



**HAL**  
open science

# Functional Ito calculus and stochastic integral representation of martingales

Rama Cont, David-Antoine Fournie

► **To cite this version:**

Rama Cont, David-Antoine Fournie. Functional Ito calculus and stochastic integral representation of martingales. 2010. hal-00455700v1

**HAL Id: hal-00455700**

**<https://hal.science/hal-00455700v1>**

Preprint submitted on 11 Feb 2010 (v1), last revised 27 Sep 2011 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Functional Ito calculus and stochastic integral representation of martingales

Rama Cont

David-Antoine Fournié

First draft: June 2009. This version: Feb 2010.\*

## Abstract

We develop a non-anticipative calculus for functionals of a continuous semimartingale, using a notion of pathwise functional derivative. A functional extension of the Ito formula is derived and used to obtain a constructive martingale representation theorem for a class of continuous martingales verifying a regularity property. By contrast with the Clark-Haussmann-Ocone formula, this representation involves non-anticipative quantities which can be computed pathwise.

These results are used to construct a weak derivative acting on square-integrable martingales, which is shown to be the inverse of the Ito integral, and derive an integration by parts formula for Ito stochastic integrals. We show that this weak derivative may be viewed as a non-anticipative “lifting” of the Malliavin derivative.

Regular functionals of an Ito martingale which have the local martingale property are characterized as solutions of a functional differential equation, for which a uniqueness result is given.

Keywords: stochastic calculus, functional calculus, Ito formula, integration by parts, Malliavin derivative, martingale representation, semimartingale, Wiener functionals, functional Feynman-Kac formula, Kolmogorov equation, Clark-Ocone formula.

---

\*We thank Bruno Dupire for sharing with us his original ideas. R. Cont thanks Hans Föllmer and especially Jean Jacod for helpful comments and discussions.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Functionals representation of non-anticipative processes</b>	<b>4</b>
2.1	Horizontal and vertical perturbation of a path . . . . .	5
2.2	Regularity for non-anticipative functionals . . . . .	6
2.3	Measurability properties . . . . .	7
<b>3</b>	<b>Pathwise derivatives of non-anticipative functionals</b>	<b>9</b>
3.1	Horizontal and vertical derivatives . . . . .	9
3.2	Obstructions to regularity . . . . .	11
3.3	Pathwise derivatives of an adapted process . . . . .	12
<b>4</b>	<b>Functional Ito formula</b>	<b>17</b>
<b>5</b>	<b>Martingale representation formula</b>	<b>21</b>
5.1	Martingale representation theorem . . . . .	21
5.2	Relation with the Malliavin derivative . . . . .	22
<b>6</b>	<b>Weak derivatives and integration by parts for stochastic integrals</b>	<b>24</b>
<b>7</b>	<b>Functional equations for martingales</b>	<b>27</b>
<b>A</b>	<b>Proof of Theorems 8 and 18</b>	<b>32</b>
A.1	Proofs of theorem 8 . . . . .	32
A.2	Proofs of Theorem 18 . . . . .	33

# 1 Introduction

Ito's stochastic calculus [15, 16, 8, 24, 20, 28] has proven to be a powerful and useful tool in analyzing phenomena involving random, irregular evolution in time.

Two characteristics distinguish the Ito calculus from other approaches to integration, which may also apply to stochastic processes. First is the possibility of dealing with processes, such as Brownian motion, which have non-smooth trajectories with infinite variation. Second is the *non-anticipative* nature of the quantities involved: viewed as a functional on the space of paths indexed by time, a non-anticipative quantity may only depend on the underlying path up to the *current* time. This notion, first formalized by Doob [9] in the 1950s via the concept of a *filtered* probability space, is the mathematical counterpart to the idea of causality.

Two pillars of stochastic calculus are the theory of **stochastic integration**, which allows to define integrals  $\int_0^T Y dX$  for of a large class of non-anticipative integrands  $Y$  with respect to a *semimartingale*  $X = (X(t), t \in [0, T])$ , and the **Ito formula** [15, 16, 24] which allows to represent smooth functions  $Y(t) = f(t, X(t))$  of a semimartingale in terms of such stochastic integrals. A central concept in both cases is the notion of *quadratic variation*  $[X]$  of a semimartingale, which differentiates Ito calculus from the calculus of smooth functions. Whereas the class of integrands  $Y$  covers a wide range of *non-anticipative path-dependent functionals* of  $X$ , the Ito formula is limited to *functions* of the current value of  $X$ .

Given that in many applications such as statistics of processes, physics or mathematical finance, one is led to consider functionals of a semimartingale  $X$  and its quadratic variation process  $[X]$  such as:

$$\int_0^t g(t, X_t) d[X](t), \quad G(t, X_t, [X]_t), \quad \text{or} \quad E[G(T, X(T), [X](T)) | \mathcal{F}_t] \quad (1)$$

(where  $X(t)$  denotes the value at time  $t$  and  $X_t = (X(u), u \in [0, t])$  the path up to time  $t$ ) there has been a sustained interest in extending the framework of stochastic calculus to such path-dependent functionals.

In this context, the Malliavin calculus [3, 4, 25, 23, 26, 29, 30] has proven to be a powerful tool for investigating various properties of Brownian functionals, in particular the smoothness of their densities.

Yet the construction of Malliavin derivative, which is a weak derivative for functionals on Wiener space, does not refer to the underlying filtration  $\mathcal{F}_t$ . Hence, it naturally leads to representations of functionals in terms of *anticipative* processes [4, 14, 26], whereas in applications it is more natural to consider non-anticipative, or causal, versions of such representations.

In a recent insightful work, B. Dupire [10] has proposed a method to extend the Ito formula to a functional setting in a *non-anticipative* manner. Building on this insight, we develop hereafter a non-anticipative calculus [6] for a class of functionals -including the above examples- which may be represented as

$$Y(t) = F_t(\{X(u), 0 \leq u \leq t\}, \{A(u), 0 \leq u \leq t\}) = F_t(X_t, A_t) \quad (2)$$

where  $A$  is the local quadratic variation defined by  $[X](t) = \int_0^t A(u) du$  and the functional

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

represents the dependence of  $Y$  on the underlying path and its quadratic variation. For such functionals, we define an appropriate notion of regularity (Section 2.2) and a non-anticipative notion of pathwise derivative (Section 3). Introducing  $A_t$  as additional variable allows us to control the dependence of  $Y$  with respect to the "quadratic variation"  $[X]$  by requiring smoothness properties of  $F_t$  with respect to the variable  $A_t$  in the supremum norm, without resorting to  $p$ -variation norms as in rough path theory [21]. This allows to consider a wider range of functionals, as in (1).

Using these pathwise derivatives, we derive a functional Ito formula (Section 4), which extends the usual Ito formula in two ways: it allows for path-dependence and for dependence with respect to quadratic variation process (Theorem 18). This result gives a rigorous mathematical framework for developing and extending the ideas proposed by B. Dupire [10] to a larger class of functionals which notably allow for dependence on the quadratic variation along a path.

We use the functional Ito formula to derive a constructive version of the martingale representation theorem (Section 5), which can be seen as a non-anticipative form of the Clark-Haussmann-Ocone formula [4, 13, 14, 26].

The martingale representation formula allows to obtain an integration by parts formula for Ito stochastic integrals (Theorem 24), which enables in turn to define a weak functional derivative for a class of square-integrable martingales (Section 6). We argue that this weak derivative may be viewed as a non-anticipative "lifting" of the Malliavin derivative (Theorem 29).

Finally, we show that regular functionals of an Ito martingale which have the local martingale property are characterized as solutions of a functional analogue of Kolmogorov's backward equation (Section 7), for which a uniqueness result is given (Theorem 32).

Our method follows the spirit of H. Föllmer's [12] pathwise approach to Ito calculus. Sections 2, 3 and 4 are essentially "pathwise" results which can in fact be restated in purely analytical terms [5]. Probabilistic considerations become prominent when applying the functional calculus to martingales (Sections 5, 6 and 7).

## 2 Functionals representation of non-anticipative processes

Let  $X : [0, T] \times \Omega \mapsto \mathbb{R}^d$  be a continuous,  $\mathbb{R}^d$ -valued semimartingale defined on a filtered probability space  $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$ . The paths of  $X$  then lie in  $C_0([0, T], \mathbb{R}^d)$ , which we will view as a subspace of  $D([0, t], \mathbb{R}^d)$  the space of cadlag functions with values in  $\mathbb{R}^d$ . For a path  $x \in D([0, T], \mathbb{R}^d)$ , denote by  $x(t)$  the value of  $x$  at  $t$  and by  $x_t = (x(u), 0 \leq u \leq t)$  the restriction of  $x$  to  $[0, t]$ . Thus  $x_t \in D([0, t], \mathbb{R}^d)$ . For a process  $X$  we shall similarly denote  $X(t)$  its value at  $t$  and  $X_t = (X(u), 0 \leq u \leq t)$  its path on  $[0, t]$ .

Denote by  $\mathcal{F}_t = \mathcal{F}_{t+}^X$  the right-continuous augmentation of the natural filtration of  $X$  and by  $[X] = ([X^i, X^j], i, j = 1..d)$  the quadratic (co-)variation process, taking values in the set  $S_d^+$  of positive  $d \times d$  matrices. We assume that

$$[X](t) = \int_0^t A(s) ds \tag{3}$$

for some cadlag process  $A$  with values in  $S_d^+$ . The paths of  $A$  lie in  $\mathcal{S}_t = D([0, t], S_d^+)$ , the space of cadlag functions with values  $S_d^+$ .

A process  $Y : [0, T] \times \Omega \mapsto \mathbb{R}^d$  which is progressively measurable with respect to  $\mathcal{F}_t$  may be represented as

$$Y(t) = F_t(\{X(u), 0 \leq u \leq t\}, \{A(u), 0 \leq u \leq t\}) = F_t(X_t, A_t) \tag{4}$$

where  $F = (F_t)_{t \in [0, T]}$  is a family of functionals

$$F_t : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t \rightarrow \mathbb{R}$$

representing the dependence of  $Y(t)$  on the underlying path of  $X$  and its quadratic variation.

Introducing the process  $A$  as additional variable may seem redundant at this stage: indeed  $A(t)$  is itself  $\mathcal{F}_t$ -measurable i.e. a functional of  $X_t$ . However, it is not a *continuous* functional with respect to the supremum norm or other usual topologies on  $D([0, t], \mathbb{R}^d)$ . Introducing  $A_t$  as a second argument in the functional will allow us to control the regularity of  $Y$  with respect to  $[X]_t = \int_0^t A(u) du$  without resorting to p-variation norms, simply by requiring continuity of  $F_t$  in supremum or  $L^p$  norms with respect to the second variable (see Section 2.2).

As a result of the non-anticipative character of the functional,  $F_t$  only depends on the path up to  $t$ . This motivates viewing  $F = (F_t)_{t \in [0, T]}$  as a map defined on the vector bundle:

$$\Upsilon = \bigcup_{t \in [0, T]} D([0, t], \mathbb{R}^d) \times D([0, t], S_d^+) \quad (5)$$

**Definition 1** (Non-anticipative functional on path space). A non-anticipative functional on  $\Upsilon$  is a family  $F = (F_t)_{t \in [0, T]}$  where

$$\begin{aligned} F_t : D([0, t], \mathbb{R}^d) \times D([0, t], S_d^+) &\mapsto \mathbb{R} \\ (x, v) &\rightarrow F_t(x, v) \end{aligned}$$

is measurable with respect to  $\mathcal{B}_t$ , the filtration generated by the canonical process on  $D([0, t], \mathbb{R}^d) \times D([0, t], S_d^+)$ .

We denote

$$\Upsilon_c = \bigcup_{t \in [0, T]} C([0, t], \mathbb{R}^d) \times D([0, t], S_d^+) \quad (6)$$

the sub-bundle where the first element is a continuous path.

## 2.1 Horizontal and vertical perturbation of a path

Consider a path  $x \in D([0, T], \mathbb{R}^d)$  and denote by  $x_t \in D([0, t], \mathbb{R}^d)$  its restriction to  $[0, t]$  for  $t < T$ . For  $h \geq 0$ , the *horizontal* extension  $x_{t,h} \in D([0, t+h], \mathbb{R}^d)$  of  $x_t$  to  $[0, t+h]$  is defined as

$$x_{t,h}(u) = x(u) \quad u \in [0, t]; \quad x_{t,h}(u) = x(t) \quad u \in ]t, t+h] \quad (7)$$

For  $h \in \mathbb{R}^d$ , we define the *vertical* perturbation  $x_t^h$  of  $x_t$  as the cadlag path obtained by shifting the endpoint by  $h$ :

$$x_t^h(u) = x_t(u) \quad u \in [0, t[ \quad x_t^h(t) = x(t) + h \quad (8)$$

or in other words  $x_t^h(u) = x_t(u) + h1_{t=u}$ .

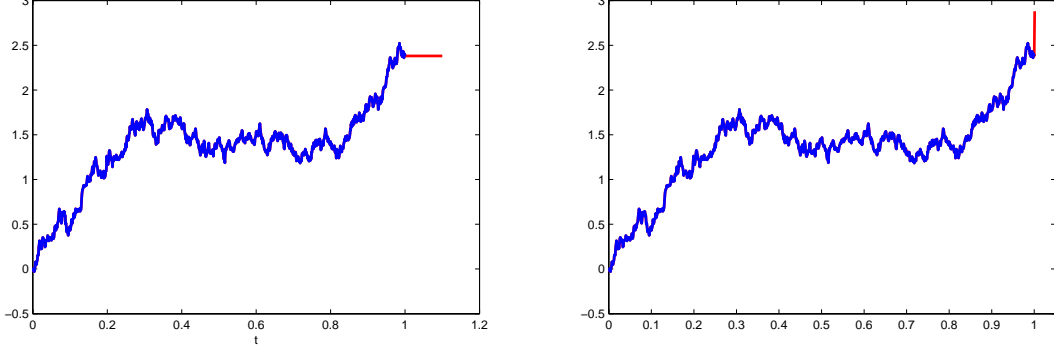


Figure 1: Left: horizontal extension  $x_{t,h}$  of a path  $x \in C_0([0, t], \mathbb{R})$ . Right: vertical extension  $x_t^h$ .

