\int_0^{\frac{\epsilon}{2}} \mathcal{D}F(t+s, U_{t+(i(s)-1)\frac{\epsilon}{2n}}) ds \xrightarrow{n \rightarrow \infty} 0.$$

It is obvious that $Z(\frac{\epsilon}{2}) = 0$ since $F(t, \omega) = 0$ whenever ω is a continuous path. On the other hand, since all derivatives of F appearing in (60) are bounded, the dominated convergence theorem allows to take expectations of both sides in (60) with respect to the Wiener measure and obtain $\alpha_{\frac{\epsilon}{2}} = 0$, a contradiction. \square

Theorem 5.28 is a key result: it enables us to define the class $\mathbb{C}_b^{1,2}(\mathcal{W}_T)$ of non-anticipative functionals such that their restriction to \mathcal{W}_T fulfills the conditions of Theorem 5.28:

$$F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T) \iff \exists \tilde{F} \in \mathbb{C}_b^{1,2}(\Lambda_T), \quad \tilde{F}|_{\mathcal{W}_T} = F, \quad (68)$$

without having to extend the functional to the full space Λ_T .

For such functionals, coupling the proof of Theorem 5.22 with Theorem 5.28 yields the following result:

Theorem 5.29 (Pathwise change of variable formula for $\mathbb{C}_b^{1,2}(\mathcal{W}_T)$ functionals). *For any $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$, $\omega \in C^0([0, T], \mathbb{R}^d) \cap Q^\pi([0, T], \mathbb{R}^d)$ the limit*

$$\int_0^T \nabla_\omega F(t, \omega_t) d^\pi \omega := \lim_{n \rightarrow \infty} \sum_{i=0}^{k(n)-1} \nabla_\omega F(t_i^n, \omega_{t_i^n}^-) \cdot (\omega(t_{i+1}^n) - \omega(t_i^n)) \quad (69)$$

exists and

$$F(T, \omega_T) - F(0, \omega_0) = \int_0^T \mathcal{D}F(t, \omega_t) dt + \int_0^T \nabla_\omega F(t, \omega_t) d^\pi \omega + \int_0^T \frac{1}{2} \text{tr}({}^t \nabla_\omega^2 F(t, \omega_t) d[\omega]).$$

5.5 Application to functionals of stochastic processes

Consider now a stochastic process $Z : [0, T] \times \Omega \mapsto \mathbb{R}^d$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$. The previous formula holds for functionals of Z along a sequence of partitions π on the set

$$\Omega_\pi(Z) = \{\omega \in \Omega, \quad Z(\cdot, \omega) \in Q^\pi([0, T], \mathbb{R}^d)\}.$$

If we can construct a sequence of partitions π such that $\mathbb{P}(\Omega_\pi(Z)) = 1$ then the functional Itô formula will hold almost surely. Fortunately this turns out to be the case for many important classes of stochastic processes:

- Wiener process: if W is a Wiener process under \mathbb{P} , then for any *nested* sequence of partitions π with $|\pi_n| \rightarrow 0$, the paths of W lie in $Q^\pi([0, T], \mathbb{R}^d)$ with probability 1 [47]:

$$\mathbb{P}(\Omega_\pi(W)) = 1 \quad \text{and} \quad \forall \omega \in \Omega_\pi(W), [W(\cdot, \omega)]_\pi(t) = t.$$

This is a classical result due to Paul Lévy [47, Sec. 4, Theorem 5]. The nesting condition may be removed if one requires that $|\pi_n| \log n \rightarrow 0$ [20].

- Fractional Brownian motion: if B^H is fractional Brownian motion with Hurst index $H \in (0.5, 1)$ then for any sequence of partitions π with $n|\pi_n| \rightarrow 0$, the paths of B^H lie in $Q^\pi([0, T], \mathbb{R}^d)$ with probability 1 [20]:

$$\mathbb{P}(\Omega_\pi(B^H)) = 1 \quad \text{and} \quad \forall \omega \in \Omega_\pi(B^H), [B^H(\cdot, \omega)](t) = 0.$$

- Brownian stochastic integrals: Let $\sigma : (\Lambda_T, d_\infty) \rightarrow \mathbb{R}$ be a Lipschitz map, B a Wiener process and consider the Itô stochastic integral

$$X = \int_0^\cdot \sigma(t, B_t) dB(t).$$

Then for any sequence of partitions π with $n|\pi_n| \log n \rightarrow 0$, the paths of X lie in $Q^\pi([0, T], \mathbb{R})$ with probability 1 and

$$[X](t, \omega) = \int_0^t |\sigma(u, B_u(\omega))|^2 du$$

- Lévy processes: if L is a Lévy process with triplet (b, A, ν) then for *any* sequence of partitions π with $n|\pi_n| \rightarrow 0$, $\mathbb{P}(\Omega_\pi(L)) = 1$ and

$$\forall \omega \in \Omega_\pi(L), [L(\cdot, \omega)](t) = tA + \sum_{s \in [0, t]} |L(s, \omega) - L(s-, \omega)|^2.$$

Note that the only property needed for the change of variable formula to hold (and, thus, for the integral to exist pathwise) is the finite quadratic variation property, which does not require the process to be a semimartingale.

The construction of the Föllmer integral depends a priori on the sequence π of partitions. But in the case of a semimartingales one can identify these limits of (non-anticipative) Riemann sums as Itô integrals, which guarantees that the limit is a.s. unique, independent of the choice of π . We now take a closer look at the semimartingale case.

6 The Functional Ito formula

6.1 Semimartingales and quadratic variation

We now consider a **semimartingale** X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with the natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ of X ; X is a cadlag process and the *stochastic integral*

$$\phi \in \mathbb{S}(\mathbb{F}) \mapsto \int \phi dX$$

defines a functional on the set $\mathbb{S}(\mathbb{F})$ of simple \mathbb{F} -predictable processes, with the following continuity property: for any sequence $\phi^n \in \mathbb{S}(\mathbb{F})$ of simple predictable processes,

$$\sup_{[0, T] \times \Omega} |\phi^n(t, \omega) - \phi(t, \omega)| \xrightarrow{n \rightarrow \infty} 0 \quad \Rightarrow \quad \int_0^\cdot \phi^n . dX \xrightarrow[n \rightarrow \infty]{\text{UCP}} \int_0^\cdot \phi . dX, \quad (70)$$

where UCP stands for uniform convergence in probability on compact sets [62]. For any caglad adapted process ϕ , the Ito integral $\int_0^\cdot \phi dX$ may be then constructed as a limit (in probability) of nonanticipative Riemann sums: for any sequence $(\pi_n)_{n \geq 1}$ of partitions of $[0, T]$ with $|\pi_n| \rightarrow 0$,

$$\sum_{\pi_n} \phi(t_k^n) \cdot (X(t_{k+1}^n) - X(t_k^n)) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^T \phi dX.$$

Let us recall some important properties of semimartingales [19, 62]:

- Quadratic variation: for any sequence of partitions $\pi = (\pi_n)_{n \geq 1}$ of $[0, T]$ with $|\pi_n| \rightarrow 0$ a.s,

$$\sum_{\pi_n} (X(t_{k+1}^n) - X(t_k^n))^2 \xrightarrow{\mathbb{P}} [X](T) = [X]^c(T) + \sum_{0 \leq s \leq T} \Delta X(s)^2 < \infty$$

- Ito formula: $\forall f \in C^2(\mathbb{R}^d, \mathbb{R})$,

$$\begin{aligned} f(X(t)) &= f(X(0)) + \int_0^t \nabla f(X) dX + \int_0^t \frac{1}{2} \text{tr}(\partial_{xx}^2 f(X)) . d[X]^c \\ &+ \sum_{0 \leq s \leq t} (f(X(s-) + \Delta X(s)) - f(X(s-)) - \nabla f(X(s-)) . \Delta X(s)) \end{aligned}$$

- Semimartingale Decomposition: X has a unique decomposition

$$X = M^d + M^c + A$$

where M^c is a continuous \mathbb{F} -local martingale, M^d pure-jump \mathbb{F} -local martingale and A a continuous \mathbb{F} -adapted finite variation process.

- The increasing process $[X]$ has a unique decomposition $[X] = [M]^d + [M]^c$ where $[M]^d(t) = \sum_{0 \leq s \leq t} \Delta X(s)^2$, $[M]^c$ a continuous increasing \mathbb{F} -adapted process.
- If X has continuous paths then it has a unique decomposition $X = M + A$ where M continuous \mathbb{F} -local martingale, A a continuous \mathbb{F} -adapted process with finite variation.

These properties have several consequences. First, if $F \in \mathbb{C}_l^{0,0}(\Lambda_T)$ then the non-anticipative Riemann sums

$$\sum_{t_k^n \in \pi_n} F(t_k^n, X_{t_k^n-}) \cdot (X(t_{k+1}^n) - X(t_k^n)) \xrightarrow{\mathbb{P}} \int_0^T F(t, X_t) dX$$

converge in probability to the Ito stochastic integral $\int F(t, X) dX$. So, by a.s. uniqueness of the limit, the Föllmer integral constructed along any sequence of partitions $\pi = (\pi_n)_{n \geq 1}$ with $|\pi_n| \rightarrow 0$ almost-surely coincides with the Ito integral. In particular, the limit is independent of the choice of partitions, so we omit in the sequel the dependence of the integrals on the sequence π .

The semimartingale property also enables to construct a partition with respect to which the paths of the semimartingale have the finite quadratic variation property with probability 1:

Proposition 6.1. *Let S be a semimartingale on $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0})$, $T > 0$. There exists a sequence of partitions $\pi = (\pi_n)_{n \geq 1}$ of $[0, T]$ with $|\pi_n| \rightarrow 0$, such that the paths of S lie in $Q^\pi([0, T], \mathbb{R}^d)$ with probability 1:*

$$\mathbb{P}(\{\omega \in \Omega, \quad S(\cdot, \omega) \in Q^\pi([0, T], \mathbb{R}^d)\}) = 1.$$

Proof. Consider the dyadic partition $t_k^n = kT/2^n, k = 0..2^n$. Since $\sum_{\pi_n} (S(t_{k+1}^n) - S(t_k^n))^2 \rightarrow [S]_T$ in probability, there exists a subsequence $(\pi_n)_{n \geq 1}$ of partitions such that

$$\sum_{\pi_n} (S(t_i^n) - S(t_{i+1}^n))^2 \xrightarrow{n \rightarrow \infty} [S]_T \quad \mathbb{P} - a.s.$$

This subsequence achieves the result. \square

The notion of semimartingale in the literature usually refers to a real-valued process but one can extend this to vector-valued semimartingales [43]. For an \mathbb{R}^d -valued semimartingale $X = (X^1, \dots, X^d)$, the above properties should be understood in the vector sense, and the quadratic (co-)variation process is an S_d^+ -valued process, defined by

$$\sum_{t_i^n \in \pi_n, t_i^n \leq t} (X(t_i^n) - X(t_{i+1}^n)) \cdot^t (X(t_i^n) - X(t_{i+1}^n)) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} [X](t).$$

$t \rightarrow [X](t)$ is a.s. increasing in the sense of the order on positive symmetric matrices:

$$\forall t \geq 0, \quad \forall h > 0, \quad [X](t+h) - [X](t) \in S_d^+.$$

6.2 The Functional Ito formula

Using Proposition 6.1, we can now apply the pathwise change of variable formula derived in Theorem 5.22 to any semimartingale. The following *functional Ito formula*, shown in [8, Proposition 6], is a consequence of Theorem 5.22 combined with Proposition 6.1:

Theorem 6.2 (Functional Ito formula: cadlag case). *Let X be an \mathbb{R}^d -valued semimartingale and denote for $t > 0$, $X_{t-}(u) = X(u) \mathbb{1}_{[0,t[}(u) + X(t-) \mathbb{1}_{[t,T]}(u)$. For any $F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$, $t \in [0, T[$,*

$$\begin{aligned} F(t, X_t) - F(0, X_0) &= \int_0^t \mathcal{D}F(u, X_u) du + \\ \int_0^t \nabla_\omega F(u, X_u) \cdot dX(u) &+ \int_0^t \frac{1}{2} \text{tr} (\nabla_\omega^2 F(u, X_u) d[X](u)) \\ &+ \sum_{u \in]0, t]} [F(u, X_u) - F(u, X_{u-}) - \nabla_\omega F(u, X_{u-}) \cdot \Delta X(u)] \quad a.s. \end{aligned} \quad (71)$$

In particular, $Y(t) = F(t, X_t)$ is a semimartingale: the class of semimartingales is stable under transformations by $\mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$ functionals.

More precisely, we can choose $\pi = (\pi_n)_{n \geq 1}$ with

$$\sum_{\pi_n} (X(t_i^n) - X(t_{i+1}^n))^t (X(t_i^n) - X(t_{i+1}^n)) \xrightarrow{n \rightarrow \infty} [X](t) \quad a.s.$$

so setting

$$\Omega_X = \left\{ \omega \in \Omega, \sum_{\pi_n} (X(t_i) - X(t_{i+1}))^t (X(t_i) - X(t_{i+1})) \xrightarrow{n \rightarrow \infty} [X](T) < \infty \right\}$$

we have that $\mathbb{P}(\Omega_X) = 1$ and for any $F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$ and any $\omega \in \Omega_X$, the limit

$$\int_0^t \nabla_\omega F(u, X_u(\omega)) \cdot d^\pi X(\omega) := \lim_{n \rightarrow \infty} \sum_{t_i^n \in \pi_n} \nabla_\omega F(t_i^n, X_{t_i^n}^n(\omega)) \cdot (X(t_{i+1}^n, \omega) - X(t_i^n, \omega))$$

exists and for all $t \in [0, T]$,

$$\begin{aligned} F(t, X_t(\omega)) - F(0, X_0(\omega)) &= \int_0^t \mathcal{D}F(u, X_u(\omega)) du + \\ \int_0^t \nabla_\omega F(u, X_u(\omega)) \cdot dX(\omega) &+ \int_0^t \frac{1}{2} \text{tr} (\nabla_\omega^2 F(u, X_u(\omega)) d[X](\omega)) \\ &+ \sum_{u \in]0, t]} [F(u, X_u(\omega)) - F(u, X_{u-}(\omega)) - \nabla_\omega F(u, X_{u-}(\omega)) \cdot \Delta X(\omega)(u)]. \end{aligned}$$

Remark 6.3. Note that, unlike the usual statement of the Ito formula (see e.g. [62, Ch. II, Section 7]), the statement here is that there exists a set Ω_X on which the equality (71) holds pathwise for *any* $F \in \mathbb{C}_b^{1,2}([0, T])$. This is particularly useful when one needs to take a supremum over such F , such as in optimal control problems, since the null set does not depend on F .

In the continuous case, the functional Ito formula reduces to the following [21, 9, 8], which we give for the sake of completeness:

Theorem 6.4 (Functional Ito formula: continuous case[9, 21, 8]). *Let X be a continuous semimartingale and $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$. For any $t \in [0, T]$,*

$$\begin{aligned} F(t, X_t) - F(0, X_0) &= \int_0^t \mathcal{D}F(u, X_u) du + \\ \int_0^t \nabla_\omega F(u, X_u) \cdot dX(u) &+ \int_0^t \frac{1}{2} \text{tr} ({}^t \nabla_\omega^2 F(u, X_u) d[X]) \quad a.s. \end{aligned} \quad (72)$$

In particular, $Y(t) = F(t, X_t)$ is a continuous semimartingale.

If $F(t, X_t) = f(t, X(t))$ where $f \in C^{1,2}([0, T] \times \mathbb{R}^d)$ this reduces to the standard Ito formula.

Note that, using Theorem 5.28, it is sufficient to require that $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$ rather than $F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$.

Theorem 6.4 shows that, for a continuous semimartingale X , any smooth non-anticipative functional $Y = F(X)$ depends on F and its derivatives only via their values on continuous paths i.e. on $\mathcal{W}_T \subset \Lambda_T$. Thus, $Y = F(X)$ can be reconstructed from the "second-order jet" $(\nabla_\omega F, \mathcal{D}F, \nabla_\omega^2 F)$ of the functional F on \mathcal{W}_T .

Although these formulas are implied by the stronger pathwise formula (Theorem 5.22), one can also give a direct probabilistic proof using the Ito formula [11]. We outline the main ideas of the probabilistic proof, which show the role played by the different assumptions. A more detailed version may be found in [11].

Sketch of proof of Theorem 6.4: Consider first a cadlag piecewise constant process:

$$X(t) = \sum_{k=1}^n 1_{[t_k, t_{k+1}[}(t) \Phi_k$$

where Φ_k are \mathcal{F}_{t_k} -measurable bounded random variables. Each path of X is a sequence of horizontal and vertical moves:

$$X_{t_{k+1}} = X_{t_k} + (\Phi_{k+1} - \Phi_k) 1_{[t_{k+1}, T]}$$

We now decompose each increment of F into a horizontal and vertical part:

$$\begin{aligned} F(t_{k+1}, X_{t_{k+1}}) - F(t_k, X_{t_k}) &= \\ F(t_{k+1}, X_{t_{k+1}}) - F(t_{k+1}, X_{t_k}) &+ \quad \text{vertical increment} \\ F(t_{k+1}, X_{t_k}) - F(t_k, X_{t_k}) &+ \quad \text{horizontal increment} \end{aligned} \quad (73)$$

The horizontal increment is the increment of $\phi(h) = F(t_k + h, X_{t_k})$. The fundamental theorem of calculus applied to ϕ yields:

$$F(t_{k+1}, X_{t_k}) - F(t_k, X_{t_k}) = \phi(t_{k+1} - t_k) - \phi(0) = \int_{t_k}^{t_{k+1}} \mathcal{D}F(t, X_t) dt \quad (74)$$

To compute the vertical increment, we apply the Ito formula to $\psi \in C^2(\mathbb{R}^d)$ defined by $\psi(u) = F(t_{k+1}, X_{t_k} + u1_{[t_{k+1}, T]})$. This yields

$$\begin{aligned} F(t_{k+1}, X_{t_{k+1}}) - F(t_{k+1}, X_{t_k}) &= \psi(X(t_{k+1}) - X(t_k)) - \psi(0) \\ &= \int_{t_k}^{t_{k+1}} \nabla_{\omega} F(t, X_t) \cdot dX(t) + \frac{1}{2} \int_{t_k}^{t_{k+1}} \text{tr}(\nabla_{\omega}^2 F(t, X_t) d[X]) \end{aligned} \quad (75)$$

Consider now the case of a continuous semimartingale X ; the piecewise-constant approximation ${}_n X$ of X along the partition π_n approximates X almost surely in supremum norm. Thus, by the previous argument we have

$$F(T, {}_n X_T) - F(0, X_0) = \int_0^T \mathcal{D}F(t, {}_n X_t) dt + \int_0^T \nabla_{\omega} F(t, {}_n X_t) \cdot d_n X + \frac{1}{2} \int_0^T \text{tr}({}^t \nabla_{\omega}^2 F({}_n X_t) \cdot d[{}_n X])$$

Since $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$ all derivatives involved in the expression are left-continuous in d_{∞} metric, which allows to control their convergence as $n \rightarrow \infty$. Using local boundedness assumption $\nabla_{\omega} F, \mathcal{D}F, \nabla_{\omega}^2 F \in \mathbb{B}(\Lambda_T)$ we can then use the dominated convergence theorem, its extension to stochastic integrals [62, Ch.IV Theorem 32] and Lemma 5.23 to conclude that the Riemann-Stieltjes integrals converge almost surely, and the stochastic integral in probability, to the terms appearing in (72) as $n \rightarrow \infty$ [11].

6.3 Functionals with dependence on quadratic variation

The results outlined in the previous sections all assume that the functionals involved, and their directional derivatives, are continuous in supremum norm. This is a severe restriction for applications in stochastic analysis, where functionals may involve quantities such as quadratic variation, which cannot be constructed as a continuous functional for the supremum norm.