We now define two notions of distance between two paths, not necessarily defined on the same time interval. For  $T \geq t' = t + h \geq t \geq 0$ ,  $(x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t^+$  and  $(x', v') \in D([0, t + h], \mathbb{R}^d) \times \mathcal{S}_{t+h}$  define

$$d_\infty((x, v), (x', v')) = \sup_{u \in [0, t+h]} |x_{t,h}(u) - x'(u)| + \sup_{u \in [0, t+h]} |v_{t,h}(u) - v'(u)| + h \quad (9)$$

$$d_{\infty,1}((x, v), (x', v')) = \sup_{u \in [0, t+h]} |x_{t,h}(u) - x'(u)| + \int_0^{t+h} |v_{t,h}(u) - v'(u)| du + h \quad (10)$$

If the paths  $(x, v), (x', v')$  are defined on the same time interval, then  $d_\infty((x, v), (x', v'))$  is simply the distance in supremum norm. The introduction of the distance  $d_{\infty,1}$  is motivated by the fact that, if  $X^i, i = 1, 2$  are continuous semimartingales with quadratic variation  $[X]^i(t) = \int_0^t A^i(u) du$  then:

$$d_{\infty,1}((X_t^1, A_t^1), (X_t^2, A_t^2)) = \|X_t^1 - X_t^2\|_\infty + \|[X_t^1] - [X_t^2]\|_{TV} \quad (11)$$

where  $\|\cdot\|_{TV}$  is the total variation norm. This will give us an appropriate definition of continuity for functionals depending on the quadratic variation process.

## 2.2 Regularity for non-anticipative functionals

Using the distances defined above, we now introduce classes of (right) continuous non-anticipative functional on  $\Upsilon$ .

**Definition 2** (Right-continuous functionals). Define  $\mathbb{F}_r^\infty$  as the set of functionals  $F = (F_t, t \in [0, T])$  on  $\Upsilon$  which are "right-continuous" for the  $d_\infty$  metric:

$$\begin{aligned} \forall t \in [0, T[, \forall h \in [0, T - t] \quad \forall \epsilon > 0, \quad \exists \eta > 0, \\ \forall (x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \quad \forall (x', v') \in D([0, t + h], \mathbb{R}^d) \times \mathcal{S}_{t+h}, \\ d_\infty((x, v), (x', v')) < \eta \Rightarrow |F_t(x, v) - F_{t+h}(x', v')| < \epsilon \end{aligned} \quad (12)$$

**Definition 3** (Continuous functionals). Define  $\mathbb{F}^\infty$  as the set of functionals  $F = (F_t, t \in [0, T])$  on  $\Upsilon$  which are continuous up to time  $T$  for the  $d_\infty$  metric:

$$\begin{aligned} \forall t \in [0, T[, \quad \forall (x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \forall \epsilon > 0, \exists \eta > 0, \forall t' \in [0, T[, \\ \forall (x', v') \in D([0, t'], \mathbb{R}^d) \times \mathcal{S}_{t'}, d_\infty((x, v), (x', v')) < \eta \Rightarrow |F_t(x, v) - F_{t'}(x', v')| < \epsilon \end{aligned} \quad (13)$$

Most examples of functionals discussed in the introduction are continuous, in total variation norm, with respect to the path  $[X]_t$  of the quadratic variation process  $[X](t) = \int_0^t A(u)du$  i.e. continuous in  $L^1$ -norm with respect to the path  $A_t$  of its derivative. This motivates the following definition:

**Definition 4.** Define  $\mathbb{F}^{\infty,1}$  as the set of functionals  $F = (F_t, t \in [0, T])$  on  $\Upsilon$  which are continuous up to time  $T$  for the  $d_{\infty,1}$  metric:

$$\begin{aligned} \forall t \in [0, T], \quad \forall (x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \forall \epsilon > 0, \exists \eta > 0, \forall t' \in [0, T], \\ \forall (x', v') \in D([0, t'], \mathbb{R}^d) \times \mathcal{S}_{t'}, d_{\infty,1}((x, v), (x', v')) < \eta \Rightarrow |F_t(x, v) - F_{t'}(x', v')| < \epsilon \end{aligned} \quad (14)$$

and

$$\begin{aligned} \forall x \in D([0, T], \mathbb{R}^d), \forall v \in \mathcal{S}_T^+, \exists \eta > 0, C > 0, \quad \forall x' \in D([0, t], \mathbb{R}^d), \forall v' \in \mathcal{S}_t, \\ d_\infty((x_t, v_t), (x', v')) < \eta \Rightarrow |F_t(x', v_t) - F_t(x', v')| \leq C \|v - v'\|_1 \end{aligned} \quad (15)$$

We call a functional ‘‘boundedness preserving’’ if it remains bounded on each bounded set of paths, in the following sense:

**Definition 5** ( Boundedness-preserving functionals). Define  $\mathbb{B}([0, T])$  as the set of non-anticipative functionals  $F$  on  $\Upsilon([0, T])$  such that for every compact subset  $K$  of  $\mathbb{R}^d$ , every  $R > 0$  and  $t_0 < T$  there exists a constant  $C_{K,R,t_0}$  such that:

$$\forall t \leq t_0, \forall (x, v) \in D([0, t], K) \times \mathcal{S}_t, \sup_{s \in [0, t]} |v(s)| < R \Rightarrow |F_t(x, v)| < C_{K,R,t_0} \quad (16)$$

*Remark 6.* We note that  $\mathbb{F}^{\infty,1} \subset \mathbb{F}^\infty \subset \mathbb{F}_r^\infty$  and that  $d_\infty$ -convergence is stronger than  $d_{\infty,1}$ -convergence.

### 2.3 Measurability properties

Composing a non-anticipative functional  $F$  with the process  $(X, A)$  yields an  $\mathcal{F}_t$ -adapted process  $Y(t) = F_t(X_t, A_t)$ . The results below link the measurability and pathwise regularity of  $Y$  to the regularity of the functional  $F$  in terms of the classes  $\mathbb{F}_r^\infty, \mathbb{F}^\infty, \mathbb{F}^{\infty,1}$  defined above.

**Lemma 7** (Pathwise regularity).



1. If  $F \in \mathbb{F}_r^\infty$  then for any  $(x, v) \in D([0, T], \mathbb{R}^d) \times D([0, T], S)$ , the path  $t \mapsto F_t(x_t, v_t)$  is right continuous.
2. If  $F \in \mathbb{F}^\infty$  then for any  $(x, v) \in D([0, T], \mathbb{R}^d) \times D([0, T], S)$ , the path  $t \mapsto F_t(x_t, v_t)$  is cadlag and continuous at all points where  $(x, v)$  is continuous.
3. If  $F \in \mathbb{F}^{\infty,1}$  then for any  $(x, v) \in D([0, T], \mathbb{R}^d) \times D([0, T], S)$ , the path  $t \mapsto F_t(x_t, v_t)$  is furthermore cadlag and continuous at all points where  $x$  is continuous.

*Proof.* 1. Let  $F \in \mathbb{F}_r^\infty$ . For  $h > 0$  sufficiently small,

$$d_\infty((x_{t+h}, v_{t+h}), (x_t, v_t)) = \sup_{u \in (t, t+h]} |x(u) - x(t)| + \sup_{u \in (t, t+h]} |v(u) - v(t)| + h \quad (17)$$

Since both  $x$  and  $v$  are cadlag, this quantity converges to 0 as  $h \rightarrow 0+$ . The  $d_\infty$  right continuity of  $F$  at  $(x, v)$  then implies

$$F_{t+h}(x_{t+h}, v_{t+h}) - F_t(x_t, v_t) \xrightarrow{h \rightarrow 0^+} 0$$

so  $t \mapsto F_t(x_t, v_t)$  is right continuous.

2. If  $F \in \mathbb{F}^\infty$  and that the jump of  $(x, v)$  at time  $t$  is  $(\delta_x, \delta_v)$ . Then

$$d_\infty((x_{t-h}, v_{t-h}), (x_t^{-\delta_x}, v_t^{-\delta_v})) = \sup_{u \in [t-h, t)} |x(u) - x(t)| + \sup_{u \in [t-h, t)} |v(u) - v(t)| + h$$

and this quantity goes to 0 because  $x$  and  $v$  have left limits. Hence the path has left-limit  $F_t(x_t^{-\delta_x}, v_t^{-\delta_v})$  at  $t$ .

3. Assume now that  $F \in \mathbb{F}^{\infty,1}$ , and that  $(x, v)$  is continuous at  $t$ .

$$d_{\infty,1}((x_{t-h}, v_{t-h}), (x_t, v_t)) = \sup_{u \in (t-h, t]} |x(u) - x(t-h)| + \int_{t-h}^t |v(u) - v(t-h)| + h \quad (18)$$

As  $h \rightarrow 0$  the integral term goes to 0 since  $v$  is cadlag hence bounded on  $[0, T]$ . So if  $x$  is continuous at  $t$ , (18) goes to zero as  $h \rightarrow 0$  and the  $d_{\infty,1}$  continuity of  $F$  at  $(x, v)$  yields the result. If  $x$  has jump  $\delta$  at  $t$ , apply the same argument to  $x^{-\delta}$  to find  $F_t(x^{-\delta}, v)$  as left limit.  $\square$

**Theorem 8.** *Let  $F \in \mathbb{F}_r^\infty$ . Then  $Y(t) = F_t(X_t, A_t)$  defines an optional process. If  $A$  is a.s. continuous, then  $Y$  is a predictable process.*

In particular, any  $F \in \mathbb{F}_r^\infty$  is a non-anticipative functional in the sense of Definition 1.

We propose first an easy-to-read proof of this theorem under the additional assumption that  $A$  is a continuous process. The (more technical) proof for the cadlag case is given in the Appendix A.1.

*Continuous case.* Assume that  $F \in \mathbb{F}_r^\infty$  and that the paths of  $(X, A)$  are almost-surely continuous. Then by Lemma 7, the paths of  $Y$  are almost-surely right continuous. so it is enough to prove that

$Y_t$  is  $\mathcal{F}_t$ -measurable. Introduce the subdivision  $t_n^i = \frac{iT}{2^n}, i = 0..2^n$  of  $[0, T]$ , as well as the following piecewise-constant approximations of  $X$  and  $A$ :

$$\begin{aligned} X^n(t) &= \sum_{k=0}^{2^n} X(t_k^n) 1_{[t_k^n, t_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \\ A^n(t) &= \sum_{k=0}^{2^n} A(t_k^n) 1_{[t_k^n, t_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \end{aligned} \quad (19)$$

The random variable  $Y^n(t) = F_t(X_t^n, A_t^n)$  is a continuous function of the random variables  $\{X(t_k^n), A(t_k^n), t_k^n \leq t\}$  hence is  $\mathcal{F}_t$ -measurable. The representation above shows in fact that  $Y^n(t)$  is  $\mathcal{F}_t$ -measurable.  $X_t^n$  and  $A_t^n$  converge respectively to  $X_t$  and  $A_t$  almost-surely so  $Y^n(t) \xrightarrow{n \rightarrow \infty} Y(t)$  a.s., hence  $Y(t)$  is  $\mathcal{F}_t$ -measurable.

To show predictability of  $Y(t)$ , we will express it as limit of caglad adapted processes. For  $t \in [0, T]$ , define  $i^n(t)$  to be the integer such that  $t \in (\frac{(i-1)T}{n}, \frac{iT}{n}]$ . Define the process:  $Y^n((x, v), t) = F_{i^n(t)}(X_{t, \frac{i^n(t)T}{n} - t}, A_{t, \frac{i^n(t)T}{n} - t})$ , which has left-continuous trajectories since as

$$d_\infty \left( (X_{t, \frac{i^n(t)T}{n} - t}, A_{t, \frac{i^n(t)T}{n} - t}), (X_t, A_t) \right) \xrightarrow{s \rightarrow t^-} 0 \quad a.s.$$

Moreover,  $Y^n(t)$  is  $\mathcal{F}_t$ -measurable by the same approximation argument on  $(X, A)$  used to prove the first part of the theorem, hence  $Y^n(t)$  is predictable. Now, by  $d_\infty$ -right continuity of  $F$ ,  $Y^n(t) \rightarrow Y(t)$  almost surely, which proves that  $Y$  is predictable.  $\square$

### 3 Pathwise derivatives of non-anticipative functionals

#### 3.1 Horizontal and vertical derivatives

We now define pathwise derivatives for a functional  $F = (F_t)_{t \in [0, T[} \in \mathbb{F}^\infty$ , following an idea of Dupire [10].

**Definition 9** (Horizontal derivative). The *horizontal derivative* at  $(x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t$  of non-anticipative functional  $F = (F_t)_{t \in [0, T[}$  is defined as

$$\mathcal{D}_t F(x, v) = \lim_{h \rightarrow 0^+} \frac{F_{t+h}(x_{t,h}, v_{t,h}) - F_t(x, v)}{h} \quad (20)$$

if the corresponding limit exists. If (20) is defined for all  $(x, v) \in \Upsilon$  the map

$$\begin{aligned} \mathcal{D}_t F : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t &\mapsto \mathbb{R}^d \\ (x, v) &\mapsto \mathcal{D}_t F(x, v) \end{aligned} \quad (21)$$

defines a non-anticipative functional  $\mathcal{D}F = (\mathcal{D}_t F)_{t \in [0, T]}$ , the *horizontal derivative* of  $F$ .

$F$  is said to be *horizontally differentiable* if  $\mathcal{D}F$  is right-continuous i.e.  $\mathcal{D}F \in \mathbb{F}_r^\infty$ .

This pathwise derivative was introduced by B. Dupire [10] as a generalization of the time-derivative to path-dependent functionals, in the case where  $F(x, v) = G(x)$  is continuous in supremum norm. It can be seen as a ‘‘Lagrangian’’ derivative along the path  $x$ .

Dupire [10] also introduced a pathwise spatial derivative for such functionals, which we now introduce. Denote  $(e_i, i = 1..d)$  the canonical basis in  $\mathbb{R}^d$ .

**Definition 10.** A non-anticipative functional  $F = (F_t)_{t \in [0, T]}$  is said to be *vertically differentiable* at  $(x, v) \in D([0, t], \mathbb{R}^d) \times D([0, t], S)$  if

$$\begin{aligned} \mathbb{R}^d &\mapsto \mathbb{R} \\ e &\rightarrow F_t(x_t^e, v_t) \end{aligned}$$

is differentiable at 0. Its gradient at 0

$$\nabla_x F_t(x, v) = (\partial_i F_t(x, v), i = 1..d) \quad \text{where} \quad \partial_i F_t(x, v) = \lim_{h \rightarrow 0} \frac{F_t(x_t^{he_i}, v) - F_t(x, v)}{h} \quad (22)$$

is called the *vertical derivative* of  $F_t$  at  $(x, v)$ . If (22) is defined for all  $(x, v) \in \Upsilon$ , the maps

$$\begin{aligned} \nabla_x F : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t &\mapsto \mathbb{R}^d \\ (x, v) &\rightarrow \nabla_x F_t(x, v) \end{aligned} \quad (23)$$

define a non-anticipative functional  $\nabla_x F = (\nabla_x F_t)_{t \in [0, T]}$ , the *vertical derivative* of  $F$ .

$F$  is said to be *vertically differentiable* on  $\Upsilon$  if  $\nabla_x F \in \mathbb{F}_r^\infty$ .

*Remark 11.*  $\partial_i F_t(x, v)$  is simply the directional derivative of  $F_t$  in direction  $(1_{\{t\}}e_i, 0)$ . Note that this involves examining cadlag perturbations of the path  $x$ , even if  $x$  is continuous.

*Remark 12.* If  $F_t(x, v) = f(t, x(t))$  with  $f \in C^{1,1}([0, T] \times \mathbb{R}^d)$  then we retrieve the usual partial derivatives:

$$D_t F_t(x, v) = \partial_t f(t, X(t)) \quad \nabla_x F_t(X_t, A_t) = \nabla_x f(t, X(t)).$$

*Remark 13.* Bismut [3] considered directional derivatives of functionals on  $D([0, T], \mathbb{R}^d)$  in the direction of purely discontinuous (e.g. piecewise constant) functions with finite variation, which is similar to Def. 10. This notion, used in [3] to derive an integration by parts formula for pure-jump processes, seems natural in that context. We will show that the directional derivative (22) also intervenes naturally when the underlying process  $X$  is *continuous*, which is less obvious.

Note that, unlike the definition of a Fréchet derivative in which  $F$  is perturbed along all directions in  $C_0([0, T], \mathbb{R}^d)$  or the case of a Malliavin derivative [22, 23] in which  $F$  is perturbed along all Cameron-Martin (i.e. absolutely continuous) functions, we only examine *local* perturbations, so  $\nabla_x F$  and  $\mathcal{D}_t F$  seem to contain *less* information on the behavior of the functional  $F$ . Nevertheless we will show in the Section 4 that these derivatives are sufficient to reconstitute the path of  $Y(t) = F_t(X_t, A_t)$ : the pieces add up to the whole.

**Definition 14.** Define  $\mathbb{C}^{j,k}([0, T])$  as the set of functionals  $F \in \mathbb{F}_r^\infty$  which are differentiable  $j$  times horizontally and  $k$  time vertically at all  $(x, v) \in \mathcal{U}_t \times \mathcal{S}_t$ ,  $t < T$ , and the derivatives  $\mathcal{D}^m F, m \leq j, \nabla_x^n F, n \leq k$  define elements of  $\mathbb{F}_r^\infty$ .

Define  $\mathbb{C}_b^{j,k}([0, T])$  as the set of functionals  $F \in \mathbb{C}^{j,k}([0, T])$  such that the horizontal derivatives up to order  $j$  and vertical derivatives up to order  $k$  are in  $\mathbb{B}$ .

*Example 1* (Smooth functions). Let us start by noting that, in the case where  $F$  reduces to a smooth *function* of  $X(t)$ ,

$$F_t(x_t, v_t) = f(t, x(t)) \quad (24)$$

where  $f \in C^{j,k}([0, T] \times \mathbb{R}^d)$ , the pathwise derivatives reduces to the usual ones:  $F \in \mathbb{C}_b^{j,k}$  with:

$$\mathcal{D}_t^i F(x_t, v_t) = \partial_t^i f(t, x(t)) \quad \nabla_x^m F_t(x_t, v_t) = \partial_x^m f(t, x(t)) \quad (25)$$

In fact  $F \in \mathbb{C}^{j,k}$  simply requires  $f$  to be  $j$  times right-differentiable in time, and that right-derivatives in time and derivatives in space be jointly continuous in space and right-continuous in time.

*Example 2* (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

$$F_t(x_t, v_t) = \int_0^t g(x(u))v(u)du \quad (26)$$

It is readily observed that  $F \in \mathbb{C}_b^{1,\infty}$ , with:

$$\mathcal{D}_t F(x_t, v_t) = g(x(t))v(t) \quad \nabla_x^j F_t(x_t, v_t) = 0 \quad (27)$$

*Example 3.* The martingale  $Y(t) = X(t)^2 - [X](t)$  is represented by the functional

$$F_t(x_t, v_t) = x(t)^2 - \int_0^t v(u)du \quad (28)$$

Then  $F \in \mathbb{C}_b^{1,\infty}$  with:

$$\begin{aligned} \mathcal{D}_t F(x, v) &= -v(t) & \nabla_x F_t(x_t, v_t) &= 2x(t) \\ \nabla_x^2 F_t(x_t, v_t) &= 2 & \nabla_x^j F_t(x_t, v_t) &= 0, j \geq 3 \end{aligned} \quad (29)$$

*Example 4* (Doléans exponential). The exponential martingale  $Y = \exp(X - [X]/2)$  may be represented by the functional

$$F_t(x_t, v_t) = e^{x(t) - \frac{1}{2} \int_0^t v(u)du} \quad (30)$$

Elementary computations show that  $F \in \mathbb{C}_b^{1,\infty}$  with:

$$\mathcal{D}_t F(x, v) = -\frac{1}{2}v(t)F_t(x, v) \quad \nabla_x^j F_t(x_t, v_t) = F_t(x_t, v_t) \quad (31)$$

Note that, although  $A_t$  may be expressed as a functional of  $X_t$ , this functional is not continuous and without introducing the second variable  $v \in \mathcal{S}_t$ , it is not possible to represent Examples 2, 3 and 4 as a right-continuous functional of  $x$  alone.

### 3.2 Obstructions to regularity

It is instructive to observe what prevents a functional from being regular in the sense of Definition 14. The examples below illustrate the fundamental obstructions to regularity:

*Example 5* (Delayed functionals).  $F_t(x_t, v_t) = x(t - \epsilon)$  defines a  $\mathbb{C}_b^{0,\infty}$  functional. All vertical derivatives are 0. However, it fails to be horizontally differentiable.

*Example 6* (Jump of  $x$  at the current time).  $F_t(x_t, v_t) = x(t) - x(t-)$  defines a functional which is infinitely differentiable and has regular pathwise derivatives:

$$\mathcal{D}_t F(x_t, v_t) = 0 \quad \nabla_x F_t(x_t, v_t) = 1 \quad (32)$$

However, the functional itself fails to be  $F_r^\infty$ .

*Example 7* (Jump of  $x$  at a fixed time).  $F_t(x_t, v_t) = 1_{t \geq t_0}(x(t_0) - x(t_0-))$  defines a functional in  $\mathbb{F}^{\infty,1}$  which admits horizontal and vertical derivatives at any order at each point  $(x, v)$ . However,  $\nabla_x F_t(x_t, v_t) = 1_{t=t_0}$  fails to be right continuous so  $F$  is not vertically differentiable in the sense of Definition 10.

*Example 8* (Maximum).  $F_t(x_t, v_t) = \sup_{s \leq t} x(s)$  is  $\mathbb{F}^{\infty,1}$  but fails to be vertically differentiable on the set

$$\{(x_t, v_t) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \quad x(t) = \sup_{s \leq t} x(s)\}.$$

### 3.3 Pathwise derivatives of an adapted process

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

$$Y(t) = F_t(X_t, A_t) \quad (33)$$

where  $F \in \mathbb{F}^{\infty,1}$  has right-continuous horizontal and vertical derivatives  $\mathcal{D}_t F \in \mathbb{F}_r^\infty$  and  $\nabla_x F \in \mathbb{F}_r^\infty$ .

Since  $X$  has continuous paths,  $Y$  only depends on the restriction of  $F$  to  $\Upsilon_c = \bigcup_{t \in [0, T]} C([0, t], \mathbb{R}^d) \times \mathcal{S}_t$ . Therefore, the representation (33) of  $Y$  by  $F : \Upsilon \rightarrow \mathbb{R}$  in (33) is not unique, as the following example shows.

*Example 9* (Non-uniqueness of functional representation). Take  $d = 1$ . The quadratic variation process  $[X]$  may be represented by the following functionals:

$$\begin{aligned} F^0(x_t, v_t) &= \int_0^t v(u) du \\ F^1(x_t, v_t) &= \left( \lim_n \sum_{i=0}^{t2^n} \left| x\left(\frac{i+1}{2^n}\right) - x\left(\frac{i}{2^n}\right) \right|^2 \right) \mathbf{1}_{\lim_n \sum_{i \leq t2^n} (x(\frac{i+1}{2^n}) - x(\frac{i}{2^n}))^2 < \infty} \\ F^2(x_t, v_t) &= \left( \lim_n \sum_{i=0}^{t2^n} \left| x\left(\frac{i+1}{2^n}\right) - x\left(\frac{i}{2^n}\right) \right|^2 - \sum_{0 \leq s < t} |\Delta x(s)|^2 \right) \mathbf{1}_{\lim_n \sum_{i=0}^{t2^n} |x(\frac{i+1}{2^n}) - x(\frac{i}{2^n})|^2 < \infty} \mathbf{1}_{\sum_{s < t} |\Delta x(s)|^2 < \infty} \end{aligned}$$

where  $\Delta x(t) = x(t) - x(t-)$  denotes the discontinuity of  $x$  at  $t$ . If  $X$  is a continuous semimartingale, then

$$F_t^0(X_t, A_t) = F_t^1(X_t, A_t) = F_t^2(X_t, A_t) = [X](t)$$

Yet  $F^0 \in \mathbb{C}_b^{1,2}$  but  $F^1, F^2$  are not even right-continuous:  $F^i \notin \mathbb{F}_r^\infty$  for  $i = 1, 2$ .

However, the definition of  $\nabla_x F$  (Definition 10), which involves evaluating  $F$  on cadlag paths, seems to depend on the choice of the representation, in particular on the values taken by  $F$  outside  $\Upsilon_c$ . This non-uniqueness, not addressed in [10], must be resolved before one can define the pathwise derivative of a process in an intrinsic manner.

The following key result shows that, if  $Y$  has a functional representation (33) where  $F$  is differentiable in the sense of Defs. 9 and 10 and the derivatives define elements of  $\mathbb{F}_r^\infty$ , then  $\nabla_x F_t(X_t, A_t)$  is uniquely defined, independently of the choice of the representation  $F$ :

**Theorem 15.** *If  $F^1, F^2 \in \mathbb{C}^{1,1}([0, T]) \cap \mathbb{F}^\infty$  coincide on continuous paths:*

$$\begin{aligned} \forall t < T, \quad \forall (x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \quad F_t^1(x_t, v_t) = F_t^2(x_t, v_t) \\ \text{then} \quad \forall t < T, \quad \forall (x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \quad \nabla_x F_t^1(x_t, v_t) = \nabla_x F_t^2(x_t, v_t) \end{aligned}$$

*Proof.* Let  $F = F^1 - F^2 \in \mathbb{F}^\infty([0, T])$  and  $(x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$ . Then  $F_t(x, v) = 0$  for all  $t \leq T$ . It is then obvious that  $\mathcal{D}_t F(x, v)$  is also 0 on continuous paths because the extension  $(x_{t,h})$  of  $x_t$  is itself a continuous path. Assume now that there exists some  $(x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$  such that for some  $1 \leq i \leq d$  and  $t \in [0, T[$ ,  $\partial_i F_t(x, v) > 0$ . Define the following extension of  $x_t$  to  $[0, T]$ :

$$\begin{aligned} z(u) &= x(u), u \leq t \\ z_j(u) &= x_j(t) + 1_{i=j}(u-t), t \leq u \leq T, 1 \leq j \leq d \end{aligned} \quad (34)$$

Let  $\alpha = \frac{1}{2} \partial_i F_t(x, v)$ . By the right-continuity of  $\partial_i F$  and  $\mathcal{D}_t F$  at  $(x, v)$ , we may choose  $h < T - t$  sufficiently small such that, for any  $t' \in [t, T[$ , for any  $(x', v') \in \mathcal{U}_{t'} \times \mathcal{S}_{t'}$ ,

$$d_\infty((x, v), (x', v')) < h \Rightarrow \partial_i F_{t'}(x', v') > \alpha \text{ and } |\mathcal{D}_t F(x', v')| < 1 \quad (35)$$

Define the following sequence of piecewise constant approximations of  $z_{t+h}$ :

$$\begin{aligned} z^n(u) &= \tilde{z}^n = z(u), u \leq t \\ z_j^n(u) &= x_j(t) + 1_{i=j} \frac{h}{n} \sum_{k=1}^n 1_{\frac{kh}{n} \leq u-t}, t \leq u \leq t+h, 1 \leq j \leq d \end{aligned} \quad (36)$$

Since  $d_\infty((z, v_{t,h}), (z^n, v_{t,h})) = \frac{h}{n} \rightarrow 0$ ,

$$|F_{t+h}(z, v_{t,h}) - F_{t+h}(z^n, v_{t,h})| \xrightarrow{n \rightarrow +\infty} 0$$

We can now decompose  $F_{t+h}(z^n, v_{t,h}) - F_t(x, v)$  as

$$\begin{aligned} F_{t+h}(z^n, v_{t,h}) - F_t(x, v) &= \sum_{k=1}^n F_{t+\frac{kh}{n}}(z_{t+\frac{kh}{n}}^n, v_{t, \frac{kh}{n}}) - F_{t+\frac{kh}{n}}(z_{t+\frac{kh}{n}-}^n, v_{t, \frac{kh}{n}}) \\ &+ \sum_{k=1}^n F_{t+\frac{kh}{n}}(z_{t+\frac{kh}{n}-}^n, v_{t, \frac{kh}{n}}) - F_{t+\frac{(k-1)h}{n}}(z_{t+\frac{(k-1)h}{n}}^n, v_{t, \frac{(k-1)h}{n}}) \end{aligned} \quad (37)$$

where the first sum corresponds to jumps of  $z^n$  at times  $t + \frac{kh}{n}$  and the second sum to its extension by a constant on  $[t + \frac{(k-1)h}{n}, t + \frac{kh}{n}]$ .

$$F_{t+\frac{kh}{n}}(z_{t+\frac{kh}{n}}^n, v_{t, \frac{kh}{n}}) - F_{t+\frac{kh}{n}}(z_{t+\frac{kh}{n}-}^n, v_{t, \frac{kh}{n}}) = \phi\left(\frac{h}{n}\right) - \phi(0) \quad (38)$$

where  $\phi$  is defined as

$$\phi(u) = F_{t+\frac{kh}{n}}((z^n)_{t+\frac{kh}{n}-}^{ue_i}, v_{t,\frac{kh}{n}})$$

Since  $F$  is vertically differentiable,  $\phi$  is differentiable and

$$\phi'(u) = \partial_i F_{t+\frac{kh}{n}}((z^n)_{t+\frac{kh}{n}-}^{ue_i}, v_{t,\frac{kh}{n}})$$

is right-continuous. Since

$$d_\infty((x, v), ((z^n)_{t+\frac{kh}{n}-}^{ue_i}, v_{t,\frac{kh}{n}})) \leq h,$$

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

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

On the other hand

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

where

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

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

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

Since  $F \in \mathbb{F}^\infty([0, T])$ ,  $\psi$  is also continuous by theorem 7 so

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

Noting that:

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

we obtain that:

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

since the path of  $z_{t+u}$  is continuous. Moreover  $|\mathcal{D}_t F_{t+u}(z_{t+u}^n, v_{t,u})| \leq 1$  since  $d_\infty((z_{t+u}^n, v_{t,u}), (x, v)) \leq h$ , so by dominated convergence the integral goes to 0 as  $n \rightarrow \infty$ . Writing:

$$F_{t+h}(z, v_{t,h}) - F_t(x, v) = [F_{t+h}(z, v_{t,h}) - F_{t+h}(z^n, v_{t,h})] + [F_{t+h}(z^n, v_{t,h}) - F_t(x, v)]$$

and taking the limit on  $n \rightarrow \infty$  leads to  $F_{t+h}(z, v_{t,h}) - F_t(x, v) \geq \alpha h$ , a contradiction.  $\square$

The above result implies in particular that, if  $\nabla_x F^i \in \mathbb{C}^{1,1}([0, T])$ , and  $F^1(x, v) = F^2(x, v)$  for any continuous path  $x$ , then  $\nabla_x^2 F^1$  and  $\nabla_x^2 F^2$  must also coincide on continuous paths.