More generally, in many applications such as statistics of processes, physics or mathematical finance, one is led to consider path-dependent functionals of a semimartingale X and its quadratic variation process $[X]$ such as

$$\int_0^t g(t, X_t) d[X](t), \quad G(t, X_t, [X]_t), \quad \text{or} \quad E[G(T, X(T), [X](T)) | \mathcal{F}_t]$$

where $X(t)$ denotes the value at time t and $X_t = (X(u), u \in [0, t])$ the path up to time t .

In Section 7 we will develop a weak functional calculus capable of handling all such examples in a weak sense. However, disposing of a pathwise interpretation is of essence in most applications and this interpretation is lost when passing to weak derivatives. In this section, we show how the pathwise calculus outlined in Sections 5 and 6 can be extended to functionals depending on quadratic variation while retaining the pathwise interpretation.

Consider a continuous \mathbb{R}^d -valued semimartingale with absolutely continuous quadratic variation:

$$[X](t) = \int_0^t A(u) du$$

where A is an S_d^+ -valued process. Denote by \mathcal{F}_t the natural filtration of X .

The idea is to 'double the variables' and represent the above examples in the form

$$Y(t) = F(t, \{X(u), 0 \leq u \leq t\}, \{A(u), 0 \leq u \leq t\}) = F(t, X_t, A_t) \quad (76)$$

where F is a non-anticipative functional defined on an enlarged domain $D([0, T], \mathbb{R}^d) \times D([0, T], S_d^+)$:

$$F : [0, T] \times D([0, T], \mathbb{R}^d) \times D([0, T], S_d^+) \rightarrow \mathbb{R}$$

and represents the dependence of Y on the stopped path $X_t = \{X(u), 0 \leq u \leq t\}$ of X and its quadratic variation.

Introducing the process A as additional variable may seem redundant: indeed $A(t)$ is itself \mathcal{F}_t -measurable i.e. a functional of X_t . However, it is not a *continuous* functional on (Λ_T, d_∞) . Introducing A_t as a second argument in the functional will allow us to control the regularity of Y with respect to $[X](t) = \int_0^t A(u)du$ simply by requiring continuity of F with respect to the "lifted process" (X, A) . This idea is analogous to the approach of rough path theory [49, 34, 35] in which one controls integral functionals of a path in terms of the path jointly with its 'Lévy area'. Here $d[X]$ is the symmetric part of the Lévy area, while the asymmetric part does not intervene in the change of variable formula. However, unlike the rough path construction, in our construction we do not need to resort to p-variation norms.

An important property in the above examples is that their dependence on the process A is either through $[X] = \int_0^\cdot A(u)du$ or through an integral functional of the form $\int_0^\cdot \phi d[X] = \int_0^\cdot \phi(u)A(u)du$. In both cases they satisfy the condition

$$F(t, X_t, A_t) = F(t, X_t, A_{t-})$$

where, as before, we denote

$$\omega_{t-} = \omega \cdot 1_{[0, t[} + \omega(t-) \cdot 1_{[t, T]}.$$

Following these ideas, we define the space $(\mathcal{S}_T, d_\infty)$ of stopped paths

$$\mathcal{S}_T = \{(t, x(t \wedge \cdot), v(t \wedge \cdot)), (t, x, v) \in [0, T] \times D([0, T], \mathbb{R}^d) \times D([0, T], S_d^+)\} \quad (77)$$

and we will assume throughout this section that

Assumption 6.1. $F : (\mathcal{S}_T, d_\infty) \rightarrow \mathbb{R}$ is a non-anticipative functional with "predictable" dependence with respect to the second argument:

$$\forall (t, x, v) \in \mathcal{S}_T, \quad F(t, x, v) = F(t, x_t, v_{t-}). \quad (78)$$

The regularity concepts introduced in Section 5 carry out without any modification to the case of functionals on the larger space \mathcal{S}_T ; we define the corresponding class of regular functionals $\mathbb{C}_{\text{loc}}^{1,2}(\mathcal{S}_T)$ by analogy with Definition 5.16).

Condition (78) entails that vertical derivatives with respect to the variable v are zero, so the vertical derivative on the product space coincides with the vertical derivative with respect to the variable x ; we continue to denote it as ∇_ω .

As the examples below show, the decoupling of the variables (X, A) leads to a much larger class of smooth functionals:

Example 6.5 (Integrals with respect to quadratic variation). A process $Y(t) = \int_0^t g(X(u))d[X](u)$ where $g \in C^0(\mathbb{R}^d)$ may be represented by the functional

$$F(t, x_t, v_t) = \int_0^t g(x(u))v(u)du. \quad (79)$$

F verifies Assumption 6.1, $F \in \mathbb{C}_b^{1,\infty}$, with:

$$\mathcal{D}F(t, x_t, v_t) = g(x(t))v(t) \quad \nabla_\omega^j F(t, x_t, v_t) = 0$$

Example 6.6. The process $Y(t) = X(t)^2 - [X](t)$ is represented by the functional

$$F(t, x_t, v_t) = x(t)^2 - \int_0^t v(u)du. \quad (80)$$

F verifies Assumption 6.1 and $F \in \mathbb{C}_b^{1,\infty}(\mathcal{S}_T)$ with:

$$\begin{aligned} \mathcal{D}F(t, x, v) &= -v(t) & \nabla_\omega F(t, x_t, v_t) &= 2x(t) \\ \nabla_\omega^2 F(t, x_t, v_t) &= 2 & \nabla_\omega^j F(t, x_t, v_t) &= 0, j \geq 3 \end{aligned}$$

Example 6.7. The process $Y = \exp(X - [X]/2)$ may be represented as $Y(t) = F(t, X_t, A_t)$

$$F(t, x_t, v_t) = e^{x(t) - \frac{1}{2} \int_0^t v(u)du}$$

Elementary computations show that $F \in \mathbb{C}_b^{1,\infty}(\mathcal{S}_T)$ with:

$$\mathcal{D}F(t, x, v) = -\frac{1}{2}v(t)F(t, x, v) \quad \nabla_\omega^j F(t, x_t, v_t) = F(t, x_t, v_t)$$

We can now state a change of variable formula for non-anticipative functionals which are allowed to depend on the path of X and its quadratic variation:

Theorem 6.8. *Let $F \in \mathbb{C}_b^{1,2}(\mathcal{S}_T)$ be a non-anticipative functional verifying (78). Then for $t \in [0, T[$,*

$$\begin{aligned} F(t, X_t, A_t) - F_0(X_0, A_0) &= \int_0^t \mathcal{D}_u F(X_u, A_u)du + \int_0^t \nabla_\omega F(u, X_u, A_u) \cdot dX(u) \\ &+ \int_0^t \frac{1}{2} \text{tr} (\nabla_\omega^2 F(u, X_u, A_u) d[X]) \quad a.s. \end{aligned} \quad (81)$$

In particular, for any $F \in \mathbb{C}_b^{1,2}(\mathcal{S}_T)$, $Y(t) = F(t, X_t, A_t)$ is a continuous semimartingale.

Importantly, (81) contains no extra term involving the dependence on A as compared to (72). As will be clear in the proof, due to the predictable dependence of F in A , one can in fact 'freeze' the variable A when computing the increment/ differential of F , so locally A acts a parameter.

Proof. Let us first assume that X does not exit a compact set K and that $\|A\|_\infty \leq R$ for some $R > 0$. Let us introduce a sequence of random partitions $(\tau_k^n, k = 0..k(n))$ of $[0, t]$, by adding the jump times of A to the dyadic partition $(t_i^n = \frac{it}{2^n}, i = 0..2^n)$:

$$\tau_0^n = 0 \quad \tau_k^n = \inf\{s > \tau_{k-1}^n \mid \frac{2^n s}{t} \in \mathbb{N} \text{ or } |A(s) - A(s-)| > \frac{1}{n}\} \wedge t \quad (82)$$

The construction of $(\tau_k^n, k = 0..k(n))$ ensures that

$$\eta_n = \sup\{|A(u) - A(\tau_i^n)| + |X(u) - X(\tau_{i+1}^n)| + \frac{t}{2^n}, i \leq 2^n, u \in [\tau_i^n, \tau_{i+1}^n[]\}_{n \rightarrow \infty} \rightarrow 0.$$

Denote ${}_n X = \sum_{i=0}^{\infty} X(\tau_{i+1}^n)1_{[\tau_i^n, \tau_{i+1}^n)} + X(t)1_{\{t\}}$ which is a cadlag piecewise constant approximation of X_t , and ${}_n A = \sum_{i=0}^{\infty} A(\tau_i^n)1_{[\tau_i^n, \tau_{i+1}^n)} + A(t)1_{\{t\}}$ which is an adapted cadlag piecewise-constant approximation of A_t . Denote $h_i^n = \tau_{i+1}^n - \tau_i^n$. Start with the decomposition:

$$\begin{aligned} F(\tau_{i+1}^n, {}_n X_{\tau_{i+1}^n-}, {}_n A_{\tau_{i+1}^n-}) - F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) &= \\ F(\tau_{i+1}^n, {}_n X_{\tau_{i+1}^n-}, {}_n A_{\tau_i^n}) - F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n}) &+ \\ F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) - F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) & \end{aligned} \quad (83)$$

where we have used (78) to deduce $F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) = F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-})$. The first term in (83) can be written $\psi(h_i^n) - \psi(0)$ where:

$$\psi(u) = F(\tau_i^n + u, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) \quad (84)$$

Since $F \in \mathbb{C}_b^{1,2}(\mathcal{S}_T)$, ψ is right-differentiable and left-continuous by Proposition 5.5, so

$$F(\tau_{i+1}^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) - F(\tau_i^n, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) = \int_0^{\tau_{i+1}^n - \tau_i^n} \mathcal{D}F(\tau_i^n + u, {}_n X_{\tau_i^n-}, {}_n A_{\tau_i^n-}) du \quad (85)$$

The second term in (83) can be written $\phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0)$ where $\phi(u) = F(\tau_i^n, {}_n X_{\tau_i^n-}^u, {}_n A_{\tau_i^n-})$. Since $F \in \mathbb{C}_b^{1,2}$, $\phi \in C^2(\mathbb{R}^d)$ and

$$\nabla \phi(u) = \nabla_\omega F(\tau_i^n, {}_n X_{\tau_i^n-}^u, {}_n A_{\tau_i^n-}), \quad \nabla^2 \phi(u) = \nabla_\omega^2 F(\tau_i^n, {}_n X_{\tau_i^n-}^u, {}_n A_{\tau_i^n-}).$$

Applying the Ito formula to $\phi(X(\tau_i^n + s) - X(\tau_i^n))$ yields:

$$\begin{aligned} \phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0) &= \int_{\tau_i^n}^{\tau_{i+1}^n} \nabla_\omega F(\tau_i^n, {}_n X_{\tau_i^n-}^{X(s) - X(\tau_i^n)}, {}_n A_{\tau_i^n-}) dX(s) \\ &+ \frac{1}{2} \int_{\tau_i^n}^{\tau_{i+1}^n} \text{tr} \left[t \nabla_\omega^2 F(\tau_i^n, {}_n X_{\tau_i^n-}^{X(s) - X(\tau_i^n)}, {}_n A_{\tau_i^n-}) d[X](s) \right] \end{aligned} \quad (86)$$

Summing over $i \geq 0$ and denoting $i(s)$ the index such that $s \in [\tau_{i(s)}^n, \tau_{i(s)+1}^n[$, we have shown:

$$\begin{aligned}
F(t, {}_n X_t, {}_n A_t) - F(0, X_0, A_0) &= \int_0^t \mathcal{D}F(s, {}_n X_{\tau_{i(s)}^n}, {}_n A_{\tau_{i(s)}^n}) ds \\
&\quad + \int_0^t \nabla_{\omega} F(\tau_{i(s)}^n, {}_n X_{\tau_{i(s)}^n}^{X(s)-X(\tau_{i(s)}^n)}, {}_n A_{\tau_{i(s)}^n}) dX(s) \\
&\quad + \frac{1}{2} \int_0^t \text{tr} \left[\nabla_{\omega}^2 F(\tau_{i(s)}^n, {}_n X_{\tau_{i(s)}^n}^{X(s)-X(\tau_{i(s)}^n)}, {}_n A_{\tau_{i(s)}^n}) \cdot d[X](s) \right] \quad (87)
\end{aligned}$$

$F(t, {}_n X_t, {}_n A_t)$ converges to $F(t, X_t, A_t)$ almost surely. Since all approximations of (X, A) appearing in the various integrals have a d_{∞} -distance from (X_s, A_s) less than $\eta_n \rightarrow 0$, the continuity at fixed times of $\mathcal{D}F$ and left-continuity $\nabla_{\omega} F$, $\nabla_{\omega}^2 F$ imply that the integrands appearing in the above integrals converge respectively to $\mathcal{D}F(s, X_s, A_s)$, $\nabla_{\omega} F(s, X_s, A_s)$, $\nabla_{\omega}^2 F(s, X_s, A_s)$ as $n \rightarrow \infty$. Since the derivatives are in \mathbb{B} the integrands in the various above integrals are bounded by a constant depending only on F, K, R and t but not on s nor on ω . The dominated convergence and the dominated convergence theorem for the stochastic integrals [62, Ch.IV Theorem 32] then ensure that the Lebesgue-Stieltjes integrals converge almost surely, and the stochastic integral in probability, to the terms appearing in (81) as $n \rightarrow \infty$.

Consider now the general case where X and A may be unbounded. Let K_n be an increasing sequence of compact sets with $\bigcup_{n \geq 0} K_n = \mathbb{R}^d$ and denote the stopping times

$$\tau_n = \inf\{s < t \mid X(s) \notin K^n \text{ or } |A(s)| > n\} \wedge t.$$

Applying the previous result to the stopped process $(X_{t \wedge \tau_n}, A_{t \wedge \tau_n})$ and noting that, by (78), $F(t, X_t, A_t) = F(t, X_t, A_{t-})$ leads to:

$$\begin{aligned}
F(t, X_{t \wedge \tau_n}, A_{t \wedge \tau_n}) - F(0, X_0, A_0) &= \int_0^{t \wedge \tau_n} \mathcal{D}F(u, X_u, A_u) du \\
&\quad + \frac{1}{2} \int_0^{t \wedge \tau_n} \text{tr} \left({}^t \nabla_{\omega}^2 F(u, X_u, A_u) d[X](u) \right) + \int_0^{t \wedge \tau_n} \nabla_{\omega} F(u, X_u, A_u) \cdot dX \\
&\quad + \int_{t \wedge \tau_n}^t \mathcal{D}F(u, X_{u \wedge \tau_n}, A_{u \wedge \tau_n}) du
\end{aligned}$$

The terms in the first line converges almost surely to the integral up to time t since $t \wedge \tau_n = t$ almost surely for n sufficiently large. For the same reason the last term converges almost surely to 0. \square

7 Weak functional calculus for square-integrable processes

The pathwise functional calculus presented in Sections 5 and 6 extends the Ito calculus to a large class of path-dependent functionals of semimartingales, of which we have already given several examples. Although in Section 6 we introduced a probability measure \mathbb{P} on the space $D([0, T], \mathbb{R}^d)$, under which X is a semimartingale, its probabilistic properties did not play any crucial role in the results, since the key results were shown to hold pathwise, \mathbb{P} -almost surely.

However, this pathwise calculus requires continuity of the functionals and their directional derivatives with respect to the supremum norm: without the (left-)continuity condition (5.4), the Functional Ito formula (Theorem 6.4) may fail to hold in a pathwise sense.³ This excludes some important examples of non-anticipative functionals, such as Ito stochastic integrals or the Ito map [52, 48], which describes the correspondence between the solution of a stochastic differential equation and the driving Brownian motion.

Another issue which arises when trying to apply the pathwise functional calculus to stochastic processes is that, in a probabilistic setting, functionals of a process X need only to be defined on a set of probability 1 with respect to \mathbb{P}^X : modifying the definition of a functional F outside the support of \mathbb{P}^X does not affect the image process $F(X)$ from a probabilistic perspective. So, in order to work with processes, we need to extend the previous framework to functionals which are not necessarily defined on the whole space Λ_T but only \mathbb{P}^X -almost everywhere, i.e. \mathbb{P}^X -equivalence classes of functionals.

This construction is well known in a finite-dimensional setting: given a reference measure μ on \mathbb{R}^d one can define notions of μ -almost everywhere regularity, weak derivatives and Sobolev regularity scales using duality relations with respect to the reference measure; this is the approach behind Schwartz's theory of distributions [65]. The Malliavin calculus [52], conceived by Paul Malliavin as a weak calculus for Wiener functionals, can be seen as an analogue of Schwartz distribution theory on the Wiener space, using the Wiener measure as reference measure.⁴ The idea is to construct an extension of the derivative operator using a duality relation, which extends the *integration by parts* formula. The closure of the derivative operator then defines the class of function(al)s to which the derivative may be applied in a weak sense.

We adopt here a similar approach, outlined in [11]: given a square-integrable Ito process X , we extend the pathwise functional calculus to a *weak non-anticipative functional calculus* whose domain of applicability includes all square integrable semimartingales adapted to the filtration \mathcal{F}^X generated by X .

We first define, in Section 7.1, a vertical derivative operator acting on *processes*, i.e. \mathbb{P} -equivalence classes of functionals of X , which is shown to be

³Failure to recognize this important point has led to several erroneous assertions in the literature, via a naive application of the functional Ito formula.

⁴This viewpoint on the Malliavin calculus was developed by Shigekawa [66], Watanabe [74] and Sugita [72, 71].

related to the martingale representation theorem (Section 7.2).

This operator ∇_X is then shown to be closable on the space of square-integrable martingales, and its closure is identified as the inverse of the Ito integral with respect to X : for $\phi \in \mathcal{L}^2(X)$, $\nabla_X (\int \phi.dX) = \phi$ (Theorem 7.7). In particular, we obtain a constructive version of the martingale representation theorem (Theorem 7.8), which states that for any square-integrable \mathcal{F}_t^X -martingale Y ,

$$Y(T) = Y(0) + \int_0^T \nabla_X Y.dX \quad \mathbb{P} - a.s.$$

This formula can be seen as a non-anticipative counterpart of the Clark-Haussmann-Ocone formula [5, 36, 37, 44, 56]. The integrand $\nabla_X Y$ is an adapted process which may be computed pathwise, so this formula is more amenable to numerical computations than those based on Malliavin calculus [12].

Section 7.4 further discusses the relation with the Malliavian derivative on the Wiener space: we show that the weak derivative ∇_X may be viewed as a non-anticipative “lifting” of the Malliavin derivative (Theorem 7.9): for square-integrable martingales Y whose terminal value $Y(T) \in \mathbf{D}^{1,2}$ is Malliavin-differentiable, we show that the vertical derivative is in fact a version of the predictable projection of the Malliavin derivative: $\nabla_X Y(t) = E[\mathbb{D}_t H | \mathcal{F}_t]$.

Finally, in Section 7.5 we extend the weak derivative to all square-integrable semimartingales and discuss an application of this construction to forward-backward SDEs (FBSDEs).

7.1 Vertical derivative of an adapted process

Throughout this Section, $X : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ will denote a continuous, \mathbb{R}^d -valued semi-martingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Since all processes we deal with are functionals of X , we will consider without loss of generality Ω to be the canonical space $D([0, T], \mathbb{R}^d)$, and $X(t, \omega) = \omega(t)$ to be the coordinate process. \mathbb{P} then denotes the law of the semimartingale. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the \mathbb{P} -completion of \mathcal{F}_{t+} .

We assume that

$$[X](t) = \int_0^t A(s)ds \tag{88}$$

for some cadlag process A with values in S_d^+ . Note that A need not be a semimartingale.