We now show that the same result can be obtained under the weaker assumption that  $F^i \in \mathbb{C}^{1,2}([0, T])$ , using a probabilistic argument. Interestingly, while the previous result on the uniqueness of the first vertical derivative is based on the fundamental theorem of calculus, the proof of the following theorem is based on its stochastic equivalent, the Itô formula [15, 16].

**Theorem 16.** If  $F^1, F^2 \in C^{1,2}([0, T]) \cap \mathbb{F}^\infty$  coincide on continuous paths:

$$\forall (x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \quad \forall t \in [0, T[, \quad F_t^1(x_t, v_t) = F_t^2(x, v) \quad (39)$$

then their second vertical derivatives also coincide on continuous paths:

$$\forall (x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \quad \forall t \in [0, T[, \quad \nabla_x^2 F_t^1(x_t, v_t) = \nabla_x^2 F_t^2(x_t, v_t)$$

*Proof.* Let  $F = F^1 - F^2$ . Assume that there exists some  $(x, v) \in D([0, T], \mathbb{R}^d) \times \mathcal{S}_T$  such that for some  $t < T$ , and some direction  $h \in \mathbb{R}^d, \|h\| = 1, {}^t h \nabla_x^2 F_t(x_t, v_t) \cdot h > 0$ , and denote  $\alpha = \frac{1}{2} {}^t h \nabla_x^2 F_t(x_t, v_t) \cdot h$ . We will show that this leads to a contradiction. We already know that  $\nabla_x F_t(x_t, v_t) = 0$  by theorem 15. Let  $\epsilon > 0$  be small enough so that:

$$\begin{aligned} & \forall t' > t, \forall (x', v') \in \mathcal{U}_{t'} \times \mathcal{S}_{t'}, d_\infty((x_t, v_t), (x', v')) < \epsilon \\ \Rightarrow & |F_{t'}(x', v')| < |F_t(x_t, v_t)| + 1, |\nabla_x F_{t'}(x', v')| < 1, {}^t h \nabla_x^2 F_{t'}(x', v') \cdot h > \alpha \end{aligned} \quad (40)$$

Let  $W$  be a one dimensional Brownian motion on some probability space  $(\tilde{\Omega}, \mathcal{B}, \mathbb{P})$ ,  $(\mathcal{B}_s)$  its natural filtration, and let

$$\tau = \inf\{s > 0, \quad |W(s)| = \frac{\epsilon}{2}\} \quad (41)$$

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

$$U(t') = x(t')1_{t' \leq t} + (x(t) + W((t' - t) \wedge \tau)h)1_{t' > t} \quad (42)$$

and notice that for all  $s < \frac{\epsilon}{2}$ ,

$$d_\infty((U_{t+s}, v_{t,s}), (x_t, v_t)) < \epsilon \quad (43)$$

Define the following piecewise constant approximations of the stopped process  $W^\tau$ :

$$W^n(s) = \sum_{i=0}^{n-1} W(i \frac{\epsilon}{2n} \wedge \tau) 1_{s \in [i \frac{\epsilon}{2n}, (i+1) \frac{\epsilon}{2n})} + W(\frac{\epsilon}{2} \wedge \tau) 1_{s = \frac{\epsilon}{2}}, 0 \leq s \leq \frac{\epsilon}{2n} \quad (44)$$

Denoting

$$Z(s) = F_{t+s}(U_{t+s}, v_{t,s}), \quad s \in [0, T - t] \quad (45)$$

$$U^n(t') = x(t')1_{t' \leq t} + (x(t) + W^n((t' - t) \wedge \tau)h)1_{t' > t} \quad Z^n(s) = F_{t+s}(U_{t+s}^n, v_{t,s}) \quad (46)$$

we have the following decomposition:

$$\begin{aligned} Z(\frac{\epsilon}{2}) - Z(0) &= Z(\frac{\epsilon}{2}) - Z^n(\frac{\epsilon}{2}) + \sum_{i=1}^n Z^n(i \frac{\epsilon}{2n}) - Z^n(i \frac{\epsilon}{2n} -) \\ &+ \sum_{i=0}^{n-1} Z^n((i+1) \frac{\epsilon}{2n} -) - Z^n(i \frac{\epsilon}{2n}) \end{aligned} \quad (47)$$



The first term in (47) goes to 0 almost surely since

$$d_\infty((U_{t+\frac{\epsilon}{2}}, v_{t,\frac{\epsilon}{2}}), (U_{t+\frac{\epsilon}{2}}^n, v_{t,\frac{\epsilon}{2}}^n)) \xrightarrow{n \rightarrow \infty} 0. \quad (48)$$

The second term in (47) may be expressed as

$$Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-) = \phi_i(W(i\frac{\epsilon}{2n}) - W((i-1)\frac{\epsilon}{2n})) - \phi_i(0) \quad (49)$$

where:

$$\phi_i(u, \omega) = F_{t+i\frac{\epsilon}{2n}}(U_{t+i\frac{\epsilon}{2n}-}^{n,uh}(\omega), v_{t,i\frac{\epsilon}{2n}})$$

Note that  $\phi_i(u, \omega)$  is measurable with respect to  $\mathcal{B}_{(i-1)\epsilon/2n}$  whereas its argument in (49) is independent with respect to  $\mathcal{B}_{(i-1)\epsilon/2n}$ . Let  $\Omega_1 = \{\omega \in \tilde{\Omega}, t \mapsto W(t, \omega) \text{ continuous}\}$ . Then  $\mathbb{P}(\Omega_1) = 1$  and for any  $\omega \in \Omega_1$ ,  $\phi_i(\cdot, \omega)$  is  $C^2$  with:

$$\begin{aligned} \phi_i'(u, \omega) &= \nabla_x F_{t+i\frac{\epsilon}{2n}}(U_{t+i\frac{\epsilon}{2n}-}^{n,uh}(\omega), v_{t,i\frac{\epsilon}{2n}})h \\ \phi_i''(u, \omega) &= {}^t h \nabla_x^2 F_{t+i\frac{\epsilon}{2n}}(U_{t+i\frac{\epsilon}{2n}-}^{n,uh}(\omega), v_{t,i\frac{\epsilon}{2n}}).h \end{aligned} \quad (50)$$

So, using the above arguments we can apply the Ito formula to (49) for each  $\omega \in \Omega_1$ . We therefore obtain, summing on  $i$  and denoting  $i(s)$  the index such that  $s \in [(i-1)\frac{\epsilon}{2n}, i\frac{\epsilon}{2n})$ :

$$\begin{aligned} \sum_{i=1}^n Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-) &= \int_0^{\frac{\epsilon}{2}} \nabla_x F_{t+i(s)\frac{\epsilon}{2n}}(U_{t+i(s)\frac{\epsilon}{2n}-}^{n,uh}, v_{t,i(s)\frac{\epsilon}{2n}})hdW(s) \\ &\quad + \int_0^{\frac{\epsilon}{2}} {}^t h \cdot \nabla_x^2 F_{t+i(s)\frac{\epsilon}{2n}}(U_{t+i(s)\frac{\epsilon}{2n}-}^{n,uh}, v_{t,i(s)\frac{\epsilon}{2n}}).h ds \end{aligned} \quad (51)$$

Since the first derivative is bounded by (40), the stochastic integral is a martingale, so taking expectation leads to:

$$E[\sum_{i=1}^n Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-)] > \alpha \frac{\epsilon}{2} \quad (52)$$

Now

$$Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) = \psi(\frac{\epsilon}{2n}) - \psi(0) \quad (53)$$

where

$$\psi(u) = F_{t+(i-1)\frac{\epsilon}{2n}+u}(U_{t+(i-1)\frac{\epsilon}{2n}+u}^n, v_{t,(i-1)\frac{\epsilon}{2n}+u}) \quad (54)$$

is right-differentiable with right derivative:

$$\psi'(u) = \mathcal{D}_t F_{t+(i-1)\frac{\epsilon}{2n}+u}(U_{t+(i-1)\frac{\epsilon}{2n}+u}^n, v_{t,(i-1)\frac{\epsilon}{2n}+u}) \quad (55)$$

Since  $F \in \mathbb{F}^\infty([0, T])$ ,  $\psi$  is continuous by theorem 8 and the fundamental theorem of calculus yields:

$$\sum_{i=0}^{n-1} Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) = \int_0^{\frac{\epsilon}{2}} \mathcal{D}_t F_{t+s}(U_{t+(i(s)-1)\frac{\epsilon}{2n}+u}^n, v_{t,s})ds \quad (56)$$

The integrand converges to  $\mathcal{D}_t F_{t+s}(U_{t+(i(s)-1)\frac{\epsilon}{2n}+u}, v_{t,s}) = 0$  since  $\mathcal{D}_t F$  is zero whenever the first argument is a continuous path. Since this term is also bounded, by dominated convergence the integral converges almost surely to 0.

It is obvious that  $Z(\frac{\epsilon}{2}) = 0$  since  $F(x, v) = 0$  whenever  $x$  is a continuous path. On the other hand, since all derivatives of  $F$  appearing in (47) are bounded, the dominated convergence theorem allows to take expectations of both sides in (47) with respect to the Wiener measure and obtain  $\alpha\frac{\epsilon}{2} = 0$ , a contradiction.  $\square$

Using Theorems 15 and 16, we can now define the horizontal and vertical derivatives for an  $\mathcal{F}_t$ -adapted process  $Y$  which admits a  $\mathbb{C}^{1,2}$ -representation, i.e. extending the pathwise derivatives introduced in Definitions 9–10 to functionals which are defined *almost-surely*.

Theorems 15 and 16 guarantee that the derivatives of  $Y$  are independent of the choice of the functional representation in (4):

**Definition 17** (Horizontal and vertical derivative of a process). Define  $\mathcal{C}^{1,2}(X)$  the set of  $\mathcal{F}_t$ -adapted processes  $Y$  which admit a  $\mathbb{C}^{1,2}$ -representation:

$$\mathcal{C}^{1,2}(X) = \{Y, \exists F \in \mathbb{C}^{1,2}([0, T]) \cap \mathbb{F}^{\infty,1}, Y(t) = F_t(X_t, A_t) \text{ } \mathbb{P} - \text{ a.s.}\} \quad (57)$$

For  $Y \in \mathcal{C}^{1,2}(X)$  the following right-continuous non-anticipative processes:

$$\mathcal{D}Y(t) = \mathcal{D}_t F(X_t, A_t) \quad \nabla_X Y(t) = \nabla_x F_t(X_t, A_t) \quad \nabla_X^2 Y(t) = \nabla_x^2 F_t(X_t, A_t) \quad (58)$$

are uniquely defined up to an evanescent set, independently of the choice of the functional representation  $F \in \mathbb{C}^{1,2}([0, T]) \cap \mathbb{F}^{\infty,1}$ .

We will call  $\mathcal{D}Y$  the *horizontal derivative* of  $Y$  and  $\nabla_X Y$  the *vertical derivative* of  $Y$  with respect to  $X$ .

Similarly, we will denote  $\mathcal{C}_b^{1,2}(X)$  the set of processes  $Y \in \mathcal{C}^{1,2}(X)$  which admit a representation  $Y(t) = F_t(X_t, A_t)$  with  $F \in \mathbb{C}_b^{1,2}([0, T]) \cap \mathbb{F}^{\infty,1}$ .

The operators

$$\mathcal{D} : \mathcal{C}^{1,2}(X) \mapsto \mathcal{C}(X) \quad (59)$$

$$\text{and } \nabla_X : \mathcal{C}^{1,2}(X) \mapsto \mathcal{C}(X) \quad (60)$$

map a process  $Y \in \mathcal{C}^{1,2}(X)$  into an optional process belonging

$$\mathcal{C}(X) = \{Y, \exists F \in \mathbb{F}_r^\infty, Y(t) = F_t(X_t, A_t) \text{ } \mathbb{P} - \text{ a.s.}\}, \quad (61)$$

the set of non-anticipative process with right-continuous path-dependence.

## 4 Functional Ito formula

We are now ready to state a functional change of variable formula which extends the Ito formula to path-dependent functionals of a semimartingale:

**Theorem 18** (Functional Ito formula). *Let  $Y \in \mathcal{C}_b^{1,2}(X)$ . For any  $t \in [0, T[$ ,*

$$Y(t) - Y(0) = \int_0^t \mathcal{D}Y(u)du + \int_0^t \frac{1}{2} \text{tr}[{}^t\nabla_X^2 Y(u) d[X](u)] + \int_0^t \nabla_X Y(u).dX(u) \quad a.s. \quad (62)$$

*In particular, for any  $F \in \mathbb{C}_b^{1,2}([0, T]) \cap \mathbb{F}^{\infty,1}([0, T])$ ,  $Y(t) = F_t(X_t, A_t)$  is a semimartingale.*

We note that:

- Note that the dependence of  $F$  on the second variable  $A$  does not enter the formula (62). Indeed, under our regularity assumptions, variations in  $A$  lead to “higher order” terms which do not contribute.
- As expected, in the case where  $X$  is continuous then  $Y$  depends on  $F$  and its derivatives only via their values on continuous paths. More precisely,  $Y$  can be reconstructed from the second-order jet of  $F$  on  $\mathcal{C} = \bigcup_{t \in [0, T[} C_0([0, t], \mathbb{R}^d) \times D([0, t], \mathcal{S}_d^+) \subset \mathfrak{Y}$ .

The basic idea of the proof, as in the classical derivation of the Ito formula [8, 24, 28], is to approximate the path of  $X$  using piecewise constant predictable processes along a subdivision of  $[0, T]$ . A crucial remark, due to Dupire [10], is that the variations of a functional along a piecewise constant path may be decomposed into successive “horizontal” and “vertical” increments, involving only the partial functions used in the definitions of the pathwise derivatives (Definitions 9 and 10). This allows to express the functional  $F$  along a piecewise constant path in the form (62). The last step is to take limits along a sequence of piecewise constant approximations of  $X$ , using the continuity properties of the pathwise derivatives. The control of the remainder terms is somewhat more involved than in the usual proof of the Ito formula given that we are dealing with functionals.

We give here the proof in the case where  $A$  is continuous. The general case where  $A$  is allowed to be discontinuous (cadlag) is treated in Appendix A.2.

*Continuous case.* Since  $Y \in \mathcal{C}_b^{1,2}(X)$ , Theorem 8 implies that all the integrands in (62) are predictable processes.