Any non-anticipative functional $F : \Lambda_T \rightarrow \mathbb{R}$ applied to X generates an \mathbb{F} -adapted process

$$Y(t) = F(t, X_t) = F(t, \{X(u \wedge t), u \in [0, T]\}) \tag{89}$$

However, the functional representation of the process Y is not unique: modifying F outside the topological support of \mathbb{P}^X does not change Y . In particular, the values taken by F outside \mathcal{W}_T do not affect Y . Yet, the definition of the vertical

$\nabla_\omega F$ seems to depend on the value of F for paths such as $\omega + e1_{[t,T]}$ which is clearly not continuous.

Theorem 5.28 gives a partial answer to this issue in the case where the topological support of \mathbb{P} is the full space $C^0([0, T], \mathbb{R}^d)$ (which is the case, for instance, for the Wiener process), but does not cover the wide variety of situations that may arise. To tackle this issue we need to define a vertical derivative operator which acts on (\mathbb{F} -adapted) processes, i.e. equivalence classes of (non-anticipative) functionals modulo an evanescent set. The functional Ito formula (72) provides the key to this construction. If $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$ then Theorem 6.4 identifies $\int_0^\cdot \nabla_\omega F(t, X_t) dM$ as the local martingale component of $Y(t) = F(t, X_t)$, which implies that it should be independent of the functional representation of Y :

Lemma 7.1. *Let $F^1, F^2 \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$, such that*

$$\forall t \in [0, T[, \quad F^1(t, X_t) = F^2(t, X_t) \quad \mathbb{P} - a.s.$$

Then

$${}^t[\nabla_\omega F^1(t, X_t) - \nabla_\omega F^2(t, X_t)]A(t-)[\nabla_\omega F^1(t, X_t) - \nabla_\omega F^2(t, X_t)] = 0 \quad (90)$$

outside an evanescent set.

Proof. Let $X = Z + M$ where Z is a continuous process with finite variation and M is a continuous local martingale. There exists $\Omega_1 \in \mathcal{F}$ such that $\mathbb{P}(\Omega_1) = 1$ and for $\omega \in \Omega_1$ the path of $t \mapsto X(t, \omega)$ is continuous and $t \mapsto A(t, \omega)$ is cadlag. Theorem 6.4 implies that the local martingale part of $0 = F^1(t, X_t) - F^2(t, X_t)$ can be written:

$$0 = \int_0^t (\nabla_\omega F^1(u, X_u) - \nabla_\omega F^2(u, X_u)) dM(u) \quad (91)$$

Computing its quadratic variation, we have, on Ω_1

$$0 = \int_0^t \frac{1}{2} {}^t[\nabla_\omega F^1(u, X_u) - \nabla_\omega F^2(u, X_u)]A(u-)[\nabla_\omega F^1(u, X_u) - \nabla_\omega F^2(u, X_u)]du \quad (92)$$

and $\nabla_\omega F^1(t, X_t) = \nabla_\omega F^1(t, X_{t-})$ since X is continuous. So on Ω_1 the integrand in (92) is left-continuous; therefore (92) implies that for $t \in [0, T[, \omega \in \Omega_1$,

$${}^t[\nabla_\omega F^1(t, X_t) - \nabla_\omega F^2(t, X_t)]A(t-)[\nabla_\omega F^1(t, X_t) - \nabla_\omega F^2(t, X_t)] = 0.$$

□

Thus, if we assume that in (88) $A(t-)$ is almost surely non-singular, then Lemma 7.1 allows to define intrinsically the pathwise derivative of any process Y with a smooth functional representation $Y(t) = F(t, X_t)$.

Assumption 7.1 (Non-degeneracy of local martingale component). $A(t)$ in (88) is non-singular almost-everywhere:

$$\det(A(t)) \neq 0 \quad dt \times d\mathbb{P} - a.e.$$

Definition 7.2 (Vertical derivative of a process). Define $\mathcal{C}_{\text{loc}}^{1,2}(X)$ the set of \mathcal{F}_t -adapted processes Y which admit a functional representation in $\mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$:

$$\mathcal{C}_{\text{loc}}^{1,2}(X) = \{Y, \exists F \in \mathbb{C}_{\text{loc}}^{1,2}, Y(t) = F(t, X_t) \quad dt \times \mathbb{P} - a.e.\} \quad (93)$$

Under Assumption 7.1, for any $Y \in \mathcal{C}_{\text{loc}}^{1,2}(X)$, the predictable process

$$\nabla_X Y(t) = \nabla_\omega F(t, X_t)$$

is uniquely defined up to an evanescent set, independently of the choice of $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$ in the representation (93). We will call the process $\nabla_X Y$ the *vertical derivative* of Y with respect to X .

Although the functional $\nabla_\omega F : \Lambda_T \rightarrow \mathbb{R}^d$ does depend on the choice of the functional representation F in (89), the *process* $\nabla_X Y(t) = \nabla_\omega F(t, X_t)$ obtained by computing $\nabla_\omega F$ along the paths of X does not.

The vertical derivative ∇_X defines a correspondence between \mathbb{F} -adapted processes: it maps any process $Y \in \mathcal{C}_b^{1,2}(X)$ to an \mathbb{F} -adapted process $\nabla_X Y$. It defines a linear operator

$$\nabla_X : \mathcal{C}_b^{1,2}(X) \mapsto \mathcal{C}_l^{0,0}(X)$$

on the algebra of smooth functionals $\mathcal{C}_b^{1,2}(X)$, with the following properties: for any $Y, Z \in \mathcal{C}_b^{1,2}(X)$, and any \mathbb{F} -predictable process λ , we have

- \mathcal{F}_t -linearity: $\nabla_X(Y + \lambda Z)(t) = \nabla_X Y(t) + \lambda(t) \nabla_X Z(t)$
- Differentiation of products: $\nabla_X(YZ)(t) = Z(t)\nabla_X Y(t) + Y(t)\nabla_X Z(t)$
- Composition rule: If $U \in \mathcal{C}_b^{1,2}(Y)$, then $U \in \mathcal{C}_b^{1,2}(X)$ and

$$\forall t \in [0, T], \quad \nabla_X U(t) = \nabla_Y U(t) \cdot \nabla_X Y(t). \quad (94)$$

The vertical derivative operator ∇_X has important links with the Ito stochastic integral. The starting point is the following remark:

Proposition 7.3 (Representation of smooth local martingales). *For any local martingale $Y \in \mathcal{C}_{\text{loc}}^{1,2}(X)$ we have the representation*

$$Y(T) = Y(0) + \int_0^T \nabla_X Y \cdot dM \quad (95)$$

where M is the local martingale component of X .

Proof. Since $Y \in \mathcal{C}_{\text{loc}}^{1,2}(X)$, there exists $F \in \mathcal{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$ such that $Y(t) = F(t, X_t)$. Theorem 6.4 then implies that for $t \in [0, T]$:

$$Y(t) - Y(0) = \int_0^t \mathcal{D}F(u, X_u) du + \frac{1}{2} \int_0^t \text{tr}({}^t \nabla_{\omega}^2 F(u, X_u) \cdot d[X](u)) \\ + \int_0^t \nabla_{\omega} F(u, X_u) dZ(u) + \int_0^t \nabla_{\omega} F(t, X_t) dM(u)$$

Given the regularity assumptions on F , all terms in this sum are continuous process with finite variation while the last term is a continuous local martingale. By uniqueness of the semimartingale decomposition of Y , we thus have: $Y(t) = \int_0^t \nabla_{\omega} F(u, X_u) dM(u)$. Since $F \in \mathcal{C}_l^{0,0}([0, T])$ $Y(t) \rightarrow F(T, X_T)$ as $t \rightarrow T$, so the stochastic integral also is well-defined for $t \in [0, T]$. \square

We now explore further the consequences of this property and its link with the martingale representation theorem.

7.2 Martingale representation formula

Consider now the case where X is a Brownian martingale:

$$X(t) = X(0) + \int_0^t \sigma(u) \cdot dW(u),$$

where σ is a process adapted to \mathcal{F}_t^W satisfying

$$E \left(\int_0^T \|\sigma(t)\|^2 dt \right) < \infty, \quad \det(\sigma(t)) \neq 0 \quad dt \times d\mathbb{P} - a.e. \quad (96)$$

Then X is a square-integrable martingale with the predictable representation property [45, 63]: for any square-integrable \mathcal{F}_T measurable random variable H , or equivalently, any square-integrable \mathbb{F} -martingale Y defined by $Y(t) = E[H | \mathcal{F}_t]$, there exists a unique \mathbb{F} -predictable process ϕ such that

$$Y(t) = Y(0) + \int_0^t \phi \cdot dX, \quad \text{i.e.,} \quad H = Y(0) + \int_0^T \phi \cdot dX \quad (97)$$

$$\text{and} \quad E \left(\int_0^T \text{tr}(\phi(u) \cdot {}^t \phi(u) d[X](u)) \right) < \infty. \quad (98)$$

The classical proof of this representation result (see e.g. [63]) is non-constructive and a lot of effort has been devoted to obtaining an explicit form for ϕ in terms of Y , using Markovian methods [17, 25, 27, 42, 58] or Malliavin calculus [5, 3, 44, 37, 56, 55]. We will now see that the vertical derivative gives a simple representation of the integrand ϕ .

If $Y \in \mathcal{C}_{\text{loc}}^{1,2}(X)$ is a martingale then Proposition 7.3 applies so, comparing (97) with (95) suggests $\phi = \nabla_X Y$ as the obvious candidate. We now show that this guess is indeed correct in the square-integrable case.

Let $\mathcal{L}^2(X)$ be the Hilbert space of \mathbb{F} -predictable processes such that

$$\|\phi\|_{\mathcal{L}^2(X)}^2 = E \left(\int_0^T \text{tr}(\phi(u) \cdot {}^t\phi(u) d[X](u)) \right) < \infty \quad (99)$$

This is the space of integrands for which one defines the usual L^2 extension of the Ito integral with respect to X .

Proposition 7.4 (Representation of smooth L^2 martingales). *If $Y \in \mathcal{C}_{\text{loc}}^{1,2}(X)$ is a square-integrable martingale then*

$$\nabla_X Y \in \mathcal{L}^2(X) \quad \text{and} \quad Y(T) = Y(0) + \int_0^T \nabla_X Y \cdot dX \quad (100)$$

Proof. Applying Proposition 7.3 to Y we obtain that $Y = Y(0) + \int_0^\cdot \nabla_X Y \cdot dX$. Since Y is square-integrable, the Ito integral $\int_0^\cdot \nabla_X Y \cdot dX$ is square-integrable and the Ito isometry formula yields

$$E(\|Y(T) - Y(0)\|^2) = E \left| \int_0^T \nabla_X Y \cdot dX \right|^2 = E \left(\int_0^T \text{tr}({}^t \nabla_X Y(u) \cdot \nabla_X Y(u) d[X](u)) \right) < \infty$$

so $\nabla_X Y \in \mathcal{L}^2(X)$. The uniqueness of the predictable representation in $\mathcal{L}^2(X)$ then allows us to conclude. \square

7.3 Weak derivative for square-integrable functionals

Let $\mathcal{M}^2(X)$ be the space of square-integrable \mathbb{F} -martingales with initial value zero, with the norm $\|Y\| = \sqrt{E|Y(T)|^2}$.

We will now use Proposition 7.4 to extend ∇_X to a continuous functional on $\mathcal{M}^2(X)$.

First, observe that the predictable representation property implies that the stochastic integral with respect to X defines a bijective isometry

$$\begin{aligned} I_X : \mathcal{L}^2(X) &\mapsto \mathcal{M}^2(X) \\ \phi &\mapsto \int_0^\cdot \phi \cdot dX \end{aligned} \quad (101)$$

Then, by Proposition 7.4, the vertical derivative ∇_X defines a continuous map

$$\nabla_X : \mathcal{C}_b^{1,2}(X) \cap \mathcal{M}^2(X) \mapsto \mathcal{L}^2(X).$$

on the set of 'smooth martingales'

$$D(X) = \mathcal{C}_b^{1,2}(X) \cap \mathcal{M}^2(X). \quad (102)$$

This isometry property can now be used to extend ∇_X to the entire space $\mathcal{M}^2(X)$, using a density argument:

Lemma 7.5 (Density of $\mathcal{C}_b^{1,2}(X)$ in $\mathcal{M}^2(X)$). $\{\nabla_X Y, Y \in D(X)\}$ is dense in $\mathcal{L}^2(X)$ and $D(X) = \mathcal{C}_b^{1,2}(X) \cap \mathcal{M}^2(X)$ is dense in $\mathcal{M}^2(X)$.

Proof. We first observe that the set of cylindrical non-anticipative processes of the form

$$\phi_{n,f,(t_1,\dots,t_n)}(t) = f(X(t_1), \dots, X(t_n))1_{t>t_n}$$

where $n \geq 1$, $0 \leq t_1 < \dots < t_n \leq T$ and $f \in C_b^\infty(\mathbb{R}^n, \mathbb{R})$ is a total set in $\mathcal{L}^2(X)$ i.e. their linear span, which we denote by U , is dense in $\mathcal{L}^2(X)$. For such an integrand $\phi_{n,f,(t_1,\dots,t_n)}$, the stochastic integral with respect to X is given by the martingale

$$Y(t) = I_X(\phi_{n,f,(t_1,\dots,t_n)})(t) = F(t, X_t)$$

where the functional F is defined on Λ_T as:

$$F(t, \omega) = f(\omega(t_1-), \dots, \omega(t_n-))(\omega(t) - \omega(t_n))1_{t>t_n}$$

so that:

$$\nabla_\omega F(t, \omega) = f(\omega(t_1-), \dots, \omega(t_n-))1_{t>t_n}, \nabla_\omega^2 F(t, \omega) = 0, \mathcal{D}F(t, \omega_t) = 0$$

which shows that $F \in \mathbb{C}_b^{1,2}$ (see Example 5.11). Hence, $Y \in \mathcal{C}_b^{1,2}(X)$. Since f is bounded, Y is obviously square integrable so $Y \in D(X)$. Hence $I_X(U) \subset D(X)$.

Since I_X is a bijective isometry from $\mathcal{L}^2(X)$ to $\mathcal{M}^2(X)$, the density of U in $\mathcal{L}^2(X)$ entails the density of $I_X(U)$ in $\mathcal{M}^2(X)$. \square

This leads to a useful characterization of the vertical derivative ∇_X :

Proposition 7.6 (Integration by parts on $D(X)$).

Let $Y \in D(X) = \mathcal{C}_b^{1,2}(X) \cap \mathcal{M}^2(X)$. Then $\nabla_X Y$ is the unique element of $\mathcal{L}^2(X)$ which satisfies

$$\forall Z \in D(X), \quad E(Y(T)Z(T)) = E\left(\int_0^T \text{tr}(\nabla_X Y(t)^t \nabla_X Z(t) d[X](t))\right). \quad (103)$$

Proof. Let $Y, Z \in D(X)$. Then Y is a square-integrable martingale with $Y(0) = 0$ and $E[|Y(T)|^2] < \infty$. Applying Proposition 7.4 to Y, Z , we obtain $Y = \int \nabla_X Y.dX, Z = \int \nabla_X Z.dX$, so

$$E[Y(T)Z(T)] = E\left[\int_0^T \nabla_X Y.dX \int_0^T \nabla_X Z.dX\right]$$

Applying the Ito isometry formula yields the (103). If we now consider another process $\psi \in \mathcal{L}^2(X)$ which verifies

$$\forall Z \in D(X), \quad E(Y(T)Z(T)) = E\left(\int_0^T \text{tr}(\psi(t)^t \nabla_X Z(t) d[X])\right),$$

then subtracting from (103) yields that

$$\forall Z \in D(X), \quad \langle \psi - \nabla_X Y, \nabla_X Z \rangle_{\mathcal{L}^2(X)} = 0.$$

which implies $\psi = \nabla_X Y$ since, by Lemma 7.5 $\{\nabla_X Z, Z \in D(X)\}$ is dense in $\mathcal{L}^2(X)$. \square

We can rewrite (103) as:

$$E \left(Y(T) \int_0^T \phi \cdot dX \right) = E \left(\int_0^T \text{tr}(\nabla_X Y(t)^t \phi(t) d[X](t)) \right) \quad (104)$$

which can be understood as an 'integration by parts formula' on $[0, T] \times \Omega$ with respect to the measure $d[X] \times d\mathbb{P}$.

The integration by parts formula, and the density of the domain are the two ingredients needed to show that the operator is closable on $\mathcal{M}^2(X)$:

Theorem 7.7 (Extension of ∇_X to $\mathcal{M}^2(X)$). *The vertical derivative $\nabla_X : D(X) \mapsto \mathcal{L}^2(X)$ admits a unique continuous extension to $\mathcal{M}^2(X)$*

$$\begin{aligned} \nabla_X : \mathcal{M}^2(X) &\mapsto \mathcal{L}^2(X) \\ \int_0^\cdot \phi \cdot dX &\mapsto \phi \end{aligned} \quad (105)$$

which is a bijective isometry characterized by the integration by parts formula: for $Y \in \mathcal{M}^2(X)$, $\nabla_X Y$ is the unique element of $\mathcal{L}^2(X)$ such that

$$\forall Z \in D(X), \quad E[Y(T)Z(T)] = E \left(\int_0^T \text{tr}(\nabla_X Y(t)^t \nabla_X Z(t) d[X](t)) \right) \quad (106)$$

In particular, ∇_X is the adjoint of the Ito stochastic integral

$$\begin{aligned} I_X : \mathcal{L}^2(X) &\mapsto \mathcal{M}^2(X) \\ \phi &\mapsto \int_0^\cdot \phi \cdot dX \end{aligned} \quad (107)$$

in the following sense:

$$\forall \phi \in \mathcal{L}^2(X), \quad \forall Y \in \mathcal{M}^2(X), \quad E[Y(T) \int_0^T \phi \cdot dX] = \langle \nabla_X Y, \phi \rangle_{\mathcal{L}^2(X)}. \quad (108)$$

Proof. Any $Y \in \mathcal{M}^2(X)$ may be written as $Y(t) = \int_0^t \phi(s) dX(s)$ with $\phi \in \mathcal{L}^2(X)$, which is uniquely defined $d[X] \times d\mathbb{P}$ a.e. The Ito isometry formula then guarantees that (106) holds for ϕ . To show that (106) uniquely characterizes ϕ , consider $\psi \in \mathcal{L}^2(X)$ which also satisfies (106), then, denoting $I_X(\psi) = \int_0^\cdot \psi dX$ its stochastic integral with respect to X , (106) then implies that

$$\forall Z \in D(X), \quad \langle I_X(\psi) - Y, Z \rangle_{\mathcal{M}^2(X)} = E[(Y(T) - \int_0^T \psi dX)Z(T)] = 0$$

which implies $I_X(\psi) = Y$ $d[X] \times d\mathbb{P}$ a.e. since by construction $D(X)$ is dense in $\mathcal{M}^2(X)$. Hence, $\nabla_X : D(X) \mapsto \mathcal{L}^2(X)$ is closable on $\mathcal{M}^2(X)$. \square

We have thus extended Dupire's pathwise vertical derivative ∇_ω to a weak derivative ∇_X defined for all square-integrable \mathbb{F} -martingales, which is the inverse of the Ito integral I_X with respect to X :

$$\forall \phi \in \mathcal{L}^2(X), \quad \nabla_X \left(\int_0^\cdot \phi dX \right) = \phi \quad (109)$$

holds in the sense of equality in $\mathcal{L}^2(X)$.

The above results now allow us to state a general version of the martingale representation formula, valid for all square-integrable martingales:

Theorem 7.8 (Martingale representation formula: general case). *For any square-integrable \mathbb{F} -martingale Y ,*

$$\forall t \in [0, T], \quad Y(t) = Y(0) + \int_0^t \nabla_X Y.dX \quad \mathbb{P} - a.s.$$

This relation, which can be understood as a stochastic version of the 'fundamental theorem of calculus' for martingales, shows that ∇_X is a 'stochastic derivative' in the sense of Zabczyk [76] and Davis [17].

Theorem 7.7 suggests that the space of square-integrable martingales can be seen as a 'Martingale Sobolev space' of order 1 constructed above $\mathcal{L}^2(X)$ [11]. However, since for $Y \in \mathcal{M}^2(X)$, $\nabla_X Y$ is not a martingale, Theorem 7.7 does not allow to iterate this weak differentiation to higher orders. We will see in Section 7.5 how to extend this construction to semimartingales and construct Sobolev spaces of arbitrary order for the operator ∇_X .

7.4 Relation with the Malliavin derivative

The above construction holds in particular in the case where $X = W$ is a Wiener process. Consider the canonical Wiener space $(\Omega_0 = C_0([0, T], \mathbb{R}^d), \|\cdot\|_\infty, \mathbb{P})$ endowed with the filtration of the canonical process W , which is a Wiener process under \mathbb{P} .

Theorem 7.7 then applies to $X = W$ is the Wiener process and allows to define, for any square-integrable Brownian martingale $Y \in \mathcal{M}^2(W)$, the *vertical derivative* $\nabla_W Y \in \mathcal{L}^2(W)$ of Y with respect to W .

Note that the Gaussian properties of W play no role in the construction of the operator ∇_W . We now compare the Brownian vertical derivative ∇_W with the *Malliavin derivative* [52, 3, 4, 68].

Consider an \mathcal{F}_T -measurable functional $H = H(X(t), t \in [0, T]) = H(X_T)$ with $E[|H|^2] < \infty$. If H is differentiable in the Malliavin sense [3, 52, 55, 68] e.g. $H \in \mathbf{D}^{1,2}$ with Malliavin derivative $\mathbb{D}_t H$, then the Clark-Haussmann-Ocone formula [56, 55] gives a stochastic integral representation of H in terms of the Malliavin derivative of H :

$$H = E[H] + \int_0^T {}^p E[\mathbb{D}_t H | \mathcal{F}_t] dW(t) \quad (110)$$

where ${}^pE[\mathbb{D}_t H|\mathcal{F}_t]$ denotes the predictable projection of the Malliavin derivative. This yields a stochastic integral representation of the martingale $Y(t) = E[H|\mathcal{F}_t]$:

$$Y(t) = E[H|\mathcal{F}_t] = E[H] + \int_0^t {}^pE[\mathbb{D}_t H|\mathcal{F}_u]dW(u)$$

Related martingale representations have been obtained under a variety of conditions [3, 17, 27, 44, 58, 55].

Denote by

- $L^2([0, T] \times \Omega)$ the set of \mathcal{F}_T -measurable maps $\phi : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ on $[0, T]$ with

$$E \int_0^T \|\phi(t)\|^2 dt < \infty$$

- $E[\cdot|\mathbb{F}]$ the conditional expectation operator with respect to the filtration $\mathbb{F} = (\mathcal{F}_t, t \in [0, T])$:

$$E[\cdot|\mathbb{F}] : H \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \rightarrow (E[H|\mathcal{F}_t], t \in [0, T]) \in \mathcal{M}^2(W)$$

- \mathbb{D} the Malliavin derivative operator, which associates to a random variable $H \in \mathbf{D}^{1,2}(0, T)$ the (anticipative) process $(\mathbb{D}_t H)_{t \in [0, T]} \in L^2([0, T] \times \Omega)$.

Theorem 7.9 (Intertwining formula). *The following diagram is commutative in the sense of $dt \times d\mathbb{P}$ equality:*

$$\begin{array}{ccc} \mathcal{M}^2(W) & \xrightarrow{\nabla_W} & \mathcal{L}^2(W) \\ \uparrow E[\cdot|\mathbb{F}] & & \uparrow E[\cdot|\mathbb{F}] \\ \mathbf{D}^{1,2} & \xrightarrow{\mathbb{D}} & L^2([0, T] \times \Omega) \end{array}$$

In other words, the conditional expectation operator intertwines ∇_W with the Malliavin derivative:

$$\forall H \in \mathbf{D}^{1,2}(\Omega_0, \mathcal{F}_T, \mathbb{P}), \quad \nabla_W (E[H|\mathcal{F}_t]) = E[\mathbb{D}_t H|\mathcal{F}_t] \quad (111)$$

Proof. The Clark-Haussmann-Ocone formula [56] gives

$$\forall H \in \mathbf{D}^{1,2}, \quad H = E[H] + \int_0^T {}^pE[\mathbb{D}_t H|\mathcal{F}_t]dW_t \quad (112)$$

where ${}^pE[\mathbb{D}_t H|\mathcal{F}_t]$ denotes the predictable projection of the Malliavin derivative. On other hand from Theorem 7.4 we know that

$$H = E[H] + \int_0^T \nabla_W Y(t) dW(t) \quad (113)$$

where $Y(t) = E[H|\mathcal{F}_t]$. Hence ${}^pE[\mathbb{D}_t H|\mathcal{F}_t] = \nabla_W E[H|\mathcal{F}_t]$, $dt \times d\mathbb{P}$ almost everywhere. \square

The *predictable projection* on \mathcal{F}_t [19, Vol. I]) can be viewed as a morphism which “lifts” relations obtained in the framework of Malliavin calculus into relations between non-anticipative quantities, where the Malliavin derivative and the Skorokhod integral are replaced, respectively, by the vertical derivative ∇_W and the Ito stochastic integral.

Rewriting(111) as

$$\forall t \leq T, \quad (\nabla_W Y)(t) = E[\mathbb{D}_t(Y(T)) | \mathcal{F}_t],$$

we can note that the choice of $T \geq t$ is arbitrary: for any $h \geq 0$, we have

$$\nabla_W Y(t) = {}^p E[\mathbb{D}_t(Y(t+h)) | \mathcal{F}_t].$$

Taking the limit $h \rightarrow 0+$ yields

$$\nabla_W Y(t) = \lim_{h \rightarrow 0+} {}^p E[\mathbb{D}_t Y(t+h) | \mathcal{F}_t] \quad (114)$$

The right hand side is sometimes written, with some abuse of notation, as $\mathbb{D}_t Y(t)$, and called the ‘diagonal Malliavin derivative’. From a computational viewpoint, unlike the Clark-Haussmann-Ocone representation which requires to simulate the *anticipative* process $\mathbb{D}_t H$ and compute conditional expectations, $\nabla_X Y$ only involves non-anticipative quantities which can be computed path by path. It is thus more amenable to numerical computations [12].

	Malliavin derivative \mathbb{D}	Vertical Derivative ∇_W
Perturbations	$H^1([0, T], \mathbb{R}^d)$	$\{e \mathbf{1}_{[t, T]}, e \in \mathbb{R}^d, t \in [0, T]\}$
Domain	$\mathbf{D}^{1,2}(\mathcal{F}_T)$	$\mathcal{M}^2(W)$
Range	$L^2([0, T] \times \Omega)$	$\mathcal{L}^2(W)$
Measurability	\mathcal{F}_T^W -measurable (anticipative)	\mathcal{F}_t^W -measurable (non-anticipative)
Adjoint	Skorokhod integral	Ito integral

Table 1: Comparison of the Malliavin derivative and the vertical derivative.

The commutative diagram above shows that ∇_W may be seen as ‘non-anticipative lifting’ of the Malliavin derivative: the Malliavin derivative on random variables (lower level) is lifted to the vertical derivative of non-anticipative processes (upper level) by the conditional expectation operator.

Having pointed out the parallels between these constructions, we must note however that our construction of the weak derivative operator ∇_X works for any square-integrable continuous martingale X and does not involve any Gaussian or probabilistic properties of X . Also, the domain of the vertical derivative spans all square-integrable functionals, whereas the Malliavin derivative has a strictly smaller domain $\mathbf{D}^{1,2}$.

Regarding this last point: in fact the above correspondence *can* be extended to the full space $L^2(\mathbb{P}, \mathcal{F}_T)$ using Hida's *white noise analysis* [38] and the use of distribution-valued processes. Aase et al [1] extend the Malliavin derivative as a distribution-valued process $\widetilde{\mathbb{D}}_t F$ on $L^2(\mathbb{P}, \mathcal{F}_T)$ where \mathbb{P} is the Wiener measure and \mathcal{F}_T the Brownian filtration, and derive a generalized Clark- Haussman-Ocone formula : for $F \in L^2(\mathbb{P}, \mathcal{F}_T)$,

$$F = E[F] + \int_0^T E[\widetilde{\mathbb{D}}_t F | \mathcal{F}_t] \diamond W(t) dt \quad (115)$$

where \diamond denotes a 'Wick product' [1].

Although this white-noise extension of the Malliavin derivative is not a stochastic process but a distribution-valued process, the uniqueness of the martingale representation shows that its predictable projection $E[\widetilde{\mathbb{D}}_t F | \mathcal{F}_t]$ is a bona-fide square-integrable process which is a version of the (weak) vertical derivative:

$$E[\widetilde{\mathbb{D}}_t F | \mathcal{F}_t] = \nabla_W Y(t) \quad dt \times d\mathbb{P} - a.e.$$

This extends the commutative diagram to the full space

$$\begin{array}{ccc} \mathcal{M}^2(W) & \xrightarrow{\nabla_W} & \mathcal{L}^2(W) \\ (E[\cdot | \mathcal{F}_t])_{t \in [0, T]} \uparrow & & \uparrow (E[\cdot | \mathcal{F}_t])_{t \in [0, T]} \\ \mathbf{L}^2(\Omega, \mathcal{F}_T, \mathbb{P}) & \xrightarrow{\widetilde{\mathbb{D}}} & \mathbf{H}^{-1}([0, T] \times \Omega) \end{array}$$

However, the formulae obtained using the vertical derivative ∇_W only involve Ito stochastic integrals of non-anticipative processes and do not require any recourse to distribution-valued processes.

But, most importantly, the construction of the vertical derivative and the associated martingale representation formulae make no use of the Gaussian properties of the underlying martingale X and extend well beyond the Wiener process, to any square-integrable martingale.

7.5 Extension to semimartingales

We now show how ∇_X can be extended to all square-integrable \mathbb{F} -*semimartingales*.

Define $\mathcal{A}^2(\mathbb{F})$ as the set of \mathbb{F} -adapted processes H with finite variation $\int_0^T |dH| < \infty$ such that

$$V_1(H) = \int_0^T |dH| \in L^2(\Omega, \mathbb{F}, \mathbb{P}), \quad \text{i.e.} \quad \mathbb{E} \left(\left| \int_0^T |dH| \right|^2 \right) < \infty,$$

equipped with the norm

$$\|H\|_{\mathcal{A}^2}^2 = \mathbb{E} (|H(0)|^2 + V_1(H)^2)$$

and consider the direct sum

$$\mathcal{S}^{1,2}(X) = \mathcal{M}^2(X) \oplus \mathcal{A}^2(\mathbb{F}). \quad (116)$$

Then any process $S \in \mathcal{S}^{1,2}(X)$ is an \mathbb{F} -adapted special semimartingale with a unique decomposition

$$S = M + H$$

where $M \in \mathcal{M}^2(X)$ is a square-integrable \mathbb{F} -martingale with $M(0) = 0$ and $H \in \mathcal{A}^2(\mathbb{F})$ with $H(0) = S(0)$. The norm defined by

$$\|S\|_{1,2}^2 = E ([S](T)) + \|H\|_{\mathcal{A}^2}^2 \quad (117)$$

defines a Hilbert space structure on $\mathcal{S}^{1,2}(X)$ for which $\mathcal{A}^2(\mathbb{F})$ and $\mathcal{M}^2(X)$ are closed subspaces of $\mathcal{S}^{1,2}(X)$.

We will refer to processes in $\mathcal{S}^{1,2}(X)$ as 'square-integrable \mathbb{F} -semimartingales'.

Theorem 7.10 (Weak derivative on $\mathcal{S}^{1,2}(X)$). *The vertical derivative $\nabla_X : \mathcal{C}^{1,2}(X) \cap \mathcal{S}^{1,2}(X) \mapsto \mathcal{L}^2(X)$ admits a unique continuous extension to the space $\mathcal{S}^{1,2}(X)$ of square-integrable \mathbb{F} -semimartingales, such that:*

1. *The restriction of ∇_X to square-integrable martingales is a bijective isometry*

$$\begin{aligned} \nabla_X : \mathcal{M}^2(X) &\mapsto \mathcal{L}^2(X) \\ \int_0^\cdot \phi.dX &\mapsto \phi \end{aligned} \quad (118)$$

which is the inverse of the Ito integral with respect to X .

2. *For any finite variation process $H \in \mathcal{A}^2(X)$, $\nabla_X H = 0$.*

Clearly, both properties are necessary since ∇_X already verifies them on $\mathcal{S}^{1,2}(X) \cap \mathcal{C}_b^{1,2}(X)$, which is dense in $\mathcal{S}^{1,2}(X)$. Continuity of the extension then imposes $\nabla_X H = 0$ for any $H \in \mathcal{A}^2(\mathbb{F})$. Uniqueness results from the fact that (116) is a continuous direct sum of closed subspaces. The semimartingale decomposition of $S \in \mathcal{S}^{1,2}(X)$ may then be interpreted as a projection on the closed subspaces $\mathcal{A}^2(\mathbb{F})$ and $\mathcal{M}^2(X)$.

The following characterization follows from the definition of $\mathcal{S}^{1,2}(X)$:

Proposition 7.11 (Description of $\mathcal{S}^{1,2}(X)$).

$$\forall S \in \mathcal{S}^{1,2}(X), \exists ! (h, \phi) \in \mathcal{L}^2(dt \times d\mathbb{P}) \times \mathcal{L}^2(X), \quad S(t) = S(0) + \int_0^t h(u)du + \int_0^t \phi.dX$$

where h, ϕ are given by

$$\phi = \nabla_X S, \quad h(t) = \frac{d}{dt} \left(S(t) - \int_0^t \nabla_X S.dX \right)$$

The map

$$\nabla_X : \mathcal{S}^{1,2}(X) \mapsto \mathcal{L}^2(X)$$

is continuous. So $\mathcal{S}^{1,2}(X)$ may be seen as a Sobolev space of order 1 constructed above $\mathcal{L}^2(X)$, which explains our notation.

We can iterate this construction and define a scale of 'Sobolev' spaces of \mathbb{F} -semimartingales with respect to ∇_X :

Theorem 7.12 (The space $\mathcal{S}^{k,2}(X)$). *Let $\mathcal{S}^{0,2}(X) = \mathcal{L}^2(X)$ and define, for $k \geq 2$,*

$$\mathcal{S}^{k,2}(X) := \{S \in \mathcal{S}^{1,2}(X), \nabla_X S \in \mathcal{S}^{k-1,2}(X)\} \quad (119)$$

equipped with the norm

$$\|S\|_{k,2}^2 = \|H\|_{\mathcal{A}^2}^2 + \sum_{j=1}^k \|\nabla_X^j S\|_{\mathcal{L}^2(X)}^2. \quad (120)$$

Then $(\mathcal{S}^{k,2}(X), \|\cdot\|_{k,2})$ is a Hilbert space, $\mathcal{S}^{k,2}(X) \cap \mathcal{C}_b^{1,2}(X)$ is dense in $\mathcal{S}^{k,2}(X)$ and the maps

$$\nabla_X : \mathcal{S}^{k,2}(X) \mapsto \mathcal{S}^{k-1,2}(X)$$

$$I_X : \mathcal{S}^{k-1,2}(X) \mapsto \mathcal{S}^{k,2}(X)$$

are continuous.

The above construction enables to define

$$\nabla_X^k : \mathcal{S}^{k,2}(X) \mapsto \mathcal{L}^2(X)$$

as a continuous operator on $\mathcal{S}^{k,2}(X)$. This construction may be viewed as a non-anticipative counterpart of Gaussian Sobolev spaces introduced in the Malliavin calculus [66, 72, 71, 74, 26]; unlike those constructions, the embeddings involve the Ito stochastic integral and do not rely on the Gaussian nature of the underlying measure.

We can now use these ingredients to extend the horizontal derivative (Definition 5.7) to $S \in \mathcal{S}^{2,2}(X)$. First, note that if $S = F(X)$ with $F \in \mathcal{C}_b^{1,2}(\mathcal{W}_T)$, then

$$S(t) - S(0) - \int_0^t \nabla_X S \cdot dX - \frac{1}{2} \int_0^t \langle \nabla_X^2 S, d[X] \rangle \quad (121)$$

is equal to $\int_0^t \mathcal{D}F(u) du$. For $S \in \mathcal{S}^{2,2}(X)$, the process defined by (121) is almost-surely absolutely continuous with respect to the Lebesgue measure: we use its Radon-Nikodym derivative to define a weak extension of the horizontal derivative.

Definition 7.13 (Extension of the horizontal derivative to $\mathcal{S}^{2,2}$). For $S \in \mathcal{S}^{2,2}$ there exists a unique \mathbb{F} -adapted process $\mathcal{D}S \in \mathcal{L}^2(dt \times d\mathbb{P})$ such that

$$\forall t \in [0, T], \quad \int_0^t \mathcal{D}S(u)du = S(t) - S(0) - \int_0^t \nabla_X S \cdot dX - \frac{1}{2} \int_0^t \nabla_X^2 S \cdot d[X] \quad (122)$$

$$\text{and } E \left(\int_0^T |\mathcal{D}S(t)|^2 dt \right) < \infty.$$

Formula (122) shows that $\mathcal{D}S$ may be interpreted as the 'Stratonovich drift' of S , i.e.

$$\int_0^t \mathcal{D}S(u)du = S(t) - S(0) - \int_0^t \nabla_X S \circ dX$$

where \circ denotes the Stratonovich integral.

The following property is a straightforward consequence of Definition 7.13:

Proposition 7.14. *Let $(Y^n)_{n \geq 1}$ be a sequence in $\mathcal{S}^{2,2}(X)$. If*

$$\|Y^n - Y\|_{2,2} \xrightarrow{n \rightarrow \infty} 0, \quad \text{then } \mathcal{D}Y^n(t) \rightarrow \mathcal{D}Y(t) \quad dt \times d\mathbb{P} - a.e.$$

The above discussion implies that Proposition 7.14 does not hold if $\mathcal{S}^{2,2}(X)$ is replaced by $\mathcal{S}^{1,2}(X)$ (or $\mathcal{L}^2(X)$).

If $S = F(X)$ with $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$, the Functional Ito formula (Theorem 6.4) then implies that $\mathcal{D}S$ is a version of the process $\mathcal{D}F(t, X_t)$. However, as the following example shows, Definition 7.13 is much more general and applies to functionals which are not even continuous in supremum norm.

Example 7.15 (Multiple stochastic integrals). Let $(\Phi(t))_{t \in [0, T]} = (\Phi_1(t), \dots, \Phi_d(t))_{t \in [0, T]}$ be an $\mathbb{R}^{d \times d}$ -valued \mathbb{F} -adapted process such that

$$E \int_0^T \sum_{i,j=1}^d \left(\int_0^t \text{tr}(\Phi_i(u)^t \Phi_j(u) \cdot d[X](u)) \right) d[X]_{i,j}(t) < \infty.$$

Then $\int_0^\cdot \Phi \cdot dX$ defines an \mathbb{R}^d -valued \mathbb{F} -adapted process whose components are in $\mathcal{L}^2(X)$. Consider now the process

$$Y(t) = \int_0^t \left(\int_0^s \Phi(u) \cdot dX(u) \right) dX(s) \quad (123)$$

Then $Y \in \mathcal{S}^{2,2}(X)$ and

$$\nabla_X Y(t) = \int_0^t \Phi \cdot dX, \quad \nabla_X^2 Y(t) = \Phi(t), \quad \text{and } \mathcal{D}Y(t) = -\frac{1}{2} \langle \Phi(t), A(t) \rangle$$

where $A(t) = d[X]/dt$ and $\langle \Phi, A \rangle = \text{tr}(\Phi \cdot A)$.