Let us first assume that  $X$  does takes values in a compact set  $K$  and that  $\|A\|_\infty \leq R$  for some  $R > 0$ . Then the integrands in (62) are a.s. bounded; in particular the stochastic integral term is well-defined.

Let  $\pi_n = (t_i^n, i = 0..2^n)$  be the dyadic subdivision of  $[0, T]$ , ie  $t_i^n = t \frac{i}{2^n}$ . The following arguments apply pathwise. Using the uniform continuity of  $X$  and  $A$  on  $[0, t]$ ,

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

Let  $\eta > 0, C > 0$  be such that, for any  $s < T$ , for any  $(x, v) \in D([0, s], \mathbb{R}^d) \times \mathcal{S}_s^+$ ,  $d_\infty((X_s, A_s), (x, v)) < \eta \Rightarrow |F_s(x, A_s) - F_s(x, v_s)| \leq C\|A_s - v_s\|_1$ , and we will assume  $n$  large enough so that  $\eta_n < \eta$ .

Denoting  ${}_n X = \sum_{i=0}^{2^n-1} X(t_i^n)1_{[t_i^n, t_{i+1}^n)} + X(t)1_{\{t\}}$  the cadlag piecewise constant approximation of  $X_t$  along  $\pi_n$ ,

$$F_t(X_t, A_t) - F_0(X_0, A_0) = F_t(X_t, A_t) - F_t({}_n X_t, A_t) + \sum_{i=0}^{k_n-1} F_{t_{i+1}^n}({}_n X_{t_{i+1}^n}, A_{t_{i+1}^n}) - F_{t_i^n}({}_n X_{t_i^n}, A_{t_i^n}) \quad (63)$$

First, note that  $|F(X_t, A_t) - F({}_n X_t, A_t)| \rightarrow 0$  as  $n \rightarrow \infty$ . Denote  $\delta_i = X_{t_{i+1}^n} - X_{t_i^n}$  and  $h_i = t_{i+1}^n - t_i^n$ . Each term in the sum can then be decomposed as

$$\begin{aligned} & [F_{t_{i+1}^n}({}_n X_{t_{i+1}^n}, A_{t_{i+1}^n}) - F_{t_{i+1}^n}({}_n X_{t_{i+1}^n}, A_{t_i^n}, h_i)] + [F_{t_{i+1}^n}({}_n X_{t_{i+1}^n}, A_{t_i^n}, h_i) - F_{t_{i+1}^n}({}_n X_{t_i^n}, h_i, A_{t_i^n}, h_i)] \\ & \quad + [F_{t_{i+1}^n}({}_n X_{t_i^n}, h_i, A_{t_i^n}, h_i) - F_{t_i^n}({}_n X_{t_i^n}, A_{t_i^n})] \end{aligned} \quad (64)$$

The first term in (64) is bounded by

$$C \|A_{t_{i+1}^n} - A_{t_i^n, h_i}\|_1 = C \int_{t_i^n}^{t_{i+1}^n} |A(s) - A(t_i^n)| ds \leq C |t_{i+1}^n - t_i^n| \eta_n.$$

Summing over  $i$  leads to a term which is bounded by  $C t \eta_n$ , hence converging to 0 as  $n \rightarrow \infty$ .

Denote by  ${}_n Y_{t_{i+1}^n} = {}_n X_{t_i^n, h_i}$  the horizontal extension of  ${}_n X_{t_i}$  to  $[t_i^n, t_{i+1}^n]$ . Since  ${}_n X$  is piecewise constant,  ${}_n Y_{t_{i+1}^n}^{\delta_i} = {}_n X_{t_{i+1}^n}$  so the second term in (64) can be written  $\phi(X(t_{i+1}^n) - X(t_i^n)) - \phi(0)$  where

$$\phi(u) = F_{t_{i+1}^n}({}_n Y_{t_{i+1}^n}^u, A_{t_i^n}, h_i) \quad (65)$$

Since  $F \in \mathbb{C}^{1,2}$ , this implies that  $\phi$  is  $C^2$  and

$$\phi'(u) = \nabla_x F_{t_{i+1}^n}({}_n Y_{t_{i+1}^n}^u, A_{t_i^n}, h_i) \quad \phi''(u) = \nabla_x^2 F_{t_{i+1}^n}({}_n Y_{t_{i+1}^n}^u, A_{t_i^n}, h_i) \quad (66)$$

Applying the Ito formula to  $\phi$  then allows to rewrite the second term in (64) as

$$\begin{aligned} \phi(X(t_{i+1}^n) - X(t_i^n)) - \phi(0) &= \int_{t_i^n}^{t_{i+1}^n} \nabla_x F_{t_{i+1}^n}({}_n Y_{t_{i+1}^n}^{X(s) - X(t_i^n)}, A_{t_i^n}, h_i) dX(s) \\ & \quad + \frac{1}{2} \int_{t_i^n}^{t_{i+1}^n} \text{tr} \left[ \frac{1}{2} {}^t \nabla_x^2 F_{t_{i+1}^n}({}_n Y_{t_{i+1}^n}^{X(s) - X(t_i^n)}, A_{t_i^n}, h_i) d[X](s) \right] \end{aligned}$$

The third term in (64) can be expressed as  $\psi(t_{i+1}^n - t_i) - \psi(0)$  where  $\psi(h) = F_{t_{i+1}^n}({}_n X_{t_i^n, h}, A_{t_i^n}, h)$ . By lemma 7,  $\psi$  is continuous and right-differentiable with  $\psi'(h) = \mathcal{D}_{t_{i+1}^n} F({}_n X_{t_i^n}, h, A_{t_i^n}, h)$  so

$$F_{t_{i+1}^n}({}_n X_{t_i^n}, h_i, A_{t_i^n}, h_i) - F_{t_i^n}({}_n X_{t_i^n}, A_{t_i^n}) = \int_{t_i^n}^{t_{i+1}^n} \mathcal{D}_s F({}_n X_{t_i^n, s-t_i}, A_{t_i^n, s-t_i}) ds \quad (67)$$

Summing over  $i = 1..2^n$  and denoting  $i(s)$  the index such that  $s \in [t_{i(s)}^n, t_{i(s)+1}^n)$ , we have shown:

$$\begin{aligned} F_t(X_t, A_t) - F_0(X_0, A_0) &= \int_0^t \mathcal{D}_s F({}_n X_{t_{i(s)}^n, s-t_{i(s)}^n}, A_{t_{i(s)}^n, s-t_{i(s)}^n}) ds \\ & \quad + \int_0^t \nabla_x F_{t_{i(s)+1}^n}({}_n Y_{t_{i(s)+1}^n}^{X(s) - X(t_{i(s)}^n)}, A_{t_{i(s)}^n}, h_{i(s)}) dX(s) \\ & \quad + \frac{1}{2} \int_0^t \text{tr} \left( {}^t \nabla_x^2 F_{t_{i(s)+1}^n}({}_n Y_{t_{i(s)+1}^n}^{X(s) - X(t_{i(s)}^n)}, A_{t_{i(s)}^n}, h_{i(s)}) d[X] \right) + r(\pi_n) \end{aligned}$$

where  $r(\pi_n) \rightarrow 0$  as  $n \rightarrow \infty$ . The  $d_\infty$ -distance to  $(X_s, A_s)$  of all terms appearing in the various integrals is less than  $\eta_n$ , hence they converge respectively to  $\mathcal{D}_s F(X_s, A_s)$ ,  $\nabla_x F_s(X_s, A_s)$ , and

$\nabla_x^2 F_s(X_s, A_s)$  as  $n \rightarrow \infty$  by  $d_\infty$  right-continuity. Since the derivatives are in  $\mathbb{B}$  the integrands in the various above integrals are bounded by a constant dependant only on  $F, K$  and  $R$  and  $t$  hence non-dependant on  $s$  nor on  $\omega$ , hence the dominated convergence theorem and the dominated convergence theorem for the stochastic integrals [28, Ch.IV Theorem 32] ensure that the integrals above converge in probability, uniformly on  $[0, t_0]$ , for any  $t_0 < T$  to the corresponding terms appearing in (62) 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

$$\tau_n = \inf\{s < t | X_s \notin K^n \text{ or } |A_s| > n\} \wedge t$$

which are optional times. Applying the previous result to the stopped process  $(X_{t \wedge \tau_n}, A_{t \wedge \tau_n})$  leads to:

$$\begin{aligned} F_t(X_{t \wedge \tau_n}, A_{t \wedge \tau_n}) - Y(0) &= \int_0^{t \wedge \tau_n} \mathcal{D}Y(u) du + \frac{1}{2} \int_0^{t \wedge \tau_n} \text{tr}({}^t \nabla_X^2 F_u(X_u, A_u) d[X](u)) \\ &+ \int_0^{t \wedge \tau_n} \nabla_X Y \cdot dX + \int_{t \wedge \tau_n}^t D_t F(X_{u \wedge \tau_n}, A_{u \wedge \tau_n}) du \end{aligned} \quad (68)$$

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.  $\square$

*Remark 19.* The above proof is probabilistic and makes use of the Ito formula (for functions of semimartingales). In the companion paper [5] we give a non-probabilistic proof of Theorem 18, which allows  $X$  to have discontinuous (cadlag) trajectories using the analytical approach of Föllmer [12].

An immediate corollary of Theorem 18 is that any regular functional of a local martingale which has finite variation is equal to the integral of its horizontal derivative:

**Corollary 20.** *If  $X$  is a local martingale and  $Y \in \mathcal{C}_b^{1,2}(X)$  is a process with finite variation then  $\nabla_X Y(t) = 0$   $d[X] \times d\mathbb{P}$ -almost everywhere and*

$$Y(t) = \int_0^t \mathcal{D}Y(u) \, du$$

*Proof.*  $Y \in \mathcal{C}_b^{1,2}(X)$  is a continuous semimartingale by Theorem 18, with canonical decomposition given by (62). If  $Y$  has finite variation, then by formula (62), its continuous martingale component should be zero i.e.  $\int_0^t \nabla_X Y \cdot dX = 0$  a.s. Computing the quadratic variation of this martingale we obtain

$$\int_0^T \text{tr}({}^t \nabla_X Y \cdot \nabla_X Y \cdot d[X]) = 0$$

which implies in particular that  $\|\nabla_X Y^i\|^2 = 0$   $d[X^i] \times d\mathbb{P}$ -almost everywhere for  $i = 1..d$ . Thus,  $\nabla_X Y(t, \omega) = 0$  for  $(t, \omega) \notin A \subset [0, T] \times \Omega$  where  $[X^i] \times \mathbb{P}(A) = 0$  for  $i = 1..d$ . From (the locality of) Definition 10 we deduce that  $\nabla_X^2 Y(t, \omega) = 0$  for  $(t, \omega) \notin A$ . In particular  $\int_0^t \text{tr}(\nabla_X^2 Y \cdot d[X]) = 0$  which entails the result.  $\square$

*Example 10.* If  $F_t(x_t, v_t) = f(t, x(t))$  where  $f \in C^{1,2}([0, T] \times \mathbb{R}^d)$ , (62) reduces to the standard Itô formula.

*Example 11.* For integral functionals of the form

$$F_t(x_t, v_t) = \int_0^t g(x(u))v(u)du \quad (69)$$

where  $g \in C_0(\mathbb{R}^d)$ , the Ito formula reduces to the trivial relation

$$F_t(X_t, A_t) = \int_0^t g(X(u))A(u)du \quad (70)$$

since the vertical derivatives are zero in this case.

*Example 12.* For a scalar semimartingale  $X$ , applying the formula to  $F_t(x_t, v_t) = x(t)^2 - \int_0^t v(u)du$  yields the well-known Ito product formula:

$$X(t)^2 - [X](t) = \int_0^t 2X.dX \quad (71)$$

*Example 13.* For the Doléans functional (Ex. 4)

$$F_t(x_t, v_t) = e^{x(t) - \frac{1}{2} \int_0^t v(u)du} \quad (72)$$

the formula (62) yields the well-known integral representation

$$\exp(X(t) - \frac{1}{2}[X](t)) = \int_0^t e^{X(u) - \frac{1}{2}[X](u)} dX(u) \quad (73)$$

## 5 Martingale representation formula

We consider now the case where the process  $X$  is a continuous martingale. We will show that, in this case, the functional Ito formula (Theorem (18)) leads to an explicit martingale representation formula for  $\mathcal{F}_t$ -martingales in  $\mathcal{C}^{1,2}(X)$ . This result may be seen as a non-anticipative counterpart of the Clark-Haussmann-Ocone formula [4, 26, 14] and generalizes explicit martingale representation formulas previously obtained in a Markovian context by Elliott and Kohlmann [11] and Jacod et al. [17].

### 5.1 Martingale representation theorem

Consider an  $\mathcal{F}_T$  measurable random variable  $H$  with  $E|H| < \infty$  and consider the martingale  $Y(t) = E[H|\mathcal{F}_t]$ . If  $Y \in \mathcal{C}_b^{1,2}(X)$ , we obtain the following martingale representation:

**Theorem 21.** *If  $Y \in \mathcal{C}^{1,2}(X)$  then*

$$Y(T) = E[Y(T)] + \int_0^T \nabla_X Y(t) dX(t) \quad (74)$$

Note that regularity assumptions are given not on  $H = Y(T)$  but on the functionals  $Y(t) = E[H|\mathcal{F}_t]$ , which is typically more regular than  $H$  itself.

*Proof.* Theorem 18 implies that for  $t \in [0, T[$ :

$$Y(t) = \left[ \int_0^t D_u F(X_u, A_u) du + \frac{1}{2} \int_0^t \text{tr} [{}^t \nabla_x^2 F_u(X_u, A_u) d[X](u)] + \int_0^t \nabla_x F_u(X_u, A_u) dX(u) \right] \quad (75)$$

Given the regularity assumptions on  $F$ , the first term in this sum is a finite variation process while the second is a local martingale. However,  $Y$  is a martingale and the decomposition of a semimartingale as sum of finite variation process and local martingale is unique. Hence the first term is 0 and:  $Y(t) = \int_0^t F_u(X_u, A_u) dX_u$ . Since  $F \in \mathbb{F}^{\infty,1}$   $Y(t)$  has limit  $F_T(X_T, A_T)$  as  $t \rightarrow T$ , the stochastic integral also converges, which concludes the proof.  $\square$

*Example 14.*

If the Doleans-Dade exponential  $e^{X(t) - \frac{1}{2}[X](t)}$  is a martingale, applying Theorem 21 to the functional  $F_t(x_t, v_t) = e^{x(t) - \int_0^t v(u) du}$  yields the familiar formula:

$$e^{X(t) - \frac{1}{2}[X](t)} = 1 + \int_0^t e^{X(s) - \frac{1}{2}[X](s)} dX(s) \quad (76)$$

If  $X(t)^2$  is integrable, applying Theorem 21 to the functional  $F_t(x(t), v(t)) = x(t)^2 - \int_0^t v(u) du$ , we obtain the well-known Ito product formula

$$X(t)^2 - [X](t) = \int_0^t 2X(s) dX(s) \quad (77)$$

## 5.2 Relation with the Malliavin derivative

The reader familiar with Malliavin calculus is by now probably intrigued by the relation between the pathwise calculus introduced above and the stochastic calculus of variations as introduced by Malliavin [23] and developed by Bismut [2, 3], Stroock [30], Shigekawa [29], Watanabe [33] and others.

To investigate this relation, consider the case where  $X(t) = W(t)$  is the Brownian motion and  $\mathbb{P}$  the Wiener measure. Denote by  $\Omega_0$  the canonical Wiener space  $(C_0([0, T], \mathbb{R}^d), \|\cdot\|_\infty, \mathbb{P})$  endowed with its Borelian  $\sigma$ -algebra, the filtration of the canonical process.

Consider an  $\mathcal{F}_T$ -measurable functional  $H = H(X(t), t \in [0, T]) = H(X_T)$  with  $E[|H|^2] < \infty$  and define the martingale  $Y(t) = E[H|\mathcal{F}_t]$ . If  $H$  is differentiable in the Malliavin sense [23, 25, 30] e.g.  $H \in \mathbf{D}^{1,2}$  with Malliavin derivative  $\mathbb{D}_t H$ , then the Clark-Haussmann-Ocone formula [18, 26, 25] gives a stochastic integral representation of the martingale  $Y$  in terms of the Malliavin derivative of  $H$ :

$$H = E[H] + \int_0^T {}^p E[\mathbb{D}_t H | \mathcal{F}_t] dW_t \quad (78)$$

where  ${}^p E[\mathbb{D}_t H | \mathcal{F}_t]$  denotes the predictable projection of the Malliavin derivative. Similar representations have been obtained under a variety of conditions [2, 7, 11, 1].

As shown by Pardoux and Peng [27, Prop. 2.2] in the Markovian case, one does not really need the full specification of the (anticipative) process  $(\mathbb{D}_t H)_{t \in [0, T]}$  in order to recover the (predictable) martingale representation of  $H$ . Indeed, when  $X$  is a (Markovian) diffusion process, Pardoux &

Peng [27, Prop. 2.2] show that in fact the integrand is given by the “diagonal” Malliavin derivative  $\mathbb{D}_t Y_t$ , which is non-anticipative.

Theorem 21 shows that this result holds beyond the Markovian case and yields an explicit non-anticipative representation for the martingale  $Y$  as a pathwise derivative of the martingale  $Y$ , provided that  $Y \in \mathcal{C}^{1,2}(X)$ .

The uniqueness of the integrand in the martingale representation (74) leads to the following result:

**Theorem 22.** *Denote by*

- $\mathcal{P}$  the set of  $\mathcal{F}_t$ -adapted processes on  $[0, T]$  with values in  $L^1(\Omega, \mathcal{F}_T, \mathbb{P})$ .
- $\mathbb{A}_p$  the set of (anticipative) processes on  $[0, T]$  with values in  $L^p(\Omega, \mathcal{F}_T, \mathbb{P})$ .
- $\mathbb{D}$  the Malliavin derivative operator, which associates to a random variable  $H \in \mathbf{D}^{1,1}(0, T)$  the (anticipative) process  $(\mathbb{D}_t H)_{t \in [0, T]} \in \mathbb{A}_1$ .
- $\mathbb{H}$  the set of Malliavin-differentiable functionals  $H \in \mathbf{D}^{1,1}(0, T)$  whose predictable projection  $H_t = {}^p E[H|\mathcal{F}_t]$  admits a  $\mathcal{C}_b^{1,2}(W)$  version:

$$\mathbb{H} = \{H \in \mathbf{D}^{1,1}, \exists Y \in \mathcal{C}_b^{1,2}(W), E[H|\mathcal{F}_t] = Y(t) \quad dt \times d\mathbb{P} - a.e\}$$

Then the following diagram is commutative, in the sense of  $dt \times d\mathbb{P}$  almost everywhere equality:

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{\mathbb{D}} & \mathbb{A}_1 \\ \downarrow ({}^p E[\cdot|\mathcal{F}_t])_{t \in [0, T]} & & \downarrow ({}^p E[\cdot|\mathcal{F}_t])_{t \in [0, T]} \\ \mathcal{C}_b^{1,2}(W) & \xrightarrow{\nabla_W} & \mathcal{P} \end{array}$$

*Proof.* The Clark-Haussmann-Ocone formula extended to  $\mathbf{D}^{1,1}$  in [18] gives

$$H = E[H] + \int_0^T {}^p E[\mathbb{D}_t H|\mathcal{F}_t] dW_t \quad (79)$$

where  ${}^p E[\mathbb{D}_t H|\mathcal{F}_t]$  denotes the predictable projection of the Malliavin derivative. On other hand theorem 21 gives:

$$H = E[H] + \int_0^T \nabla_W E[H|\mathcal{F}_t] dW(t) \quad (80)$$

Hence:

$${}^p E[\mathbb{D}_t H|\mathcal{F}_t] = \nabla_W E[H|\mathcal{F}_t] \quad (81)$$

$dt \times d\mathbb{P}$  almost everywhere. □

Let us conclude with a note on potential applications to numerical simulation. 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 in a pathwise manner. This implies the usefulness of (74) for the numerical computation of martingale representations, a topic which we further explore in a forthcoming work.



## 6 Weak derivatives and integration by parts for stochastic integrals

Assume now that  $X$  is a continuous, square-integrable real-valued martingale. We will now extend the operator  $\nabla_X$  to a *weak derivative* over a space of stochastic integrals, that is, an operator which verifies

$$\nabla_X \left( \int \phi \cdot dX \right) = \phi, \quad dt \times d\mathbb{P} - a.s. \quad (82)$$

for square-integrable stochastic integrals of the form:

$$Y(t) = \int_0^t \phi_s dX(s) \quad \text{where} \quad E \left[ \int_0^t \phi_s^2 d[X](s) \right] < \infty \quad (83)$$

Let  $\mathcal{L}^2(X)$  be the Hilbert space of progressively-measurable processes  $\phi$  such that:

$$\|\phi\|_{\mathcal{L}^2(X)}^2 = E \left[ \int_0^t \phi_s^2 d[X](s) \right] < \infty \quad (84)$$

and  $\mathcal{I}^2(X)$  be the space of square-integrable stochastic integrals with respect to  $X$ :

$$\mathcal{I}^2(X) = \left\{ \int_0^\cdot \phi(t) dX(t), \phi \in \mathcal{L}^2(X) \right\} \quad (85)$$

endowed with the norm

$$\|Y\|_2^2 = E[Y(T)^2] \quad (86)$$

The Ito integral  $\phi \mapsto \int_0^\cdot \phi_s dX(s)$  is then a bijective isometry from  $\mathcal{L}^2(X)$  to  $\mathcal{I}^2(X)$  [28].

**Definition 23** (Space of test processes). The space of *test processes*  $D(X)$  is defined as

$$D(X) = \mathcal{C}_b^{1,2}(X) \cap \mathcal{I}^2(X) \quad (87)$$

**Theorem 24** (Integration by parts on  $D(X)$ ). *Let  $Y, Z \in D(X)$ . Then:*

$$E[Y(T)Z(T)] = E \left[ \int_0^T \nabla_X Y(t) \nabla_X Z(t) d[X](t) \right] \quad (88)$$

*Proof.* Let  $Y, Z \in D(X) \subset \mathcal{C}_b^{1,2}(X)$ . Then  $Y, Z$  are martingales with  $Y(0) = Z(0) = 0$  and  $E[|Y(T)|^2] < \infty, E[|Z(T)|^2] < \infty$ . Applying Theorem 21 to  $Y$  and  $Z$ , we obtain

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

Applying the Ito isometry formula yields the result.  $\square$

Using this result, we can extend the operator  $\nabla_X$  in a weak sense to a suitable space of the space of (square-integrable) stochastic integrals, where  $\nabla_X Y$  is characterized by (88) being satisfied against all test processes.

The following definition introduces the Hilbert space  $\mathcal{W}^{1,2}(X)$  of martingales on which  $\nabla_X$  acts as a weak derivative, characterized by integration-by-part formula (88). This definition may be also viewed as a non-anticipative counterpart of Wiener-Sobolev spaces in the Malliavin calculus [23, 29].

**Definition 25** (Martingale Sobolev space). The Martingale Sobolev space  $\mathcal{W}^{1,2}(X)$  is defined as the closure in  $\mathcal{I}^2(X)$  of  $D(X)$ .

The Martingale Sobolev space  $\mathcal{W}^{1,2}(X)$  is in fact none other than  $\mathcal{I}^2(X)$ , the set of square-integrable stochastic integrals:

**Lemma 26.**  $\{\nabla_X Y, Y \in D(X)\}$  is dense in  $\mathcal{L}^2(X)$  and

$$\mathcal{W}^{1,2}(X) = \mathcal{I}^2(X).$$

*Proof.* We first observe that the set of “cylindrical” integrands 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  $f \in C_b^\infty(\mathbb{R}^n \rightarrow \mathbb{R})$  is a total set in  $\mathcal{L}^2(X)$  i.e. the linear span of  $U$  of such functions 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

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

where the functional  $F$  is defined on  $\Upsilon$  as:

$$F_t(x_t, v_t) = f(x(t_1-), \dots, x(t_n-))(x(t) - x(t_n-))1_{t>t_n} \in \mathbb{F}^{\infty,1}$$

so that:

$$\begin{aligned} \nabla_x F_t(x_t, v_t) &= f(x_{t_1-}, \dots, x_{t_n-})1_{t>t_n} \in F_r^\infty \cap \mathbb{B} \\ \nabla_x^2 F_t(x_t, v_t) &= 0, \mathcal{D}_t F(x_t, v_t) = 0 \end{aligned}$$

which prove that  $F \in \mathbb{C}_b^{1,2} \cap \mathbb{F}^{\infty,1}$ . Hence,  $Y \in \mathbb{C}_b^{1,2}(X)$ . 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  $\mathcal{I}^2(X)$ , the density of  $U$  in  $\mathcal{L}^2(X)$  entails the density of  $I_X(U)$  in  $\mathcal{I}^2(X)$ , so  $\mathcal{W}^{1,2}(X) = \mathcal{I}^2(X)$ .  $\square$

**Theorem 27** (Weak derivative on  $\mathcal{W}^{1,2}(X)$ ). The vertical derivative  $\nabla_X : D(X) \mapsto \mathcal{L}^2(X)$  is closable on  $\mathcal{W}^{1,2}(X)$ . Its closure defines a bijective isometry

$$\begin{aligned} \nabla_X : \mathcal{W}^{1,2}(X) &\mapsto \mathcal{L}^2(X) \\ \int_0^T \phi \cdot dX &\mapsto \phi \end{aligned} \tag{89}$$

characterized by the following integration by parts formula: for  $Y \in \mathcal{W}^{1,2}(X)$ ,  $\nabla_X Y$  is the unique element of  $\mathcal{L}^2(X)$  such that

$$\forall Z \in D(X), \quad E[Y(T)Z(T)] = E \left[ \int_0^T \nabla_X Y(t) \nabla_X Z(t) d[X](t) \right]. \tag{90}$$

In particular,  $\nabla_X$  is the adjoint of the Ito stochastic integral

$$\begin{aligned} I_X : \mathcal{L}^2(X) &\mapsto \mathcal{W}^{1,2}(X) \\ \phi &\mapsto \int_0^\cdot \phi \cdot dX \end{aligned} \quad (91)$$

in the following sense:

$$\forall \phi \in \mathcal{L}^2(X), \quad \forall Y \in \mathcal{W}^{1,2}(X), \quad \langle Y, I_X(\phi) \rangle_{\mathcal{W}^{1,2}(X)} = \langle \nabla_X Y, \phi \rangle_{\mathcal{L}^2(X)} \quad (92)$$

$$i.e. \quad E[Y(T) \int_0^T \phi \cdot dX] = E[\int_0^T \nabla_X Y \phi \cdot d[X]] \quad (93)$$

*Proof.* Any  $Y \in \mathcal{W}^{1,2}(X)$  may be written as  $Y(t) = \int_0^t \phi(s) dX(s)$  for some  $\phi \in \mathcal{L}^2(X)$ , which is uniquely defined  $d[X] \times d\mathbb{P}$  a.e. The Ito isometry formula then guarantees that (90) holds for  $\phi$ . One still needs to prove that (90) uniquely characterizes  $\phi$ . If some process  $\psi$  also satisfies (90), then, denoting  $Y' = \mathcal{I}_X(\psi)$  its stochastic integral with respect to  $X$ , (90) then implies that  $U = Y' - Y$  verifies

$$\forall Z \in D(X), \quad \langle U, Z \rangle_{\mathcal{W}^{1,2}(X)} = E[U(T)Z(T)] = 0$$

which implies  $U = 0$   $d[X] \times d\mathbb{P}$  a.e. since by construction  $D(X)$  is dense in  $\mathcal{W}^{1,2}(X)$ . Hence,  $\nabla_X : D(X) \mapsto \mathcal{L}^2(X)$  is closable on  $\mathcal{W}^{1,2}(X)$

This construction shows that  $\nabla_X : \mathcal{W}^{1,2}(X) \mapsto \mathcal{L}^2(X)$  is a bijective isometry which coincides with the adjoint of the Ito integral on  $\mathcal{W}^{1,2}(X)$ .  $\square$

Thus, Ito's stochastic integral  $\mathcal{I}_X$  with respect to  $X$ , viewed as the map

$$I_X : \mathcal{L}^2(X) \mapsto \mathcal{W}^{1,2}(X)$$

admits an inverse on  $\mathcal{W}^{1,2}(X)$  which is a weak form of the vertical derivative  $\nabla_X$  introduced in Definition 10.

*Remark 28.* In other words, we have established that for any  $\phi \in \mathcal{L}^2(X)$  the relation

$$\nabla_X(\phi \cdot X)(t) = \phi(t) \quad \text{where} \quad (\phi \cdot X)(t) = \int_0^t \phi(u) dX(u) \quad (94)$$

holds in a weak sense.

In particular these results hold when  $X = W$  is a Brownian motion. We can now restate a square-integrable version of theorem 22, which holds on  $\mathbf{D}^{1,2}$ , and where the operator  $\nabla_W$  is defined in the weak sense of theorem 27.

**Theorem 29** (Lifting theorem). *Consider  $\Omega_0 = C_0([0, T], \mathbb{R}^d)$  endowed with its Borelian  $\sigma$ -algebra, the filtration of the canonical process and the Wiener measure  $\mathbb{P}$ . Then the following diagram is commutative in the sense of  $dt \times d\mathbb{P}$  equality:*

$$\begin{array}{ccc} \mathcal{I}^2(W) & \xrightarrow{\nabla_W} & \mathcal{L}^2(W) \\ \uparrow (E[\cdot | \mathcal{F}_t])_{t \in [0, T]} & & \uparrow (E[\cdot | \mathcal{F}_t])_{t \in [0, T]} \\ \mathbf{D}^{1,2} & \xrightarrow{\mathbb{D}} & \mathbb{A}_2 \end{array}$$

*Remark 30.* With a slight abuse of notation, the above result can be also written as

$$\forall H \in L^2(\Omega_0, \mathcal{F}_T, \mathbb{P}), \quad \nabla_W (E[H|\mathcal{F}_t]) = E[\mathbb{D}_t H|\mathcal{F}_t] \quad (95)$$

In other words, the conditional expectation operator intertwines  $\nabla_W$  with the Malliavin derivative.