Note that, in the example above, the matrix-valued process Φ need not be symmetric:

Example 7.16. In Example 7.15, take $d = 2$, $X = (W^1, W^2)$ is a two-dimensional standard Wiener process with $\text{cov}(W^1(t), W^2(t)) = \rho t$. Let

$$\Phi(t) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad \text{Then}$$

$$Y(t) = \int_0^t \left(\int_0^s \Phi \cdot dX \right) dX(s) = \int_0^t \begin{pmatrix} 0 \\ W^1 \end{pmatrix} \cdot dX = \int_0^t W^1 dW^2$$

verifies $Y \in \mathcal{S}^{2,2}(X)$, with

$$\nabla_X Y(t) = \int_0^t \Phi \cdot dX = \begin{pmatrix} 0 \\ W^1(t) \end{pmatrix}, \quad \nabla_X^2 Y(t) = \Phi = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{D}Y = -\frac{\rho}{2}.$$

In particular, $\nabla_X^2 Y$ is not symmetric. Note that a naive pathwise calculation of the horizontal derivative gives ' $\mathcal{D}Y = 0$ ', which is incorrect: the correct calculation is based on (122).

This example should be contrasted with the situation for pathwise differentiable functionals: as discussed in Section 5.2.3, for $F \in \mathcal{C}_b^{1,2}(\Lambda_T)$, $\nabla_\omega^2 F(t, \omega)$ is a symmetric matrix. This implies that Y has no smooth functional representation: $Y \notin \mathcal{C}_b^{1,2}(X)$.

Example 7.15 is fundamental and can be used to construct in a similar manner martingale elements of $\mathcal{S}^{k,2}(X)$ for any $k \geq 2$ using multiple stochastic integrals of square-integrable tensor-valued processes. The same reasoning as above can then be used to show that for any square-integrable *non-symmetric* k -tensor process, its multiple (k -th order) stochastic integral with respect to X yields an element of $\mathcal{S}^{k,2}(X)$ which does not belong to $\mathcal{C}_b^{1,k}(X)$.

The following example clarifies some incorrect remarks found in the literature:

Example 7.17 (Quadratic variation). The quadratic variation process $Y = [X]$ is a square-integrable functional of X : $Y \in \mathcal{S}^{2,2}(X)$ and, from Definition 7.13 we obtain:

$$\nabla_X Y = 0, \quad \nabla_X^2 Y = 0, \quad \text{and} \quad \mathcal{D}Y(t) = A(t)$$

where $A(t) = d[X]/dt$.

Indeed, since $[X]$ has finite variation: $Y \in \ker(\nabla_X)$. This is consistent with the pathwise approach used in [11] (see Section 6.3) which yields the same result.

These ingredients now allow us to state a weak form of the Functional Ito formula, *without any pathwise differentiability* requirements:

Proposition 7.18 (Functional Ito formula: weak form). *For any semimartingale $S \in \mathcal{S}^{2,2}(X)$, the following equality holds $dt \times d\mathbb{P}$ -a.e. :*

$$S(t) = S(0) + \int_0^t \nabla_X S \cdot dX + \int_0^t \mathcal{D}S(u) du + \frac{1}{2} \int_0^t \text{tr}({}^t \nabla_X^2 S \cdot d[X]) \quad (124)$$

For $S \in \mathcal{C}_{\text{loc}}^{1,2}(X)$ the terms in the formula coincide with the pathwise derivatives in Theorem 6.4. The requirement $S \in \mathcal{S}^{2,2}(X)$ is necessary for defining all terms in (124) as square-integrable processes: although one can compute a weak derivative $\nabla_X S$ for $S \in \mathcal{S}^{1,2}(X)$, if the condition $S \in \mathcal{S}^{2,2}(X)$ is removed, in general the terms in the decomposition of $S - \int \nabla_X S.dX$ are distribution-valued processes.

7.6 Changing the reference martingale

Up to now we have considered a fixed reference process X , assumed in Sections 7.2–7.5 to be a square-integrable martingale satisfying Assumption 7.1. One of the strong points of this construction is precisely that the choice of the reference martingale X is arbitrary; unlike the Malliavin calculus, we are not restricted to choosing X to be Gaussian. We will now show that the operator ∇_X transforms in a simple way under a change of the reference martingale X .

Proposition 7.19 (‘Change of reference martingale’). *Let X be a square-integrable Ito martingale verifying Assumption 7.1. If $M \in \mathcal{M}^2(X)$ and $S \in \mathcal{S}^{1,2}(M)$, then $S \in \mathcal{S}^{1,2}(X)$ and*

$$\nabla_X S(t) = \nabla_M S(t) \cdot \nabla_X M(t) \quad dt \times d\mathbb{P} - a.e. \quad (125)$$

Proof. $M \in \mathcal{M}^2(X)$ so by Theorem 7.8, $M = \int_0^\cdot \nabla_X M.dX$ and

$$[M](t) = \int_0^t \text{tr}(\nabla_X M \cdot {}^t \nabla_X M.d[X])$$

$\mathcal{F}_t^M \subset \mathcal{F}_t^X$ so $\mathcal{A}^2(M) \subset \mathcal{A}^2(\mathbb{F})$. Since $S \in \mathcal{S}^{1,2}(M)$ there exists $H \in \mathcal{A}^2(M) \subset \mathcal{A}^2(\mathbb{F})$ and $\nabla_M S \in \mathcal{L}^2(M)$ such that

$$S = H + \int_0^\cdot \nabla_M S.dM = H + \int_0^\cdot \nabla_M S \nabla_X M.dX, \quad \text{and}$$

$$E \left(\int_0^T |\nabla_M S(t)|^2 \text{tr}(\nabla_X M \cdot {}^t \nabla_X M.d[X]) \right) = E \left(\int_0^T |\nabla_M S|^2 d[M] \right) = \|\nabla_M S\|_{\mathcal{L}^2(M)}^2 < \infty,$$

so $S \in \mathcal{S}^{1,2}(X)$. Uniqueness of the semimartingale decomposition of S then entails $\int_0^\cdot \nabla_M S \nabla_X M.dX = \int_0^\cdot \nabla_X S.dX$. Using Assumption 7.1 and following the same steps as in the proof of Lemma 7.1, we conclude that

$$\nabla_X S(t) = \nabla_M S(t) \cdot \nabla_X M(t) \quad dt \times d\mathbb{P} - a.e. \quad \square$$

7.7 Application to Forward-Backward SDEs

Consider now a forward-backward stochastic differential equation (FBSDE) [57, 58] driven by a continuous semimartingale X :

$$X(t) = X(0) + \int_0^t b(u, X_{u-}, X(u))du + \int_0^t \sigma(u, X_{u-}, X(u))dW(u) \quad (126)$$

$$Y(t) = H(X_T) - \int_t^T f(s, X_{s-}, X(s), Y(s), Z(s))ds - \int_t^T Z.dX, \quad (127)$$

where we have distinguished in our notation the dependence with respect to path X_{t-} on $[0, t[$ from the dependence with respect to its endpoint $X(t)$, i.e. the stopped path (t, X_t) is represented as $(t, X_{t-}, X(t))$. X is called the *forward* process, $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ is the \mathbb{P} -completed natural filtration of X , $H \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and one looks for a pair (Y, Z) of \mathbb{F} -adapted processes satisfying (126)–(127).

This form of (127), with a stochastic integral with respect to the forward process X rather than Brownian motion, is the one which naturally appears in problems in stochastic control and mathematical finance (see e.g. [39] for a discussion). The coefficients

$$b : \mathcal{W}_T \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \sigma : \mathcal{W}_T \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}, \quad f : \mathcal{W}_T \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$$

are assumed to satisfy the following standard assumptions [57]:

Assumption 7.2 (Assumptions on BSDE coefficients).

- (i) For each $(x, y, z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$, $b(\cdot, \cdot, x)$, $\sigma(\cdot, \cdot, x)$ and $f(\cdot, \cdot, x, y, z)$ are non-anticipative functionals.
- (ii) $b(t, \omega, \cdot)$, $\sigma(t, \omega, \cdot)$ and $f(t, \omega, \cdot, \cdot, \cdot)$ are Lipschitz-continuous, uniformly with respect to $(t, \omega) \in \mathcal{W}_T$.
- (iii) $E \int_0^T (|b(t, \cdot, 0)|^2 + |\sigma(t, \cdot, 0)|^2 + |f(t, \cdot, 0, 0, 0)|^2) dt < \infty$.
- (iv) $H \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Under these assumptions, the forward SDE (126) has a unique strong solution X with

$$E(\sup_{t \in [0, T]} |X(t)|^2) < \infty$$

and its local martingale component $M(t) = \int_0^t \sigma(u, X_{u-}, X(u)) dW(u)$ is a square-integrable martingale, with $M \in \mathcal{M}^2(W)$. One can then construct the weak vertical derivative operator $\nabla_M : \mathcal{M}^2(M) \rightarrow \mathcal{L}^2(M)$ as explained above.

Theorem 7.20. *Assume*

$$\det \sigma(t, X_{t-}, X(t)) \neq 0 \quad dt \times d\mathbb{P} - a.e.$$

The FBSDE (126)–(127) admits a unique solution $(Y, Z) \in \mathcal{S}^{1,2}(M) \times \mathcal{L}^2(M)$ such that $E(\sup_{t \in [0, T]} |Y(t)|^2) < \infty$, and

$$Z(t) = \nabla_M Y(t).$$

Proof. Let us rewrite (127) in the classical form

$$\begin{aligned} Y(t) &= H(X_T) - \int_t^T \underbrace{Z(u) \sigma(u, X_{u-}, X(u))}_{U(u)} \cdot dW(u) \\ &- \int_t^T \underbrace{(f(u, X_{u-}, X(u), Y(u), Z(u)) + Z(u) b(u, X_{u-}, X(u)))}_{g(u, X_{u-}, X(u), Y(u), Z(u))} du \end{aligned} \quad (128)$$

The map g then also verifies Assumption 7.2, so the classical result of Pardoux & Peng [57] then implies that there exists a unique pair of \mathbb{F} -adapted processes (Y, U) such that

$$E \sup_{t \in [0, T]} |Y(t)|^2 + E \int_0^T \|U(t)\|^2 dt < \infty$$

which satisfy (126)-(128). Using the decomposition (127) we observe that that Y is an \mathcal{F}_t^M -semimartingale. Since $E \sup_{t \in [0, T]} |Y(t)|^2 < \infty$, $Y \in \mathcal{S}^{1,2}(W)$ with $\nabla_W Y(t) = Z(t)\sigma(t, X_{t-}, X(t))$. The non-singularity assumption on $\sigma(t, X_{t-}, X(t))$, implies that $\mathcal{F}_t^M = \mathcal{F}_t^W$. Let us define

$$Z(t) = U(t)\sigma(t, X_{t-}, X(t))^{-1}.$$

Then Z is \mathcal{F}_t^M -measurable, and

$$\|Z\|_{\mathcal{L}^2(M)}^2 = E \left(\int_0^T Z.dM \right)^2 = E \left(\int_0^T U.dW \right)^2 = E \left(\int_0^T \|U(t)\|^2 .dt \right) < \infty$$

so $Z \in \mathcal{L}^2(M)$. Therefore, Y also belongs to $\mathcal{S}^{1,2}(M)$ and by Proposition (7.19),

$$\nabla_W Y(t) = \nabla_M Y(t).\nabla_W M(t) = \nabla_M Y(t)\sigma(t, X_{t-}, X(t)),$$

$$\text{so } \nabla_M Y(t) = \nabla_W Y(t)\sigma(t, X_{t-}, X(t))^{-1} = U(t)\sigma(t, X_{t-}, X(t))^{-1} = Z(t).$$

□

8 Functional Kolmogorov Equations

Rama CONT & David Antoine FOURNIE

One of the key topics in stochastic analysis is the deep link between Markov processes and partial differential equations, which can be used to characterize a diffusion process in terms of its infinitesimal generator [70]. Consider a second-order differential operator $L : C_b^{1,2}([0, T] \times \mathbb{R}^d) \rightarrow C_b^0([0, T] \times \mathbb{R}^d)$ defined, for test functions $f \in C^{1,2}([0, T] \times \mathbb{R}^d, \mathbb{R})$, by

$$Lf(t, x) = \frac{1}{2} \text{tr}(\sigma(t, x) \cdot {}^t\sigma(t, x) \partial_x^2 f) + b(t, x) \partial_x f(t, x)$$

with continuous, bounded coefficients

$$b \in C_b^0([0, T] \times \mathbb{R}^d, \mathbb{R}^d), \quad \sigma \in C_b^0([0, T] \times \mathbb{R}^d, \mathbb{R}^{d \times n}).$$

Under various sets of assumptions on b, σ , such as

- (i) Lipschitz continuity and linear growth or
 - (ii) uniform ellipticity : $\sigma^t \sigma(t, x) \geq \epsilon I_d$,
- the stochastic differential equation

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t)$$

has a unique weak solution $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, W, \mathbb{P})$ and X is a Markov process under \mathbb{P} whose evolution operator $(P_{t,s}^X, s \geq t \geq 0)$ has infinitesimal generator L .

This weak solution $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, W, \mathbb{P})$ is characterized by the property that

$$\forall f \in C_b^{1,2}([0, T] \times \mathbb{R}^d), \quad f(u, X(u)) - \int_0^u (\partial_t f + Lf)(t, X(t))dt$$

is a \mathbb{P} -martingale [70]. Indeed, applying Ito's formula yields

$$\begin{aligned} f(u, X(u)) &= f(0, X(0)) + \int_0^u (\partial_t f + Lf)(t, X(t))dt \\ &\quad + \int_0^u \partial_x f(t, X(t)) \sigma(t, X(t)) dW \end{aligned}$$

In particular, if

$$(\partial_t f + Lf)(t, x) = 0 \quad \forall t \in [0, T], \quad \forall x \in \text{supp}(X(t))$$

then $M(t) = f(t, X(t))$ is a martingale.

More generally for $f \in C^{1,2}([0, T] \times \mathbb{R}^d)$, $M(t) = f(t, X(t))$ is a local martingale. if and only if f is a solution of

$$\partial_t f(t, x) + Lf(t, x) = 0 \quad \forall t \geq 0, x \in \text{supp}(X(t)).$$

This PDE is the Kolmogorov (backward) equation associated with X , whose solutions are the *space-time (L-)harmonic functions* associated with the differential operator L :

$$f \text{ space-time harmonic} \iff f(t, X(t)) \text{ is a (local) martingale.}$$

The tools of Functional Ito calculus may be used to extend these relations beyond the Markovian setting, to a large class of processes with path-dependent characteristics. This leads us to a new class of partial differential equations on path space –*functional Kolmogorov equations*– which have many properties in common with their finite dimensional counterparts and lead to new Feynman-Kac formulas for path-dependent functionals of semimartingales. This is an exciting new topic, with many connections to other strands of stochastic analysis, and which is just starting to be explored [10, 14, 23, 28, 59].

8.1 Functional Kolmogorov equations and harmonic functionals

8.1.1 Stochastic differential equations with path-dependent coefficients

Let $(\Omega = D([0, T], \mathbb{R}^d), \mathcal{F}^0)$ be the canonical space and consider a semimartingale X which can be represented as the solution of a stochastic differential equation whose coefficients are allowed to be path-dependent, left-continuous functionals:

$$dX(t) = b(t, X_t)dt + \sigma(t, X_t)dW(t) \tag{129}$$

where b, σ are non-anticipative functionals with values in \mathbb{R}^d (resp. $\mathbb{R}^{d \times n}$) whose coordinates are in $\mathbb{C}_t^{0,0}(\Lambda_T)$ such that equation (129) has a unique weak solution \mathbb{P} on (Ω, \mathcal{F}^0) .

This class of processes is a natural ‘path-dependent’ extension of the class of diffusion processes; various conditions (such as functional Lipschitz property and boundedness) may be given for existence and uniqueness of solutions (see e.g. [62]). We give here a sample of such a result, which will be useful in the examples.

Proposition 8.1 (Strong solutions for path-dependent SDEs). *Assume that the non-anticipative functionals b and σ satisfy the following Lipschitz and linear growth conditions:*

$$|b(t, \omega) - b(t, \omega')| + |\sigma(t, \omega) - \sigma(t, \omega')| \leq K \sup_{s \leq t} |\omega(s) - \omega'(s)| \quad (130)$$

$$|b(t, \omega)| + |\sigma(t, \omega)| \leq K(1 + \sup_{s \leq t} |\omega(s)| + |t|) \quad (131)$$

for all $t \geq t_0$, $\omega, \omega' \in C^0([0, t], \mathbb{R}^d)$. Then for any $\xi \in C^0([0, T], \mathbb{R}^d)$, the stochastic differential equation 129 has a unique strong solution X with initial condition $X_{t_0} = \xi_{t_0}$. The paths of X lie in $C^0([0, T], \mathbb{R}^d)$ and

1. There exists a constant C depending only on T, K such that, for $t \in [t_0, T]$:

$$E[\sup_{s \in [0, t]} |X(s)|^2] \leq C(1 + \sup_{s \in [0, t_0]} |\xi(s)|^2) e^{C(t-t_0)} \quad (132)$$

2. $\int_0^{t-t_0} [|b(t_0 + s, X_{t_0+s})| + |\sigma(t_0 + s, X_{t_0+s})|^2] ds < +\infty$ a.s.

3. $X(t) - X(t_0) = \int_0^{t-t_0} b(t_0 + s, X_{t_0+s}) ds + \sigma(t_0 + s, X_{t_0+s}) dW(s)$

We denote $\mathbb{P}_{(t_0, \xi)}$ the law of this solution.

Proofs of the uniqueness and existence, as well as the useful bound (132), can be found in [62].

Topological support of a stochastic process In the statement of the link between a diffusion and the associated Kolmogorov equation, the domain of the PDE is related to the support of the random variable $X(t)$. In the same way, the support of the law (on space of paths) of the process plays a key role in the path-dependent extension of the Kolmogorov equation.

Recall that the *topological support* of a random variable X with values in a metric space is the smallest closed set $\text{supp}(X)$ such that $\mathbb{P}(X \in \text{supp}(X)) = 1$.

Viewing a process X with continuous paths as a random variable on $(C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty)$ leads to the notion of *topological support* of a continuous process.

The support may be characterized by the following property: it is the set $\text{supp}(X)$ of paths $\omega \in C^0([0, T], \mathbb{R}^d)$ for which every Borel neighborhood has strictly positive measure:

$$\text{supp}(X) = \{\omega \in C^0([0, T], \mathbb{R}^d) \mid \text{for any Borel neighborhood } V \text{ of } \omega, \mathbb{P}(X_T \in V) > 0\}$$

Example 8.2. If W is a d -dimensional Wiener process with non-singular covariance matrix, then

$$\text{supp}(W) = \{\omega \in C^0([0, T], \mathbb{R}^d), \quad \omega(0) = 0\}.$$

Example 8.3. For an Ito process $X(t, \omega) = x + \int_0^t \sigma(t, \omega).dW$, if $\mathbb{P}(\sigma(t, \omega).^t\sigma(t, \omega) \geq \epsilon \quad Id) = 1$ then

$$\text{supp}(X) = \{\omega \in C^0([0, T], \mathbb{R}^d), \quad \omega(0) = x\}.$$

A classical result due to Stroock and Varadhan [69] characterizes the topological support of a multidimensional diffusion process (for a simple proof see Millet & Sanz-Solé [54]):

Example 8.4 (Stroock-Varadhan support theorem). Let

$$b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$$

be Lipschitz-continuous in x , uniformly on $[0, T]$. Then the law of the diffusion process

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t) \quad X(0) = x$$

has a support given by the closure in $(C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty)$ of the 'skeleton'

$$x + \{\omega \in H^1([0, T], \mathbb{R}^d), \exists h \in H^1([0, T], \mathbb{R}^d), \dot{\omega}(t) = b(t, \omega(t)) + \sigma(t, \omega(t))\dot{h}(t)\}$$

defined as the set of all solutions of the system of ODEs obtained when substituting $h \in H^1([0, T], \mathbb{R}^d)$ for W in the SDE.

We will denote by $\Lambda(T, X)$ the set of all stopped paths obtained from paths in $\text{supp}(X)$:

$$\Lambda(T, X) = \{(t, \omega) \in \Lambda_T, \omega \in \text{supp}(X)\} \quad (133)$$

For a continuous process X , $\Lambda(T, X) \subset \mathcal{W}_T$.

8.1.2 Local martingales and harmonic functionals

Functionals of a process X which have the local martingale property play an important role in stochastic control theory [18] and mathematical finance. By analogy with the Markovian case, we call such functionals \mathbb{P} -harmonic functionals:

Definition 8.5 (\mathbb{P} -harmonic functionals). A non-anticipative functional F is called \mathbb{P} -harmonic if $Y(t) = F(t, X_t)$ is a \mathbb{P} -local martingale.

The following result [10] characterizes smooth \mathbb{P} -harmonic functionals as solutions to a *functional Kolmogorov equation*:

Theorem 8.6 (Functional equation for $\mathcal{C}^{1,2}$ martingales). If $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$ and $DF \in \mathbb{C}_l^{0,0}$, then $Y(t) = F(t, X_t)$ is a local martingale if and only if F satisfies

$$DF(t, \omega_t) + b(t, \omega_t)\nabla_\omega F(t, \omega_t) + \frac{1}{2}\text{tr}[\nabla_\omega^2 F(t, \omega)\sigma^t\sigma(t, \omega)] = 0, \quad (134)$$

on the topological support of X in $(C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty)$.

Proof. If $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$ is a solution of (134), the functional Ito formula (Theorem 6.4) applied to $Y(t) = F(t, X_t)$ shows that Y has the semimartingale decomposition

$$Y(t) = Y(0) + \int_0^t Z(u)du + \int_0^t \nabla_\omega F(u, X_u)\sigma(u, X_u)dW(u) \quad \text{where}$$

$$Z(t) = \mathcal{D}F(t, X_t) + b(t, X_t)\nabla_\omega F(t, X_t) + \frac{1}{2}\text{tr}(\nabla_\omega^2 F(t, X_t)\sigma^t\sigma(t, X_t)).$$

(134) then implies that the finite variation term in (72) is almost-surely zero: so $Y(t) = Y(0) + \int_0^t \nabla_\omega F(u, X_u)\cdot\sigma(u, X_u)dW(u)$ and thus Y is a continuous local martingale.

Conversely, assume that $Y = F(X)$ is a local martingale. Let $A(t, \omega) = \sigma(t, \omega)^t\sigma(t, \omega)$. Then Y is left-continuous by Lemma 5.5. Suppose (134) is not satisfied for some $\omega \in \text{supp}(X) \subset C^0([0, T], \mathbb{R}^d)$ and some $t_0 < T$. Then there exists $\eta > 0$ and $\epsilon > 0$ such that

$$|\mathcal{D}F(t, \omega) + b(t, \omega)\nabla_\omega F(t, \omega) + \frac{1}{2}\text{tr}[\nabla_\omega^2 F(t, \omega)A(t, \omega)]| > \epsilon \quad (135)$$

for $t \in [t_0 - \eta, t_0]$, by left-continuity of the expression. By left-continuity of the expression, there exist a neighborhood V of ω in $C^0([0, T], \mathbb{R}^d)$, $\|\cdot\|_\infty$ such that, for all $\omega' \in V$ and $t \in [t_0 - \eta, t_0]$,

$$|\mathcal{D}F(t, \omega') + b(t, \omega')\nabla_\omega F(t, \omega') + \frac{1}{2}\text{tr}[\nabla_\omega^2 F(t, \omega')A(t, \omega')]| > \frac{\epsilon}{2} \quad (136)$$

Since $\omega \in \text{supp}(X)$, $\mathbb{P}(X \in V) > 0$ so

$$\{(t, \omega) \in \mathcal{W}_T, \quad |\mathcal{D}F(t, \omega) + b(t, \omega)\nabla_\omega F(t, \omega) + \frac{1}{2}\text{tr}(\nabla_\omega^2 F(t, \omega)A(t, \omega))| > \frac{\epsilon}{2}\}$$

has non-zero measure with respect to $dt \times d\mathbb{P}$. Applying the functional Ito formula (72) to the process $Y(t) = F(t, X_t)$ then leads to a contradiction, because as a continuous local martingale its finite variation component should be zero $dt \times d\mathbb{P}$ -almost everywhere. \square

In the case where $F(t, \omega) = f(t, \omega(t))$ and the coefficients b, σ are not path-dependent, this equation reduces to the well-known backward Kolmogorov equation [46].

Remark 8.7 (Relation with infinite-dimensional Kolmogorov equations in Hilbert spaces). Note that the second operator ∇_ω^2 is an iterated directional derivative, and not a second-order Fréchet or "H-derivative", so this equation is different from the infinite-dimensional Kolmogorov equations in Hilbert spaces as described, for example, by Da Prato & Zabczyk [16]. The relation between these two classes of infinite-dimensional Kolmogorov equations has been studied recently by Flandoli & Zanco [28] who show that, although one can partly recover some results for (134) using the theory of infinite-dimensional Kolmogorov equations in Hilbert-spaces, this can only be done at the price of higher regularity requirements (not the least of them being Fréchet differentiability), which exclude many interesting examples.

When $X = W$ is a d -dimensional Wiener process, Theorem 8.6 gives a characterization of regular Brownian local martingales as solutions of a 'functional heat equation':

Corollary 8.8 (Harmonic functionals on \mathcal{W}_T). *Let $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$. Then $(Y(t) = F(t, W_t), t \in [0, T])$ is a local martingale if and only if*

$$\forall t \in [0, T], \forall \omega \in C^0([0, T], \mathbb{R}^d), \quad \mathcal{D}F(t, \omega) + \frac{1}{2} \operatorname{tr}(\nabla_\omega^2 F(t, \omega)) = 0.$$

8.1.3 Sub-solutions and super-solutions

By analogy with the case of finite-dimensional parabolic PDEs, we introduce the notion of sub- and super-solution to (134).

Definition 8.9 (Subsolutions and supersolutions). $F \in \mathbb{C}^{1,2}(\Lambda_T)$ is called a supersolution of (134) on a domain $U \subset \Lambda_T$ if: $\forall (t, \omega) \in U$,

$$\mathcal{D}F(t, \omega) + b(t, \omega) \cdot \nabla_\omega F(t, \omega) + \frac{1}{2} \operatorname{tr}[\nabla_\omega^2 F(t, \omega) \sigma^t \sigma(t, \omega)] \leq 0 \quad (137)$$

$F \in \mathbb{C}_{\text{loc}}^{1,2}(\Lambda_T)$ is called a subsolution of (134) on U if: $\forall (t, \omega) \in U$,

$$\mathcal{D}F(t, \omega) + b(t, \omega) \cdot \nabla_\omega F(t, \omega) + \frac{1}{2} \operatorname{tr}[\nabla_\omega^2 F(t, \omega) \sigma^t \sigma(t, \omega)] \geq 0 \quad (138)$$

The following result shows that \mathbb{P} -integrable sub- and super-solutions have a natural connection with \mathbb{P} -submartingales and \mathbb{P} -supermartingales:

Proposition 8.10. *Let $F \in \mathbb{C}_b^{1,2}(\Lambda_T)$ such that: $\forall t \in [0, T], E(|F(t, X_t)|) < \infty$. Set $Y(t) = F(t, X_t)$.*

- Y is a submartingale if and only if F is a subsolution of (134) on $\operatorname{supp}(X)$.
- Y is a supermartingale if and only if F is a supersolution of (134) on $\operatorname{supp}(X)$.

Proof. If $F \in \mathbb{C}^{1,2}$ is a supersolution of (134), the application of the Functional Ito formula (Theorem 6.4, Eq. (72)) gives

$$\begin{aligned} Y(t) - Y(0) &= \int_0^t \nabla_\omega F(u, X_u) \sigma(u, X_u) \cdot dW(u) \\ &+ \int_0^t \left(\mathcal{D}F(u, X_u) + \nabla_\omega F(u, X_u) \cdot b(u, X_u) + \frac{1}{2} \operatorname{tr}(\nabla_\omega^2 F(u, X_u) A(u, X_u)) \right) du \quad a.s. \end{aligned}$$

where the first term is a local martingale and, since F is a supersolution, the finite variation term (second line) is a decreasing process, so Y is a supermartingale.

Conversely, let $F \in \mathbb{C}^{1,2}(\Lambda_T)$ such that $Y(t) = F(t, X_t)$ is a supermartingale. Let $Y = M - H$ be its Doob-Meyer decomposition, where M is a local martingale

and H an increasing process. From the functional Ito formula, Y has the above decomposition so it is a continuous supermartingale. By uniqueness of the Doob-Meyer decomposition, the second term is thus a decreasing process so

$$\mathcal{D}F(t, X_t) + \nabla_\omega F(t, X_t) \cdot b(t, X_t) + \frac{1}{2} \text{tr}(\nabla_\omega^2 F(t, X_t) A(t, X_t)) \leq 0 \, dt \times d\mathbb{P} - a.e.$$

This implies that the set

$$S = \{\omega \in C^0([0, T], \mathbb{R}^d), (137) \text{ holds for all } t \in [0, T], \}$$

includes any Borel set $B \in \mathcal{B}(C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty)$ with $\mathbb{P}(X_T \in B) > 0$. Using continuity of paths,

$$S = \cap_{t \in [0, T] \cap \mathbb{Q}} \{\omega \in C^0([0, T], \mathbb{R}^d), (137) \text{ holds for } (t, \omega)\}$$

Since $F \in \mathbb{C}^{1,2}(\Lambda_T)$, this is a closed set in $(C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty)$, so S contains the topological support of X , i.e. (137) holds on $\Lambda_T(X)$. \square

8.1.4 Comparison principle and uniqueness

A key property of the functional Kolmogorov equation (134) is the comparison principle. As in the finite dimensional case, this requires imposing an integrability condition:

Theorem 8.11 (Comparison principle). *Let $\underline{F} \in \mathbb{C}^{1,2}(\Lambda_T)$ be a subsolution of (134) and $\overline{F} \in \mathbb{C}^{1,2}(\Lambda_T)$ be a supersolution of (134) such that*

$$\forall \omega \in C^0([0, T], \mathbb{R}^d), \quad \underline{F}(T, \omega) \leq \overline{F}(T, \omega)$$

$$\text{and } E \left(\sup_{t \in [0, T]} |\underline{F}(t, X_t) - \overline{F}(t, X_t)| \right) < \infty. \quad (139)$$

$$\text{Then : } \quad \forall t \in [0, T], \forall \omega \in \text{supp}(X), \quad \underline{F}(t, \omega) \leq \overline{F}(t, \omega). \quad (140)$$

Proof. $F = \underline{F} - \overline{F} \in \mathbb{C}^{1,2}(\Lambda_T)$ is a subsolution of (134) so by Proposition 8.10, the process Y defined by $Y(t) = F(t, X_t)$ is a submartingale. Since $Y(T) = \underline{F}(T, X_T) - \overline{F}(T, X_T) \leq 0$, we have

$$\forall t \in [0, T[, \quad Y(t) \leq 0 \quad \mathbb{P} - a.s.$$

Define

$$S = \{\omega \in C^0([0, T], \mathbb{R}^d), \forall t \in [0, T], \underline{F}(t, \omega) \leq \overline{F}(t, \omega)\}.$$

Using continuity of paths, and since $F \in \mathbb{C}^{0,0}(\mathcal{W}_T)$,

$$S = \cap_{t \in \mathbb{Q} \cap [0, T]} \{\omega \in \text{supp}(X), \underline{F}(t, \omega) \leq \overline{F}(t, \omega)\}$$

is closed in $(C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty)$. If (140) does not hold then there exists $t_0 \in [0, T[, \omega \in \text{supp}(X)$ such that $F(t_0, \omega) > 0$. Since $O = \text{supp}(X) - S$ is a non-empty open subset of $\text{supp}(X)$, there exists an open set $A \subset O \subset \text{supp}(X)$ containing ω and $h > 0$ such that

$$\forall t \in [t_0 - h, t], \forall \omega \in A, \quad F(t, \omega) > 0.$$

But since $\mathbb{P}(X_T \in A) > 0$ this implies $\int dt \times d\mathbb{P} \quad 1_{F(t, \omega) > 0} > 0$ which contradicts the above. \square

A straightforward consequence of the comparison principle is the following uniqueness result for \mathbb{P} -uniformly integrable solutions of the Functional Kolmogorov equation (134) :

Theorem 8.12 (Uniqueness of solutions). *Let $H : (C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty) \rightarrow \mathbb{R}$ be a continuous functional. If $F^1, F^2 \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$ are solutions of the Functional Kolmogorov equations (134) satisfying*

$$F^1(T, \omega) = F^2(T, \omega) = H(\omega) \tag{141}$$

$$E\left[\sup_{t \in [0, T]} |F^1(t, X_t) - F^2(t, X_t)|\right] < \infty \tag{142}$$

then they coincide on the topological support of X :

$$\forall \omega \in \text{supp}(X), \quad \forall t \in [0, T] \quad F^1(t, \omega) = F^2(t, \omega). \tag{143}$$

The integrability condition in this result (or some variation of it) is required for uniqueness: indeed, even in the finite-dimensional setting of a heat equation in \mathbb{R}^d , one cannot do without an integrability condition with respect to the heat kernel. Without this condition, we can only assert that $F(t, X_t)$ is a local martingale, so it is not uniquely determined by its terminal value and counterexamples to uniqueness are easy to construct.

Note also that uniqueness holds on the support of X , the process associated with the (path-dependent) operator appearing in the Kolmogorov equation. The local martingale property of $F(X)$ implies no restriction on the behavior of F outside $\text{supp}(X)$ so there is no reason to expect uniqueness or comparison of solutions on the whole path space.

8.1.5 Feynman-Kac formula for path-dependent functionals

As observed above, any \mathbb{P} -uniformly integrable solution F of the functional Kolmogorov equation yields a \mathbb{P} -martingale $Y(t) = F(t, X_t)$. Combined with the uniqueness result (Theorem 8.12), this leads to an extension of the well-known Feynman-Kac representation to the path-dependent case:

Theorem 8.13 (A Feynman-Kac formula for path-dependent functionals). *Let $H : (C^0([0, T], \|\cdot\|_\infty) \mapsto \mathbb{R}$ continuous. If $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$ verifies*

$$\begin{aligned} \forall \omega \in C^0([0, T]), \quad \mathcal{D}F(t, \omega) + b(t, \omega) \cdot \nabla_\omega F(t, \omega) + \frac{1}{2} \text{tr}[\nabla_\omega^2 F(t, \omega) \sigma^t \sigma(t, \omega)] &= 0 \\ F(T, \omega) = H(\omega), \quad E\left[\sup_{t \in [0, T]} |F(t, X_t)|\right] &< \infty \end{aligned}$$

then F has the probabilistic representation

$$\forall \omega \in \text{supp}(X), \quad F(t, \omega) = E^{\mathbb{P}}[H(X_T)|X_t = \omega_t] = E^{\mathbb{P}^{(t, \omega)}}[H(X_T)]$$

where $\mathbb{P}^{(t, \omega)}$ is the law of the unique solution of the SDE

$$dX(u) = b(u, X_u)du + \sigma(u, X_u)dW(u) \quad u \geq t$$

with initial condition $X_t = \omega_t$. In particular:

$$F(t, X_t) = E^{\mathbb{P}}[H(X_T)|\mathcal{F}_t] \quad dt \times d\mathbb{P} - a.s.$$

8.2 FBSDEs and semilinear functional PDEs

The relation between stochastic processes and partial differential equations extends to the semilinear case: in the Markovian setting, there is a well-known relation between semilinear PDEs of parabolic type and Markovian forward-backward stochastic differential equations [58, 24, 75]. The tools of Functional Ito calculus allows to extend the relation between semilinear PDEs and FBSDEs beyond the Markovian case, to non-Markovian FBSDEs with path-dependent coefficients.

Let

$$b : \mathcal{W}_T \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \sigma : \mathcal{W}_T \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}, \quad f : \mathcal{W}_T \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$$

be non-anticipative functionals satisfying Assumption 7.2. Then the path-dependent SDE

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

has a unique strong solution X with $E(\sup_{t \in [0, T]} |X(t)|^2) < \infty$, whose martingale part we denote by M .

Consider the forward-backward stochastic differential equation (FBSDE) with path-dependent coefficients:

$$X(t) = X(0) + \int_0^t b(u, X_{u-}, X(u))du + \int_0^t \sigma(u, X_{u-}, X(u))dW(u) \quad (144)$$

$$Y(t) = H(X_T) + \int_t^T f(s, X_{s-}, X(s), Y(s), Z(s))ds - \int_t^T Z.dM \quad (145)$$

whose coefficients

$$b : \mathcal{W}_T \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \sigma : \mathcal{W}_T \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}, \quad f : \mathcal{W}_T \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$$

are assumed to satisfy Assumption 7.2.

Under Assumption 7.2, the FBSDE (144)-(145) admits a unique \mathbb{F} -adapted solution

$$(Y, Z) \in \mathcal{S}^{1,2}(M) \times \mathcal{L}^2(M) \quad \text{with} \quad E\left(\sup_{t \in [0, T]} |Y(t)|^2\right) < \infty.$$

Define

$$B(t, \omega) = b(t, \omega_{t-}, \omega(t)), \quad A(t, \omega) = \sigma(t, \omega_{t-}, \omega(t))^t \sigma(t, \omega_{t-}, \omega(t)).$$

In the 'Markovian' case where the coefficients are not path-dependent, the solution of the FBSDE can be obtained by solving a semilinear PDE: if $v \in C^{1,2}([0, T] \times \mathbb{R}^d)$ is a solution of

$$\partial_t v(t, x) + f(t, x, v(t, x), \nabla v(t, x)) + b(t, x) \cdot \nabla v(t, x) + \frac{1}{2} \text{tr}[a(t, \omega) \cdot \nabla^2 v(t, x)] = 0$$

then a solution of the FBSDE is given by $(Y(t), Z(t)) = (v(t, X(t)), \nabla v(t, X(t)))$. Here the function v 'decouples' the solution of the backward equation from the solution of the forward equation. In the general, path-dependent case (144)-(145) a similar object, named the 'decoupling random field' by Ma et al. [51], following earlier ideas from [50].

The following result shows that such a 'decoupling random field' for the path-dependent FBSDE (144)-(145) may be constructed as a solution of a semilinear functional PDE, analogously to the Markovian case:

Theorem 8.14 (FBSDEs as semilinear path-dependent PDEs). *Let $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$ be a solution of the path-dependent semilinear equation:*

$$\mathcal{D}F(t, \omega) + f(t, \omega_{t-}, \omega(t), F(t, \omega), \nabla_\omega F(t, \omega)) + B(t, \omega) \cdot \nabla_\omega F(t, \omega) + \frac{1}{2} \text{tr}[A(t, \omega) \cdot \nabla_\omega^2 F(t, \omega)] = 0$$

for $\omega \in \text{supp}(X)$, $t \in [0, T]$ with $F(T, \omega) = H(\omega)$. Then the pair of processes (Y, Z) given by

$$(Y(t), Z(t)) = (F(t, X_t), \nabla_\omega F(t, X_t))$$

is a solution of the FBSDE (144)-(145).