Thus, the conditional expectation operator (more precisely: the *predictable* projection on  $\mathcal{F}_t$ ) 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 by the weak derivative operator  $\nabla_W$  and the Ito stochastic integral. Obviously, making this last statement precise is a whole research program, beyond the scope of this paper.

## 7 Functional equations for martingales

Consider now a semimartingale  $X$  whose characteristics are right-continuous functionals:

$$dX(t) = b_t(X_t, A_t)dt + \sigma_t(X_t, A_t)dW(t) \quad (96)$$

where  $b, \sigma$  are non-anticipative functionals on  $\Upsilon$  (in the sense of Definition 1) with values in  $\mathbb{R}^d$ -valued (resp.  $\mathbb{R}^{d \times n}$ , whose coordinates are in  $\mathbb{F}_r^\infty$ ). The *topological support* of the law of  $(X, A)$  in  $(C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \|\cdot\|_\infty)$  is defined to be the subset  $\text{supp}(X, A)$  of all paths  $(x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$  for which every (open) neighborhood has positive measure:

$$\text{supp}(X, A) = \{(x, v) \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T \mid \text{for any Borel neighborhood } V \text{ of } (x, v), \mathbb{P}((X, A) \in V) > 0\}$$

Functionals of  $X$  which have the (local) martingale property play an important role in control theory and harmonic analysis. The following result characterizes a functional  $F \in \mathbb{C}_b^{1,2} \cap \mathbb{F}^{\infty,1}$  which define a *local martingale* as the solution to a functional version of the Kolmogorov backward equation:

**Theorem 31** (Functional equation for  $\mathcal{C}^{1,2}$  martingales). *If  $F \in \mathbb{C}_b^{1,2} \cap \mathbb{F}^{\infty,1}$ , then  $Y(t) = F_t(X_t, A_t)$  is a local martingale if and only if  $F$  satisfies the functional partial differential equation:*

$$D_t F(x_t, v_t) + b_t(x_t, v_t) \nabla_x F_t(x_t, v_t) + \frac{1}{2} \text{tr}[\nabla_x^2 F(x_t, v_t) \sigma_t^t \sigma_t(x_t, v_t)] = 0, \quad (97)$$

on the topological support of the law of the process  $(X, A)$  in  $(C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \|\cdot\|_\infty)$ .

*Proof.* If  $F \in \mathbb{C}_b^{1,2} \cap \mathbb{F}^{\infty,1}$ , then applying Theorem 18 to  $Y(t) = F_t(X_t, A_t)$ , (97) implies that the finite variation term in (62) is almost-surely zero:  $Y(t) = \int_0^t \nabla_x F_t(X_t, A_t) dX(t)$ . Hence  $Y$  is a local martingale.

Conversely, assume that  $Y$  is a local martingale. Note that  $Y$  is continuous by Theorem 7. Suppose the functional relation (97) is not satisfied at some  $(x, v)$  belongs to the  $\text{supp}(X, A) \subset C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$ . Then there exists  $t_0 < T$ ,  $\eta > 0$  and  $\epsilon > 0$  such that

$$|D_t F(x_t, v_t) + b_t(x_t, v_t) \nabla_x F_t(x_t, v_t) + \frac{1}{2} \text{tr}[\nabla_x^2 F(x_t, v_t) \sigma_t^t \sigma_t(x_t, v_t)]| > \epsilon \quad (98)$$

for  $t \in [t_0, t_0 + \eta]$ , by right-continuity of the expression. By continuity of the expression for the  $d_\infty$  norm, there exist an open neighborhood of  $(x, v)$  in  $C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$  such that, for all  $(x', v')$  in this neighborhood and all  $t \in [t_0, t_0 + \eta]$ :

$$|D_t F(x'_t, v'_t) + b_t(x'_t, v'_t) \nabla_x F_t(x_t, v_t) + \frac{1}{2} \text{tr}[\nabla_x^2 F(x'_t, v'_t) \sigma_t^t \sigma_t(x'_t, v'_t)]| > \frac{\epsilon}{2} \quad (99)$$

Since  $(X, A)$  belongs to this neighborhood with non-zero probability, it proves that:

$$D_t F(X_t, A_t) + b_t(X_t, A_t) \nabla_x F_t(x_t, v_t) + \frac{1}{2} \text{tr}[\nabla_x^2 F(X_t, A_t) \sigma_t^t \sigma_t(X_t, A_t)] > \frac{\epsilon}{2} \quad (100)$$

with non-zero  $dt \times d\mathbb{P}$  measure. Applying theorem 18 to the process  $Y(t) = F_t(X_t, A_t)$  then leads to a contradiction, because as a continuous local martingale its finite variation part should be null.  $\square$

The martingale property of  $F(X, A)$  implies no restriction on the behavior of  $F$  outside  $\text{supp}(X, A)$  so one cannot hope for uniqueness of  $F$  on  $\Upsilon$  in general. However, the following result gives a condition for uniqueness of a solution of (97) on  $\text{supp}(X, A)$ :

**Theorem 32** (Uniqueness result). *Let  $h$  be a continuous functional on  $(C_0([0, T]) \times \mathcal{S}_T, \|\cdot\|_\infty)$ . Any solution  $F \in \mathbb{C}_b^{1,2}$  of the functional equation (97), verifying*

$$F_T(x, v) = h(x, v) \quad (101)$$

$$E[\sup_{t \in [0, T]} |F_t(X_t, A_t)|] < \infty \quad (102)$$

*is uniquely defined on the topological support  $\text{supp}(X, A)$  of  $(X, A)$  in  $(C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T, \|\cdot\|)$ : if  $F^1, F^2 \in \mathbb{C}_b^{1,2}([0, T])$  verify (97)-(101)-(102) then*

$$\forall (x, v) \in \text{supp}(X, A), \quad \forall t \in [0, T] \quad F_t^1(x_t, v_t) = F_t^2(x_t, v_t). \quad (103)$$

*Proof.* Let  $F^1$  and  $F^2$  be two such solutions. Theorem 31 shows that they are local martingales. The integrability condition (102) guarantees that they are true martingales, so that we have the equality:  $F_t^1(X_t, A_t) = F_t^2(X_t, A_t) = E[h(X_T, A_T) | \mathcal{F}_t]$  almost surely. Hence reasoning along the lines of the proof of theorem 31 shows that  $F_t^1(x_t, v_t) = F_t^2(x_t, v_t)$  if  $(x, v) \in \text{supp}(X, A)$ .  $\square$

*Example 15.* Consider a scalar diffusion

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t) \quad X(0) = x_0 \quad (104)$$

whose law  $\mathbb{P}^{x_0}$  is defined as the solution of the martingale problem [32] for the operator

$$L_t f = \frac{1}{2} \sigma^2(t, x) \partial_x^2 f(t, x) + b(t, x) \partial_x f(t, x)$$

where  $b$  and  $\sigma$  are continuous and bounded functions, with  $\sigma$  bounded away from zero. We are interested in computing the martingale

$$Y(t) = E\left[\int_0^T g(t, X(t))d[X](t) | \mathcal{F}_t\right] \quad (105)$$

for a continuous bounded function  $g$ . The topological support of the process  $(X, A)$  under  $\mathbb{P}^{x_0}$  is then given by the Stroock-Varadhan support theorem [31, Theorem 3.1.] which yields:

$$\{(x, (\sigma^2(t, x(t)))_{t \in [0, T]}) \mid x \in C_0(\mathbb{R}^d, [0, T]), x(0) = x_0\}, \quad (106)$$

From theorem 31 a necessary condition for  $Y$  to have a functional representation  $Y = F(X, A)$  with  $F \in \mathbb{C}^{1,2}([0, T])$  is

$$\begin{aligned} \mathcal{D}_t F(x_t, (\sigma^2(u, x(u)))_{u \leq t}) + b(t, x(t)) \nabla_x F_t(x_t, (\sigma^2(u, x(u)))_{u \in [0, t]}) \\ + \frac{1}{2} \sigma^2(t, x(t)) \nabla_x^2 F_t(x_t, (\sigma^2(u, x(u)))_{u \in [0, t]}) = 0 \end{aligned} \quad (107)$$

together with the terminal condition:

$$F_T(x_T, (\sigma^2(u, x(u)))_{u \in [0, T]}) = \int_0^T g(t, x(t)) \sigma^2(t, x(t)) dt \quad (108)$$

for all  $x \in C_0(\mathbb{R}^d), x(0) = x_0$ . Moreover, from theorem 32, we know that there any solution satisfying the integrability condition:

$$E\left[ \sup_{t \in [0, T]} |F_t(X_t, A_t)| \right] < \infty \quad (109)$$

is unique on  $\text{supp}(X, A)$ . If such a solution exists, then the martingale  $F_t(X_t, A_t)$  is a version of  $Y$ . To find such a solution, we look for a functional of the form:

$$F_t(x_t, v_t) = \int_0^t g(u, x(u)) v(u) du + f(t, x(t))$$

where  $f$  is a smooth  $C^{1,2}$  function. Elementary computation show that  $F \in \mathbb{C}^{1,2}([0, T])$ ; so  $F$  is solution of the functional equation (107) if and only if  $f$  satisfies the Partial Differential Equation with source term:

$$\begin{aligned} \frac{1}{2} \sigma^2(t, x) \partial_x^2 f(t, x) + b(t, x) \partial_x f(t, x) + \partial_t f(t, x) = -g(t, x) \sigma^2(t, x) \\ \text{with terminal condition} \quad f(T, x) = 0 \end{aligned} \quad (110)$$

The existence of a solution  $f$  with at most exponential growth is then guaranteed by standard results on parabolic PDEs [19]. In particular, theorem 32 guarantees that there is at most one solution such that:

$$E\left[ \sup_{t \in [0, T]} |f(t, X(t))| \right] < \infty \quad (111)$$

Hence the martingale  $Y$  in (105) is given by

$$Y(t) = \int_0^t g(u, X(u)) d[X](u) + f(t, X(t))$$

where  $f$  is the unique solution of the PDE (110).

## References

- [1] H. AHN, *Semimartingale integral representation*, Ann. Probab., 25 (1997), pp. 997–1010.
- [2] J.-M. BISMUT, *A generalized formula of Itô and some other properties of stochastic flows*, Z. Wahrsch. Verw. Gebiete, 55 (1981), pp. 331–350.
- [3] J.-M. BISMUT, *Calcul des variations stochastique et processus de sauts*, Z. Wahrsch. Verw. Gebiete, 63 (1983), pp. 147–235.
- [4] J. M. C. CLARK, *The representation of functionals of Brownian motion by stochastic integrals*, Ann. Math. Statist., 41 (1970), pp. 1282–1295.
- [5] R. CONT AND D.-A. FOURNIÉ, *Change of variable formulas for non-anticipative functionals on path space*, working paper, 2009.
- [6] R. CONT AND D.-A. FOURNIÉ, *A functional extension of the Ito formula*, Comptes Rendus Mathématique Acad. Sci. Paris Ser. I, 348 (2010), pp. 57–61.
- [7] M. H. DAVIS, *Functionals of diffusion processes as stochastic integrals*, Math. Proc. Comb. Phil. Soc., 87 (1980), pp. 157–166.
- [8] C. DELLACHERIE AND P.-A. MEYER, *Probabilities and potential*, vol. 29 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 1978.
- [9] J. L. DOOB, *Stochastic Processes*, Wiley, 1953.
- [10] B. DUPIRE, *Functional Itô calculus*, Portfolio Research Paper 2009-04, Bloomberg, 2009.
- [11] R. J. ELLIOTT AND M. KOHLMANN, *A short proof of a martingale representation result*, Statistics & Probability Letters, 6 (1988), pp. 327–329.
- [12] H. FÖLLMER, *Calcul d'Itô sans probabilités*, in Séminaire de Probabilités XV, vol. 850 of Lecture Notes in Math., Springer, Berlin, 1981, pp. 143–150.
- [13] U. G. HAUSSMANN, *Functionals of Itô processes as stochastic integrals*, SIAM J. Control Optimization, 16 (1978), pp. 252–269.
- [14] U. G. HAUSSMANN, *On the integral representation of functionals of Itô processes*, Stochastics, 3 (1979), pp. 17–27.
- [15] K. ITO, *On a stochastic integral equation*, Proceedings of the Imperial Academy of Tokyo, 20 (1944), pp. 519–524.
- [16] K. ITO, *On stochastic differential equations*, Proceedings of the Imperial Academy of Tokyo, 22 (1946), pp. 32–35.
- [17] J. JACOD, S. MÉLÉARD, AND P. PROTTER, *Explicit form and robustness of martingale representations*, Ann. Probab., 28 (2000), pp. 1747–1780.
- [18] I. KARATZAS, D. L. OCONE, AND J. LI, *An extension of Clark's formula*, Stochastics Stochastics Rep., 37 (1991), pp. 127–131.

- [19] N. V. KRYLOV, *Lectures on elliptic and parabolic equations in Hölder spaces*, vol. 12 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1996.
- [20] H. KUNITA AND S. WATANABE, *On square integrable martingales*, Nagoya Math. J., 30 (1967), pp. 209–245.
- [21] T. J. LYONS, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana, 14 (1998), pp. 215–310.
- [22] P. MALLIAVIN, *Stochastic calculus of variation and hypoelliptic operators*, in Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), New York, 1978, Wiley, pp. 195–263.
- [23] P. MALLIAVIN, *Stochastic analysis*, Springer, 1997.
- [24] P. MEYER, *Un cours sur les integrales stochastiques*. Semin. Probab. X, Univ. Strasbourg 1974/75, Lect. Notes Math. 511, 245–400 (1976)., 1976.
- [25] D. NUALART, *Malliavin calculus and its applications*, vol. 110 of CBMS Regional Conference Series in Mathematics, CBMS, Washington, DC, 2009.
- [26] D. L. OCONE, *Malliavin’s calculus and stochastic integral representations of functionals of diffusion processes*, Stochastics, 12 (1984), pp. 161–185.
- [27] E. PARDOUX AND S. PENG, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, in Stochastic partial differential equations and their applications, vol. 716 of Lecture Notes in Control and Informatic Science, Springer, 1992.
- [28] P. E. PROTTER, *Stochastic integration and differential equations*, Springer-Verlag, Berlin, 2005. Second edition.
- [29] I. SHIGEKAWA, *Derivatives of Wiener functionals and absolute continuity of induced measures*, J. Math. Kyoto Univ., 20 (1980), pp. 263–289.
- [30] D. W. STROOCK, *The Malliavin calculus, a functional analytic approach*, J. Funct. Anal., 44 (1981), pp. 212–257.
- [31] D. W. STROOCK AND S. R. S. VARADHAN, *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, Berkeley, Calif., 1972, Univ. California Press, pp. 333–359.
- [32] D. W. STROOCK AND S. R. S. VARADHAN, *Multidimensional diffusion processes*, vol. 233 of Grundlehren der Mathematischen Wissenschaften, Berlin, 1979, Springer-Verlag, pp. xii+338.
- [33] S. WATANABE, *Lectures on stochastic differential equations and Malliavin calculus*, vol. 73 of Lectures on Mathematics and Physics, Tata Institute of Fundamental Research, Bombay, 1984.