Remark 8.15. In particular, under the conditions of Theorem 8.14 the random field $u : [0, T] \times \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}$ defined by

$$u(t, x, \omega) = F(t, X_{t-}(t, \omega) + x1_{[t, T]})$$

is a 'decoupling random field' for the path-dependent FBSDE (144)-(145) in the sense of Ma et al. [51].

Proof. Let $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T)$ be a solution of the semilinear equation above. Then the processes Y, Z defined by $(Y(t), Z(t)) = (F(t, X_t), \nabla_\omega F(t, X_t))$ are \mathbb{F} -adapted. Applying the functional Ito formula to $F(t, X_t)$ yields

$$\begin{aligned} H(X_T) - F(t, X_t) &= \int_t^T \nabla_\omega F(u, X_u) \cdot dM \\ &+ \int_t^T \left(\mathcal{D}F(u, X_u) + B(u, X_u) \cdot \nabla_\omega F(u, X_u) + \frac{1}{2} \text{tr}[A(u, X_u) \cdot \nabla_\omega^2 F(u, X_u)] \right) du \end{aligned}$$

Since F is a solution of the equation, the term in the last integral is equal to $-f(t, X_{t-}, X(t), F(t, X_t), \nabla_\omega F(t, X_t)) = -f(t, X_{t-}, X(t), Y(t), Z(t))$. So $(Y(t), Z(t)) = (F(t, X_t), \nabla_\omega F(t, X_t))$ verifies

$$Y(t) = H(X_T) + \int_t^T f(s, X_{s-}, X(s), Y(s), Z(s))ds - \int_t^T Z.dM.$$

So (Y, Z) is a solution of the FBSDE (144)-(145). \square

This result provides the "Hamiltonian" version of the FBSDE (144)-(145), in the terminology of [75].

8.3 Non-Markovian stochastic control and path-dependent HJB equations

An interesting application of the Functional Ito calculus is the study of non-Markovian stochastic control problems, where both the evolution of the state of the system and the control policy are allowed to be path-dependent.

Non-Markovian stochastic control problems were studied by Peng [61], who showed that, when the characteristics of the controlled process are allowed to be stochastic (i.e. path-dependent), the optimality conditions can be expressed in terms of a stochastic Hamilton-Jacobi-Bellman equation. The relation with FBSDEs is explored systematically in the monograph [75].

We formulate a mathematical framework for stochastic optimal control where the evolution of the state variable, the control policy and the objective are allowed to be non-anticipative path-dependent functionals, and show how the Functional Itô formula 6.4 characterizes the value functional as the solution to a *functional Hamilton-Jacobi-Bellman equation*.

This section is based on [32].

Let W be a d -dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the \mathbb{P} -augmented natural filtration of W .

Let C be a subset of \mathbb{R}^m . We consider the controlled stochastic differential equation

$$dX(t) = b(t, X_t, \alpha(t))dt + \sigma(t, X_t, \alpha(t))dW(t) \quad (146)$$

where the coefficients

$$b : \mathcal{W}_T \times C \rightarrow \mathbb{R}^d, \quad \sigma : \mathcal{W}_T \times C \rightarrow \mathbb{R}^{d \times d}$$

are allowed to be path-dependent and assumed to satisfy the conditions of Proposition 8.1 uniformly with respect to $\alpha \in C$, and the control α belongs to the set $\mathcal{A}(\mathbb{F})$ of \mathbb{F} -progressively measurable processes with values in C satisfying

$$E \left(\int_0^T \|\alpha(t)\|^2 dt \right) < \infty.$$

Then for any initial condition $(t, \omega) \in \mathcal{W}_T$ and for each $\alpha \in \mathcal{A}(\mathbb{F})$ the stochastic differential equation (146) admits a unique strong solution, which we denote by $X^{(t, \omega, \alpha)}$.

Let $g : (C^0([0, T], \mathbb{R}^d), \|\cdot\|_\infty) \rightarrow \mathbb{R}$ be a continuous map representing a terminal cost and

$$L : \mathcal{W}_T \times C \rightarrow L(t, \omega, u)$$

a non-anticipative functional representing a possibly path-dependent running cost. We assume

$$1. \exists K > 0, \quad -K \leq g(\omega) \leq K(1 + \sup_{s \in [0, t]} |\omega(s)|^2).$$

$$2. \exists K' > 0, \quad -K' \leq L(t, \omega, u) \leq K'(1 + \sup_{s \in [0, t]} |\omega(s)|^2).$$

Then, thanks to (132), the cost functional

$$J(t, \omega, \alpha) = E \left[g(X_T^{(t, \omega, \alpha)}) + \int_t^T L(s, X_s^{(t, \omega, \alpha)}, \alpha(s)) ds \right] < \infty \quad (147)$$

for $(t, \omega, \alpha) \in \mathcal{W}_T \times \mathcal{A}$ and defines a non-anticipative map $J : \mathcal{W}_T \times \mathcal{A}(\mathbb{F}) \rightarrow \mathbb{R}$.

We consider the optimal control problem

$$\inf_{\alpha \in \mathcal{A}(\mathbb{F})} J(t, \omega, \alpha) \quad (148)$$

whose value functional is denoted, for $(t, \omega) \in \mathcal{W}_T$,

$$V(t, \omega) = \inf_{\alpha \in \mathcal{A}(\mathbb{F})} J(t, \omega, \alpha). \quad (149)$$

From the above assumptions and the estimate (132), $V(t, \omega)$ verifies

$$\exists K'' > 0, \quad -K'' \leq V(t, \omega) \leq K''(1 + \sup_{s \in [0, t]} |\omega(s)|^2).$$

Introduce now the *Hamiltonian* associated to the control problem [75] as the non-anticipative functional $H : \mathcal{W}_T \times \mathbb{R}^d \times S_d^+ \rightarrow \mathbb{R}$ given by

$$H(t, \omega, \rho, A) = \inf_{u \in C} \left\{ \frac{1}{2} \text{tr} (A \cdot \sigma(t, \omega, u)^t \sigma(t, \omega, u)) + b(t, \omega, u) \cdot \rho + L(t, \omega, u) \right\} \quad (150)$$

It is readily observed that for any $\alpha \in \mathcal{A}(\mathbb{F})$ the process

$$V(t, X_t^{(t_0, \omega, \alpha)}) + \int_{t_0}^t L(s, X_s^{(t_0, \omega, \alpha)}, \alpha(s)) ds$$

has the submartingale property. The martingale optimality principle [18] then characterizes an optimal control α^* as one for which this process is a martingale.

We can now use the functional Itô formula (Theorem 6.4) to give a sufficient condition for a functional U to be equal to the value functional V and for a control α^* to be optimal. This condition takes the form of a path-dependent Hamilton-Jacobi-Bellman equation:

Theorem 8.16 (Verification Theorem). *Let $U \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$ be a solution of the functional HJB equation:*

$$\forall (t, \omega) \in \mathcal{W}_T, \quad \mathcal{D}U(t, \omega) + H(t, \omega, \nabla_\omega U(t, \omega), \nabla_\omega^2 U(t, \omega)) = 0 \quad (151)$$

satisfying $U(T, \omega) = g(\omega)$ and the quadratic growth condition:

$$|U(t, \omega)| \leq C \sup_{s \leq t} |\omega(s)|^2 \quad (152)$$

Then for any $\omega \in C^0([0, T], \mathbb{R}^d)$ and any admissible control $\alpha \in \mathcal{A}(\mathbb{F})$,

$$U(t, \omega) \leq J(t, \omega, \alpha). \quad (153)$$

If furthermore for $\omega \in C^0([0, T], \mathbb{R}^d)$ there exists $\alpha^ \in \mathcal{A}(\mathbb{F})$ such that*

$$\begin{aligned} H(s, X_s^{(t, \omega, \alpha^*)}, \nabla_\omega U(s, X_s^{(t, \omega, \alpha^*)}), \nabla_\omega^2 U(s, X_s^{(t, \omega, \alpha^*)})) = & (154) \\ \frac{1}{2} \text{tr}[\nabla_\omega^2 U(s, X_s^{(t, \omega, \alpha^*)}) \sigma(s, X_s^{(t, \omega, \alpha^*)}, \alpha^*(s))^t \sigma(s, X_s^{(t, \omega, \alpha^*)}, \alpha^*(s))] & \\ + \nabla_\omega U(s, X_s^{(t, \omega, \alpha^*)}) b(s, X_s^{(t, \omega, \alpha^*)}, \alpha^*(s)) + L(s, X_s^{(t, \omega, \alpha^*)}, \alpha^*(s)) & \end{aligned}$$

ds \times d \mathbb{P} almost-everywhere for $s \in [t, T]$, then

$$U(t, \omega) = V(t, \omega) = J(t, \omega, \alpha^*)$$

and α^ is an optimal control.*

Proof. Let $\alpha \in \mathcal{A}(\mathbb{F})$ be an admissible control, $t < T$ and $\omega \in C^0([0, T], \mathbb{R}^d)$. Applying functional Itô formula yields, for $s \in [t, T]$

$$\begin{aligned} U(s, X_s^{(t, \omega, \alpha)}) - U(t, \omega) &= \int_t^s \nabla_\omega U(u, X_u^{(t, \omega, \alpha)}) \sigma(u, X_u^{(t, \omega, \alpha)}, \alpha(u)) dW(u) \\ &+ \int_t^s \mathcal{D}U(u, X_u^{(t, \omega, \alpha)}) + \nabla_\omega U(u, X_u^{(t, \omega, \alpha)}) b(u, X_u^{(t, \omega, \alpha)}, \alpha(u)) du \\ &+ \int_t^s \frac{1}{2} \text{tr} \left(\nabla_\omega^2 U(u, X_u^{(t, \omega, \alpha)}) \sigma^t \sigma(u, X_u^{(t, \omega, \alpha)}) \right) du \end{aligned}$$

Since U verifies the functional Hamilton-Jacobi-Bellman equation, we have

$$\begin{aligned} U(s, X_s^{(t, \omega, \alpha)}) - U(t, \omega) &\geq \int_t^s \nabla_\omega U(u, X_u^{(t, \omega, \alpha)}) \sigma(u, X_u^{(t, \omega, \alpha)}, \alpha(u)) dW(u) \\ &- \int_t^s L(u, X_u^{(t, \omega, \alpha)}, \alpha(u)) du \end{aligned} \quad (155)$$

In other words, $U(s, X_s^{(t, \omega, \alpha)}) - U(t, \omega) + \int_t^s L(u, X_u^{(t, \omega, \alpha)}, \alpha(u)) du$ is a local submartingale. The estimate (152) and the L^2 estimate (132) guarantees that it is actually a submartingale, hence, taking $s \rightarrow T$, the left-continuity of U yields:

$$E \left[g(X_T^{(t, \omega, \alpha)}) + \int_0^{T-t} L(t+u, X_{t+u}^{(t, \omega, \alpha)}, \alpha(u)) du \right] \geq U(t, \omega) \quad (156)$$

This being true for any $\alpha \in \mathcal{A}(\mathbb{F})$, we conclude that $U(t, \omega) \leq J(t, \omega, \alpha)$.

Assume now that $\alpha^* \in \mathcal{A}(\mathbb{F})$ verifies (154). Taking $\alpha = \alpha^*$ transforms inequalities to equalities, submartingale to martingale, hence establishes the second part of the theorem. \square

This proof can be adapted [32] to the case where all coefficients, except σ , depend on the quadratic variation of X as well, using the approach outlined in Section 6.3. The subtle point is that if σ depends on the control, the functional Itô formula (Theorem 6.4) does not apply to $U(s, X_s^\alpha, [X^\alpha]_s)$ because $\frac{d[X^\alpha](s)}{ds}$ would not necessarily admit a right-continuous representative.

8.4 Weak solutions

The above results are 'verification theorems' which, assuming the existence of a solution of the functional Kolmogorov equation with a given regularity, derive a probabilistic property or probabilistic representation of this solution. The key issue when applying such results is to be able to prove the regularity conditions needed to apply the functional Ito formula. In the case of linear Kolmogorov equations, this is equivalent to constructing, given a functional H , a *smooth* version of the conditional expectation $E[H|\mathcal{F}_t]$ i.e. a functional representation which admits vertical and horizontal derivative.

As with the case of finite-dimensional PDEs, such *strong solutions* –with the required differentiability– may fail to exist in many examples of interest and, even if they exist, proving pathwise regularity is not easy in general and requires imposing many smoothness conditions on the terminal functional H , due to the absence of any regularizing effect as in the finite dimensional parabolic case [13, 60].

A more general approach is provided by the notion of *weak solution*, which we now define, using the tools developed in Section 7. Consider, as above, a semimartingale X defined as the solution of a stochastic differential equation

$$X(t) = X(0) + \int_0^t b(u, X_u) du + \int_0^t \sigma(u, X_u) dW(u) = X(0) + \int_0^t b(u, X_u) du + M(t)$$

whose coefficients b, σ are non-anticipative path-dependent functionals, satisfying the assumptions of Proposition 8.1. We assume that M , the martingale part of X , is a square-integrable martingale.

We can then use the notion of weak derivative ∇_M on the Sobolev space $\mathcal{S}^{1,2}(M)$ of semimartingales introduced in Section 7.5.

The idea is to introduce a probability measure \mathbb{P} on $D([0, T], \mathbb{R}^d)$, under which X is a semimartingale, and requiring (134) to hold $dt \times d\mathbb{P}$ -a.e.

For a classical solution $F \in \mathbb{C}_b^{1,2}(\mathcal{W}_T)$, requiring (134) to hold on the support of M is equivalent to requiring

$$\mathcal{A}F(t, X_t) = \mathcal{D}F(t, X_t) + \nabla_\omega F(t, X_t) \cdot b(t, X_t) + \frac{1}{2} \text{tr}[\nabla_\omega^2 F(t, X_t) \sigma^t \sigma(t, X_t)] = 0$$

$dt \times d\mathbb{P}$ -a.e.

Let $\mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P})$ be the space of \mathbb{F} -adapted processes $\phi : [0, T] \times \Omega \rightarrow \mathbb{R}$ with

$$E\left(\int_0^T |\phi(t)|^2 dt\right) < \infty,$$

and $\mathcal{A}^2(\mathbb{F})$ the space of absolutely continuous processes whose Radon-Nikodym derivative lies in $\mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P})$:

$$\mathcal{A}^2(\mathbb{F}) = \left\{ \Phi \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P}), \frac{d\Phi}{dt} \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P}) \right\}$$

Using the density of $\mathcal{A}^2(\mathbb{F})$ in $\mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P})$, (157) is equivalent to requiring

$$\forall \Phi \in \mathcal{A}^2(\mathbb{F}), \quad E\left(\int_0^T \Phi(t) \mathcal{A}F(t, X_t) dt\right) = 0. \quad (157)$$

where we can choose $\Phi(0) = 0$ without loss of generality. But (157) is well-defined as soon as $\mathcal{A}F(X) \in \mathcal{L}^2(dt \times d\mathbb{P})$. This motivates the following definition:

Definition 8.17 (Sobolev space of non-anticipative functionals). We define $\mathbb{W}^{1,2}(\mathbb{P})$ as the space of ($dt \times d\mathbb{P}$ -equivalence classes of) non-anticipative functionals $F : (\Lambda_T, d_\infty) \rightarrow \mathbb{R}$ such that the process $S = F(X)$ defined by $S(t) = F(t, X_t)$ belongs to $\mathcal{S}^{1,2}(M)$. Then $S = F(X)$ has a semimartingale decomposition

$$S(t) = F(t, X_t) = S(0) + \int_0^t \nabla_M S.dM + H(t) \quad \text{with} \quad \frac{dH}{dt} \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P}).$$

$\mathbb{W}^{1,2}(\mathbb{P})$ is a Hilbert space, equipped with the norm

$$\|F\|_{1,2}^2 = \|F(X)\|_{S^{1,2}}^2 = E^{\mathbb{P}}\left(|F(0, X_0)|^2 + \int_0^T \text{tr}(\nabla_M F(X) \cdot {}^t \nabla_M F(X) \cdot d[M]) + \int_0^T \left|\frac{dH}{dt}\right|^2 dt\right)$$

Denoting by $\mathbb{L}^2(\mathbb{P})$ the space of square-integrable non-anticipative functionals, we have that for $F \in \mathbb{W}^{1,2}(\mathbb{P})$, its vertical derivative lies in $\mathbb{L}^2(\mathbb{P})$ and

$$\begin{aligned} \nabla_\omega : \mathbb{W}^{1,2}(\mathbb{P}) &\mapsto \mathbb{L}^2(\mathbb{P}) \\ F &\rightarrow \nabla_\omega F \end{aligned}$$

is a continuous map.

Using linear combinations of smooth cylindrical functionals, one can show that this definition is equivalent to the following alternative construction:

Definition 8.18 ($\mathbb{W}^{1,2}(\mathbb{P})$: alternative definition). $\mathbb{W}^{1,2}(\mathbb{P})$ is the completion of $\mathbb{C}_b^{1,2}(\mathcal{W}_T)$ with respect to the norm

$$\begin{aligned} \|F\|_{1,2}^2 = E^{\mathbb{P}} &\left(|F(0, X_0)|^2 + \int_0^T \text{tr}(\nabla_\omega F(t, X_t) \cdot {}^t \nabla_\omega F(t, X_t) \cdot d[M]) \right) \\ &+ E^{\mathbb{P}} \left(\left| \int_0^T \frac{d}{dt} \left(F(t, X_t) - \int_0^t \nabla_\omega F(u, X_u) \cdot dM \right) \right|^2 dt \right) \end{aligned}$$

In particular $\mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T) \cap \mathbb{W}^{1,2}(\mathbb{P})$ is dense in $(\mathbb{W}^{1,2}(\mathbb{P}), \|\cdot\|_{1,2})$.

For $F \in \mathbb{W}^{1,2}(\mathbb{P})$, let $S(t) = F(t, X_t)$ and M be the martingale part of X . The process $S(t) - \int_0^t \nabla_M S(u).dM$ has absolutely continuous paths and one can then define:

$$\mathcal{A}F(t, X_t) := \frac{d}{dt} \left(F(t, X_t) - \int_0^t \nabla_\omega F(u, X_u).dM \right) \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P}).$$

Remark 8.19. Note that in general it is not possible to define " $\mathcal{D}F(t, X_t)$ " as a real-valued process for $F \in \mathbb{W}^{1,2}(\mathbb{P})$. As noted in Section 7.5, this requires $F \in \mathbb{W}^{2,2}(\mathbb{P})$.

Let U be the process defined by

$$U(t) = F(T, X_T) - F(t, X_t) - \int_t^T \nabla_M F(u, X_u).dM.$$

Then the paths of U lie in the Sobolev space $H^1([0, T], \mathbb{R})$,

$$\frac{dU}{dt} = -\mathcal{A}F(t, X_t) \quad \text{and} \quad U(T) = 0.$$

Thus, we can apply the integration by parts formula for H^1 Sobolev functions [65] pathwise and rewrite (157) as

$$\begin{aligned} \forall \Phi \in \mathcal{A}^2(\mathbb{F}), \quad & \int_0^T dt \quad \Phi(t) \quad \frac{d}{dt} \left(F(t, X_t) - \int_0^t \nabla_M F(u, X_u).dM \right) \\ & = \int_0^T dt \quad \phi(t) \quad \left(F(T, X_T) - F(t, X_t) - \int_t^T \nabla_M F(u, X_u).dM \right) \end{aligned}$$

We are now ready to formulate the notion of weak solution in $\mathbb{W}^{1,2}(\mathbb{P})$:

Definition 8.20 (Weak solution). $F \in \mathbb{W}^{1,2}(\mathbb{P})$ is said to be a weak solution of

$$\mathcal{D}F(t, \omega) + b(t, \omega) \nabla_\omega F(t, \omega) + \frac{1}{2} \text{tr}[\nabla_\omega^2 F(t, \omega) \sigma^t \sigma(t, \omega)] = 0$$

on $\text{supp}(X)$ with terminal condition $H(\cdot) \in L^2(\mathcal{F}_T, \mathbb{P})$ if

$$F(T, \cdot) = H(\cdot) \quad \text{and} \quad \forall \phi \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P}),$$

$$E \left[\int_0^T dt \quad \phi(t) \left(H(X_T) - F(t, X_t) - \int_t^T \nabla_M F(u, X_u).dM \right) \right] = 0. \quad (158)$$

The preceding discussion shows that any square-integrable classical solution $F \in \mathbb{C}_{\text{loc}}^{1,2}(\mathcal{W}_T) \cap \mathcal{L}^2(dt \times d\mathbb{P})$ of the functional Kolmogorov equation (134) is also a weak solution.

However, the notion of weak solution is much more general, since the form (158) only requires the vertical derivative $\nabla_\omega F(t, X_t)$ to exist in a weak sense, i.e. in $\mathcal{L}^2(X)$ and does not require any continuity of the derivative, only square-integrability.

Existence and uniqueness of such weak solutions is much simpler to study than for classical solutions:

Theorem 8.21 (Existence and uniqueness of weak solutions). *Let*

$$H \in L^2(C^0([0, T], \mathbb{R}^d), \mathbb{P}).$$

There exists a unique weak solution $F \in \mathbb{W}^{1,2}(\mathbb{P})$ of the functional Kolmogorov equation (134) with terminal conditional $F(T, \cdot) = H(\cdot)$.

Proof. Let M be the martingale component of X and $F : \mathcal{W}_T \rightarrow \mathbb{R}$ be a regular version of the conditional expectation $E(H|\mathcal{F}_t)$:

$$F(t, \omega) = E(H|\mathcal{F}_t)(\omega) \quad dt \times d\mathbb{P} - a.e.$$

and let $Y(t) = F(t, X_t)$. Then $Y \in \mathcal{M}^2(M)$ is a square-integrable martingale so by the martingale representation formula 7.8, $Y \in \mathcal{S}^{1,2}(X)$ so $F \in \mathbb{W}^{1,2}(\mathbb{P})$, and

$$F(t, X_t) = F(T, X_T) - \int_t^T \nabla_\omega F(u, X_u).dM$$

thus F satisfies (158).

Uniqueness: To show uniqueness, we use an 'energy' method. Let F^1, F^2 be weak solutions of (134) with terminal condition H . Then $F = F^1 - F^2 \in \mathbb{W}^{1,2}(\mathbb{P})$ is a weak solution of (134) with $F(T, \cdot) = 0$. Let $Y(t) = F(t, X_t)$. Then $Y \in \mathcal{S}^{1,2}(M)$ so $U(t) = F(T, X_T) - F(t, X_t) - \int_t^T \nabla_M F(u, X_u).dM$ has absolutely continuous paths and

$$\frac{dU}{dt} = -\mathcal{A}F(t, X_t)$$

is well defined. Ito's formula then yields

$$Y(t)^2 = -2 \int_t^T Y(u). \nabla_\omega F(u, X_u).dM - 2 \int_t^T Y(u)\mathcal{A}F(u, X_u)du + [Y](t) - [Y](T).$$

The first term is a \mathbb{P} -local martingale. Let $(\tau_n)_{n \geq 1}$ be an increasing sequence of stopping times such that $\int_t^{T \wedge \tau_n} Y(u). \nabla_\omega F(u, X_u).dX$ is a martingale. Then for any $n \geq 1$,

$$E \left(\int_t^{T \wedge \tau_n} Y(u). \nabla_\omega F(u, X_u).dM \right) = 0.$$

Therefore for any $n \geq 1$,

$$E(Y(t \wedge \tau_n)^2) = 2E \left(\int_t^{T \wedge \tau_n} Y(u)\mathcal{A}F(u, X_u)du \right) + E([Y](t \wedge \tau_n) - [Y](T \wedge \tau_n)).$$

Given that F is a weak solution of (134),

$$E \left(\int_t^{T \wedge \tau_n} Y(u)\mathcal{A}F(u, X_u)du \right) = E \left(\int_t^T Y(u)1_{[0, \tau_n]} \mathcal{A}F(u, X_u)du \right) = 0$$

since $Y(u)1_{[0, \tau_n]} \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P})$. Thus,

$$E(Y(t \wedge \tau_n)^2) = E([Y](t \wedge \tau_n) - [Y](T \wedge \tau_n))$$

which is a positive increasing function of t , so

$$\forall n \geq 1, \quad E(Y(t \wedge \tau_n)^2) \leq E(Y(T \wedge \tau_n)^2).$$

Since $Y \in \mathcal{L}^2(\mathbb{F}, dt \times d\mathbb{P})$ we can use a dominated convergence argument to take $n \rightarrow \infty$ and conclude that

$$0 \leq E(Y(t)^2) \leq E(Y(T)^2) = 0$$

so $Y = 0$, i.e. $F^1(t, \omega) = F^2(t, \omega)$ outside a \mathbb{P} -evanescent set. \square

Note that uniqueness holds outside a \mathbb{P} -evanescent set: there is no assertion on uniqueness outside the support of \mathbb{P} .

The flexibility of the notion of weak solution comes from the following characterization of square-integrable martingale functionals as weak solutions, which is an analog of Theorem 8.6 but *without* any differentiability requirement:

Proposition 8.22 (Characterization of harmonic functionals).

Let $F \in \mathbb{W}^{1,2}(\mathbb{P})$ be a square-integrable non-anticipative functional. Then $Y(t) = F(t, X_t)$ is a \mathbb{P} -martingale if and only if F is a weak solution of the functional Kolmogorov equation (134).

Proof. Let F be a weak solution of (134). Then $S = F(X) \in \mathcal{S}^{1,2}(X)$ so S is weakly differentiable with $\nabla_M S \in \mathcal{L}^2(X)$, $N = \int_0^\cdot \nabla_M S.dM \in \mathcal{M}^2(X)$ is a square-integrable martingale, and $A = S - N \in \mathcal{A}^2(\mathbb{F})$. Since F verifies (158), we have

$$\forall \phi \in \mathcal{L}^2(dt \times d\mathbb{P}), \quad E\left(\int_0^T dt \phi(t)A(t)\right) = 0$$

so $A(t) = 0$ $dt \times d\mathbb{P}$ -a.e.; therefore S is a martingale.

Conversely, let $F \in \mathcal{L}^2(dt \times d\mathbb{P})$ be a non-anticipative functional such that $S(t) = F(t, X_t)$ is a martingale. Then $S \in \mathcal{M}^2(X)$ so by Theorem 7.8 S is weakly differentiable and $U(t) = F(T, X_T) - F(t, X_t) - \int_t^T \nabla_M S.dM = 0$, so F verifies (158), so F is a weak solution of the functional Kolmogorov equation with terminal condition $F(T, \cdot)$. \square

This result characterizes *all* square-integrable *harmonic functionals* as weak solutions of the functional Kolmogorov equation (134) and thus illustrates the relevance of Definition 8.20.

Comments and references

Path-dependent Kolmogorov equations first appeared in the work of Dupire [21]. Uniqueness and comparison of classical solutions were first obtained in [32, 11]. The path-dependent HJB equation and its relation with non-Markovian

stochastic control problems was first explored in [32]. The relation between FBSDEs and path-dependent PDEs is an active research topic and the fully nonlinear case, corresponding to non-Markovian optimal control and optimal stopping problems, is currently under development, using an extension of the notion of *viscosity solution* to the functional case, introduced by Ekren et al. [22, 23]. A full discussion of this notion is beyond the scope of these lecture notes and we refer to [10] and recent work by Ekren et al [22, 23], Peng [59, 60] and Cosso [14].

References

- [1] Aase, K., Oksendal, B., Di Nunno, G.: White noise generalizations of the Clark-Haussmann-Ocone theorem with application to mathematical finance. *Finance and Stochastics* **4**, 465–496 (2000)
- [2] Bass, R.F.: *Diffusions and elliptic operators*. Springer (1998)
- [3] Bismut, J.M.: A generalized formula of Itô and some other properties of stochastic flows. *Z. Wahrsch. Verw. Gebiete* **55**(3), 331–350 (1981)
- [4] Bismut, J.M.: Calcul des variations stochastique et processus de sauts. *Z. Wahrsch. Verw. Gebiete* **63**(2), 147–235 (1983)
- [5] Clark, J.M.C.: The representation of functionals of Brownian motion by stochastic integrals. *Ann. Math. Statist.* **41**, 1282–1295 (1970)
- [6] Cont, R.: Functional Ito calculus. Part I: Pathwise calculus for non-anticipative functionals. Working paper, Laboratoire de Probabilités et Modèles Aléatoires (2011)
- [7] Cont, R.: Functional Ito calculus. Part II: Weak calculus for functionals of martingales. Working paper, Laboratoire de Probabilités et Modèles Aléatoires (2012)
- [8] Cont, R., Fournié, D.A.: Change of variable formulas for non-anticipative functionals on path space. *Journal of Functional Analysis* **259**(4), 1043–1072 (2010). URL <http://dx.doi.org/10.1016/j.jfa.2010.04.017>
- [9] Cont, R., Fournié, D.A.: A functional extension of the Ito formula. *Comptes Rendus Mathématique Acad. Sci. Paris Ser. I* **348**, 57–61 (2010). URL <http://dx.doi.org/10.1016/j.crma.2009.11.013>
- [10] Cont, R., Fournié, D.A.: Functional Kolmogorov equations. Working Paper (2010)
- [11] Cont, R., Fournié, D.A.: Functional Ito calculus and stochastic integral representation of martingales. *Annals of Probability* **41**(1), 109–133 (2013). URL <http://projecteuclid.org/euclid.aop/1358951982>
- [12] Cont, R., Lu, Y.: Weak approximations for martingale representations. *Stochastic Processes and Applications* **in press** (2014). DOI 10.1016/j.spa.2015.10.002. URL <http://arxiv.org/abs/1501.00383>
- [13] Cont, R., Riga, C.: Robustness and pathwise analysis of hedging strategies for path-dependent derivatives. Working Paper (2014)
- [14] Cosso, A.: Viscosity solutions of path-dependent PDEs and non-markovian forward-backward stochastic equations. arXiv:1202.2502 (2013)

- [15] Da Prato, G., Zabczyk, J.: Stochastic equations in infinite dimensions, *Encyclopedia of Mathematics and its Applications*, vol. 44. Cambridge University Press (1992)
- [16] Da Prato, G., Zabczyk, J.: Second order partial differential equations in Hilbert spaces, *London Mathematical Society Lecture Note Series*, vol. 293. Cambridge University Press, Cambridge (2002)
- [17] Davis, M.H.: Functionals of diffusion processes as stochastic integrals. *Math. Proc. Camb. Phil. Soc.* **87**, 157–166 (1980)
- [18] Davis, M.H.A.: Martingale methods in stochastic control. In: Stochastic control theory and stochastic differential systems, *Lecture Notes in Control and Information Sci.*, vol. 16, pp. 85–117. Springer, Berlin (1979)
- [19] Dellacherie, C., Meyer, P.A.: Probabilities and potential, *North-Holland Mathematics Studies*, vol. 29. North-Holland Publishing Co., Amsterdam (1978)
- [20] Dudley, R.M.: Sample functions of the Gaussian process. *Annals of Probability* **1**(1), 66–103 (1973)
- [21] Dupire, B.: Functional Itô calculus. Portfolio Research Paper 2009-04, Bloomberg (2009)
- [22] Ekren, I., Keller, C., Touzi, N., Zhang, J.: On viscosity solutions of path dependent PDEs. *Annals of Probability* **42**, 204–236 (2014)
- [23] Ekren, I., Touzi, N., Zhang, J.: Viscosity solutions of fully nonlinear parabolic path dependent pdes: Part I. arXiv:1210.0006 (2013)
- [24] El Karoui, N., Mazliak, L.: Backward stochastic differential equations, vol. 364. CRC Press (1997)
- [25] Elliott, R.J., Kohlmann, M.: A short proof of a martingale representation result. *Statistics & Probability Letters* **6**(5), 327–329 (1988)
- [26] Feyel, D., De la Pradelle, A.: Espaces de Sobolev gaussiens. *Annales de l’Institut Fourier* **39**(4), 875–908 (1989)
- [27] Fitzsimmons, P., Rajeev, B.: A new approach to the martingale representation theorem. *Stochastics* **81**, 467–476 (2009)
- [28] Flandoli, F., Zanco, G.: An infinite-dimensional approach to path-dependent Kolmogorov equations. Working Paper arXiv:1312.6165 (2013)
- [29] Fleming, W., Soner, H.: Controlled Markov Processes and Viscosity Solutions. Springer Verlag (2005)
- [30] Fliess, M.: Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bulletin de la société mathématique de France* **109**, 3–40 (1981)

- [31] Föllmer, H.: Calcul d'Itô sans probabilités. In: Séminaire de Probabilités XV, *Lecture Notes in Math.*, vol. 850, pp. 143–150. Springer, Berlin (1981)
- [32] Fournié, D.A.: Functional Ito calculus and applications. Ph.D. thesis (2010)
- [33] Fournié, E., Lasry, J.M., Lebuchoux, J., Lions, P.L., Touzi, N.: Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics* **3**(4), 391–412 (1999)
- [34] Friz, P.K., Victoir, N.B.: Multidimensional stochastic processes as rough paths: theory and applications. Cambridge University Press (2010)
- [35] Gubinelli, M.: Controlling rough paths. *Journal of Functional Analysis* **216**(1), 86 – 140 (2004)
- [36] Haussmann, U.G.: Functionals of Itô processes as stochastic integrals. *SIAM J. Control Optimization* **16**(2), 252–269 (1978)
- [37] Haussmann, U.G.: On the integral representation of functionals of Itô processes. *Stochastics* **3**(1), 17–27 (1979)
- [38] Hida, T.: White noise: an infinite dimensional calculus, vol. 253. Springer (1993)
- [39] Imkeller, P., Réveillac, A., Richter, A.: Differentiability of quadratic BSDEs generated by continuous martingales. *Ann. Appl. Probab.* **22**(1), 285–336 (2012)
- [40] Ito, K.: On a stochastic integral equation. *Proceedings of the Imperial Academy of Tokyo* **20**, 519–524 (1944)
- [41] Ito, K.: On stochastic differential equations. *Proceedings of the Imperial Academy of Tokyo* **22**, 32–35 (1946)
- [42] Jacod, J., Méléard, S., Protter, P.: Explicit form and robustness of martingale representations. *Ann. Probab.* **28**(4), 1747–1780 (2000)
- [43] Jacod, J., Shiryaev, A.N.: Limit theorems for stochastic processes. Springer, Berlin (2003)
- [44] Karatzas, I., Ocone, D.L., Li, J.: An extension of Clark's formula. *Stochastics Stochastics Rep.* **37**(3), 127–131 (1991)
- [45] Karatzas, I., Shreve, S.: Brownian motion and Stochastic calculus. Springer (1987)
- [46] Kolmogoroff, A.: Über die analytischen methoden in der wahrscheinlichkeitsrechnung. *Mathematische Annalen* **104**(1), 415–458 (1931)
- [47] Lévy, P.: Le mouvement brownien plan. *American Journal of Mathematics* **62**, 487–550 (1940)

- [48] Li, X.D., Lyons, T.J.: Smoothness of Itô maps and diffusion processes on path spaces. I. *Ann. Sci. École Norm. Sup. (4)* **39**(4), 649–677 (2006)
- [49] Lyons, T.J.: Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14**(2), 215–310 (1998)
- [50] Ma, J., Yin, H., Zhang, J.: Solving forward–backward SDEs explicitly: a four-step scheme. *Probability theory and related fields* **98**, 339–359 (1994)
- [51] Ma, J., Yin, H., Zhang, J.: On non-markovian forward–backward SDEs and backward stochastic PDEs. *Stochastic Processes and Their Applications* **122**(12), 3980–4004 (2012)
- [52] Malliavin, P.: *Stochastic analysis*. Springer (1997)
- [53] Meyer, P.: Un cours sur les integrales stochastiques. In: *Séminaire de Probabilités X, Lecture Notes in Math.*, vol. 511, pp. 245–400. Springer (1976)
- [54] Millet, A., Sanz-Solé, M.: A simple proof of the support theorem for diffusion processes. In: *Séminaire de Probabilités XXVIII*, pp. 36–48. Springer (1994)
- [55] Nualart, D.: Malliavin calculus and its applications, *CBMS Regional Conference Series in Mathematics*, vol. 110. CBMS, Washington, DC (2009)
- [56] Ocone, D.L.: Malliavin’s calculus and stochastic integral representations of functionals of diffusion processes. *Stochastics* **12**(3–4), 161–185 (1984)
- [57] Pardoux, E., Peng, S.: Adapted solutions of backward stochastic equations. *Systems and Control Letters* **14**, 55–61 (1990)
- [58] Pardoux, E., Peng, S.: Backward stochastic differential equations and quasilinear parabolic partial differential equations. In: *Stochastic partial differential equations and their applications, Lecture Notes in Control and Informatic Science*, vol. 716, pp. 200–217. Springer (1992)
- [59] Peng, S.: Note on viscosity solution of path-dependent PDE and g-martingales. arXiv (2011)
- [60] Peng, S., Wang, F.: BSDE, path-dependent PDE and nonlinear feynman-kac formula. <http://arxiv.org/abs/1108.4317> (2011)
- [61] Peng, S.G.: Stochastic Hamilton-Jacobi-Bellman equations. *SIAM J. Control Optim.* **30**(2), 284–304 (1992)
- [62] Protter, P.E.: *Stochastic integration and differential equations*. Springer-Verlag, Berlin (2005). Second edition.

- [63] Revuz, D., Yor, M.: Continuous martingales and Brownian motion, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 293, third edn. Springer-Verlag, Berlin (1999)
- [64] Schied, A.: On a class of generalized Takagi functions with linear pathwise quadratic variation. *Journal of Mathematical Analysis and Applications* **433**(2), 974–990 (2016)
- [65] Schwartz, L.: *Théorie des distributions*. Editions Hermann (1969)
- [66] Shigekawa, I.: Derivatives of Wiener functionals and absolute continuity of induced measures. *J. Math. Kyoto Univ.* **20**(2), 263–289 (1980)
- [67] Sondermann, D.: Introduction to stochastic calculus for finance, *Lecture Notes in Economics and Mathematical Systems*, vol. 579. Springer-Verlag, Berlin (2006)
- [68] Stroock, D.W.: The Malliavin calculus, a functional analytic approach. *J. Funct. Anal.* **44**(2), 212–257 (1981)
- [69] Stroock, D.W., Varadhan, S.R.S.: On the support of diffusion processes with applications to the strong maximum principle. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (1970/1971)*, Vol. III: Probability theory, pp. 333–359. Univ. California Press, Berkeley, Calif. (1972)
- [70] Stroock, D.W., Varadhan, S.R.S.: Multidimensional diffusion processes, *Grundlehren der Mathematischen Wissenschaften*, vol. 233. Springer-Verlag, Berlin (1979)
- [71] Sugita, H.: On the characterization of Sobolev spaces on an abstract Wiener space. *J. Math. Kyoto Univ.* **25**(4), 717–725 (1985)
- [72] Sugita, H.: Sobolev spaces of Wiener functionals and Malliavin’s calculus. *J. Math. Kyoto Univ.* **25**(1), 31–48 (1985)
- [73] Taylor, S.J.: Exact asymptotic estimates of Brownian path variation. *Duke Mathematical Journal* **39**, 219–241 (1972)
- [74] Watanabe, S.: Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Ann. Probab.* **15**(1), 1–39. (1987)
- [75] Yong, J., Zhou, X.Y.: *Stochastic controls: Hamiltonian systems and HJB equations*, vol. 43. Springer (1999)
- [76] Zabczyk, J.: Remarks on stochastic derivation. *Bulletin de l’Académie Polonaise des Sciences* **21**(3), 263–269 (1973)