## A Proof of Theorems 8 and 18

### A.1 Proofs of theorem 8

In order to prove theorem 8 in the general case where  $A$  is just required to be cadlag, we need the following three lemmas:

**Lemma 33.** *Let  $f$  be a cadlag function on  $[0, T]$  and define  $\Delta f(t) = f(t) - f(t-)$ . Then*

$$\forall \epsilon > 0, \quad \exists \eta > 0, \quad |x - y| \leq \eta \Rightarrow |f(x) - f(y)| \leq \epsilon + \sup_{t \in [x, y]} \{|\Delta f(t)|\} \quad (112)$$

*Proof.* Assume the conclusion does not hold. Then there exists a sequence  $(x_n, y_n)_{n \geq 1}$  such that  $x_n \leq y_n$ ,  $y_n - x_n \rightarrow 0$  but  $|f(x_n) - f(y_n)| > \epsilon + \sup_{t \in [x_n, y_n]} \{|\Delta f(t)|\}$ . We can extract a convergent subsequence  $(x_{\psi(n)})$  such that  $x_{\psi(n)} \rightarrow x$ . Noting that either an infinity of terms of the sequence are less than  $x$  or an infinity are more than  $x$ , we can extract *monotone* subsequences  $(u_n, v_n)_{n \geq 1}$  which converge to  $x$ . If  $(u_n), (v_n)$  both converge to  $x$  from above or from below,  $|f(u_n) - f(v_n)| \rightarrow 0$  which yields a contradiction. If one converges from above and the other from below,  $\sup_{t \in [u_n, v_n]} \{|\Delta f(t)|\} > |\Delta f(x)|$  but  $|f(u_n) - f(v_n)| \rightarrow |\Delta f(x)|$ , which results in a contradiction as well. Therefore (112) must hold.  $\square$

**Lemma 34.** *If  $\alpha \in \mathbb{R}$  and  $V$  is an adapted cadlag process defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and  $\sigma$  is a optional time, then:*

$$\tau = \inf\{t > \sigma, \quad |V(t) - V(t-)| > \alpha\} \quad (113)$$

*is a stopping time.*

*Proof.* We can write that:

$$\{\tau \leq t\} = \bigcup_{q \in \mathbb{Q} \cap [0, t)} (\{\sigma \leq t - q\} \cap \{\sup_{u \in (t-q, t]} |V(u) - V(u-)| > \alpha\}) \quad (114)$$

and

$$\{\sup_{u \in (t-q, t]} |V(u) - V(u-)| > \alpha\} = \bigcup_{n_0 > 1} \bigcap_{n > n_0} \left\{ \sup_{1 \leq i \leq 2^n} |V(t - q \frac{i-1}{2^n}) - V(t - q \frac{i}{2^n})| > \alpha \right\} \quad (115)$$

thanks to the lemma 33.  $\square$

The following lemma is a consequence of lemma 33:

**Lemma 35** (Uniform approximation of cadlag functions by step functions).

*Let  $h$  be a cadlag function on  $[0, T]$  and  $(t_k^n)_{n \geq 0, k=0..n}$  is a sequence of subdivisions  $0 = t_0^n < t_1^n < \dots < t_{k_n}^n = T$  of  $[0, T]$  such that:*

$$\sup_{0 \leq i \leq k-1} |t_{i+1}^n - t_i^n| \xrightarrow{n \rightarrow \infty} 0 \quad \sup_{u \in [0, T] \setminus \{t_0^n, \dots, t_{k_n}^n\}} |\Delta f(u)| \xrightarrow{n \rightarrow \infty} 0$$

*then*

$$\sup_{u \in [0, T]} |h(u) - \sum_{i=0}^{k_n-1} h(t_i^n) 1_{[t_i^n, t_{i+1}^n)}(u) + h(t_{k_n}^n) 1_{\{t_{k_n}^n\}}(u)| \xrightarrow{n \rightarrow \infty} 0 \quad (116)$$

We can now prove Theorem 8 in the general case where  $A$  is only assumed to be cadlag.

**Proof of Theorem 8:** Since the trajectories of  $Y(t)$  are right continuous we just have to prove that the process is adapted. For this we introduce a sequence of random subdivision of  $[0, T]$ , indexed by  $n$ , as follows: starting with the deterministic subdivision  $t_i^n = \frac{iT}{2^n}, i = 0..2^n$  we add the time of jumps of  $X$  and  $A$  of size greater or equal to  $\frac{1}{n}$ . We define the following sequence of stopping times:

$$\tau_0^n = 0 \quad \tau_k^n = \inf\{t > \tau_{k-1}^n | 2^n t \in \mathbb{N} \text{ or } |A(t) - A(t-)| > \frac{1}{n}\} \wedge T \quad (117)$$

We define the stepwise approximations of  $X$  and  $A$  along the subdivision of index  $n$ :

$$\begin{aligned} X^n(t) &= \sum_{k=0}^{\infty} X_{\tau_k^n} 1_{[\tau_k^n, \tau_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \\ A^n(t) &= \sum_{k=0}^{\infty} A_{\tau_k^n} 1_{[\tau_k^n, \tau_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \end{aligned} \quad (118)$$

as well as their truncations of rank  $K$ :

$$\begin{aligned} {}_K X^n(t) &= \sum_{k=0}^K X_{\tau_k^n} 1_{[\tau_k^n, \tau_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \\ {}_K A^n(t) &= \sum_{k=0}^K A_{\tau_k^n} 1_{[\tau_k^n, \tau_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \end{aligned} \quad (119)$$

The random variable  $Y^n(t) = F_t(X_t^n, A_t^n)$  can be written as the following almost-sure limit:

$$Y^n(t) = \lim_{K \rightarrow \infty} F_t({}_K X_t^n, {}_K A_t^n) \quad (120)$$

because  ${}_K X_t^n, {}_K A_t^n$  coincides with  $X_t^n, A_t^n$  for  $K$  sufficiently large. The truncations  $F_t({}_K X_t^n, {}_K A_t^n)$  are  $\mathcal{G}_t$ -measurable as they are continuous functions of the random variables  $\{X(\tau_k^n) 1_{\tau_k^n \leq t}, A(\tau_k^n) 1_{\tau_k^n \leq t}\}$ , so  $Y^n(t)$  is  $\mathcal{G}_t$ -measurable. Thanks to lemma 35,  $X_t^n$  and  $A_t^n$  almost surely converge uniformly to  $X_t$  and  $A_t$ , hence  $Y^n(t)$  converges almost surely to  $Y(t)$ , which concludes the proof.

## A.2 Proofs of Theorem 18

Following is the proof of theorem 18 in the general case where  $A$  is just assumed to be cadlag.

*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 subdivision of  $[0, T]$ , indexed by  $n$ , as follows: starting with the deterministic subdivision  $t_i^n = \frac{iT}{2^n}, i = 0..2^n$  we add the time of jumps of  $X$  and  $A$  of size greater or equal to  $\frac{1}{n}$ . We define the following sequence of stopping times:

$$\tau_0^n = 0 \quad \tau_k^n = \inf\{t > \tau_{k-1}^n | 2^n t \in \mathbb{N} \text{ or } |A(t) - A(t-)| > \frac{1}{n}\} \wedge t \quad (121)$$

The following arguments apply pathwise. Lemma 35 ensures that  $\eta_n = \sup\{|A(u) - A(\tau_i^n)| + |X(u) - X(\tau_i^n)| + \frac{t}{2^n}, i \leq 2^n, u \in [\tau_i^n, \tau_{i+1}^n]\} \rightarrow_{n \rightarrow \infty} 0$ . Let  $\eta > 0, C > 0$  be such that, for any  $s < T$ ,

for any  $(x, v) \in D([0, s], \mathbb{R}^d) \times \mathcal{S}_s^+$ ,  $d_\infty((X_s, A_s), (x, v)) < \eta \Rightarrow |F_s(x, A_s) - F_s(x, v_s)| \leq C\|A_s - v_s\|_1$ , and we will assume  $n$  large enough so that  $\eta_n < \eta$ .

Denoting  ${}_n X = \sum_{i=0}^{\infty} X(\tau_i^n) 1_{[\tau_i^n, \tau_{i+1}^n)} + X(t) 1_{\{t\}}$  the cadlag piecewise constant approximation of  $X_t$ ,

$$F_t(X_t, A_t) - F_0(X_0, A_0) = F_t(X_t, A_t) - F_t({}_n X_t, A_t) + \sum_{i=0}^{k_n-1} F_{\tau_{i+1}^n}({}_n X_{\tau_{i+1}^n}, A_{\tau_{i+1}^n}) - F_{\tau_i^n}({}_n X_{\tau_i^n}, A_{\tau_i^n}) \quad (122)$$

It is first obvious that  $|F(X_t, A_t) - F({}_n X_t, A_t)| \rightarrow 0$  as  $n \rightarrow \infty$ . Denote  $\delta_i = X_{\tau_{i+1}^n} - X_{\tau_i^n}$  and  $h_i = \tau_{i+1}^n - \tau_i^n$ . Each term in the sum can then be decomposed as

$$[F_{\tau_{i+1}^n}({}_n X_{\tau_{i+1}^n}, A_{\tau_{i+1}^n}) - F_{\tau_{i+1}^n}({}_n X_{\tau_{i+1}^n}, A_{\tau_i^n, h_i})] + [F_{\tau_{i+1}^n}({}_n X_{\tau_{i+1}^n}, A_{\tau_i^n, h_i}) - F_{\tau_{i+1}^n}({}_n X_{\tau_i^n, h_i}, A_{\tau_i^n, h_i})] + F_{\tau_{i+1}^n}({}_n X_{\tau_i^n, h_i}, A_{\tau_i^n, h_i}) - F_{\tau_i^n}({}_n X_{\tau_i^n}, A_{\tau_i^n}) \quad (123)$$

The first term in (123) is bounded by

$$C\|A_{\tau_{i+1}^n} - A_{\tau_i^n, h_i}\|_1 = C \int_{\tau_i^n}^{\tau_{i+1}^n} |A(s) - A(\tau_i^n)| ds \leq C|\tau_{i+1}^n - \tau_i^n| \eta_n$$

by right continuity of  $A$ . Summing over  $i$  leads to a term which is bounded by  $Ct\eta_n$ , hence converging to 0 as  $n \rightarrow \infty$ .

Denote by  ${}_n Y_{\tau_{i+1}^n} = {}_n X_{\tau_i^n, h_i}$  the horizontal extension of  ${}_n X_{\tau_i^n}$  to  $[\tau_i^n, \tau_{i+1}^n]$ . Noting that  ${}_n Y_{\tau_{i+1}^n}^{\delta_i} = {}_n X_{\tau_{i+1}^n}$ , the second term in (123) can be written  $\phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0)$  where  $\phi(u) = F_{\tau_{i+1}^n}({}_n Y_{\tau_{i+1}^n}^u, A_{\tau_i^n, h_i})$ . Since  $F \in \mathbb{C}^{1,2}([0, T])$ ,  $\phi$  is  $C^2$  and  $\phi'(u) = \nabla_x F_{\tau_{i+1}^n}({}_n Y_{\tau_{i+1}^n}^u, A_{\tau_i^n, h_i})$ ,  $\phi''(u) = \nabla_x^2 F_{\tau_{i+1}^n}({}_n Y_{\tau_{i+1}^n}^u, A_{\tau_i^n, h_i})$ . Applying the Ito formula yields

$$\begin{aligned} \phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0) &= \int_{\tau_i^n}^{\tau_{i+1}^n} \nabla_x F_{\tau_{i+1}^n}({}_n Y_{\tau_{i+1}^n}^{X(s) - X(\tau_i^n)}, A_{\tau_i^n, h_i}) dX(s) \\ &\quad + \frac{1}{2} \int_{\tau_i^n}^{\tau_{i+1}^n} \text{tr}[\nabla_x^2 F_{\tau_{i+1}^n}({}_n Y_{\tau_{i+1}^n}^{X(s) - X(\tau_i^n)}, A_{\tau_i^n, h_i}) d[X](s)] \end{aligned} \quad (124)$$

The third term in (123) can be expressed as  $\psi(\tau_{i+1}^n - t_i) - \psi(0)$  where  $\psi(h) = F_{\tau_{i+1}^n}({}_n X_{\tau_i^n, h}, A_{\tau_i^n, h})$ . By lemma 7,  $\psi$  is continuous and right-differentiable with  $\psi'(h) = \mathcal{D}_{\tau_{i+1}^n + h} F({}_n X_{\tau_i^n, h}, A_{\tau_i^n, h})$  so

$$F_{\tau_{i+1}^n}({}_n X_{\tau_i^n, h_i}, A_{\tau_i^n, h_i}) - F_{\tau_i^n}({}_n X_{\tau_i^n}, A_{\tau_i^n}) = \int_{\tau_i^n}^{\tau_{i+1}^n} \mathcal{D}_s F({}_n X_{\tau_i^n, s-t_i}, A_{\tau_i^n, s-t_i}) ds \quad (125)$$

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

$$\begin{aligned} F_t(X_t, A_t) - F_0(X_0, A_0) &= \int_0^t \mathcal{D}_s F({}_n X_{\tau_{i(s)}^n, s-\tau_{i(s)}^n}, A_{\tau_{i(s)}^n, s-\tau_{i(s)}^n}) ds \\ &\quad + \int_0^t \nabla_x F_{\tau_{i(s)+1}^n}({}_n Y_{\tau_{i(s)+1}^n}^{X(s) - X(\tau_{i(s)}^n)}, A_{\tau_{i(s)}^n, h_{i(s)}}) dX(s) \\ &\quad + \frac{1}{2} \int_0^t \text{tr}(\nabla_x^2 F_{\tau_{i(s)+1}^n}({}_n Y_{\tau_{i(s)+1}^n}^{X(s) - X(\tau_{i(s)}^n)}, A_{\tau_{i(s)}^n, h_{i(s)}}) \cdot A(s)) ds \end{aligned} \quad (126)$$

where  $r(\pi_n) \rightarrow 0$  as  $n \rightarrow \infty$ . All the approximations of  $(X, A)$  appearing in the various integrals have a  $d_\infty$ -distance from  $(X_s, A_s)$  less than  $\eta_n$  hence all the integrands appearing in the above integrals converge respectively to  $\mathcal{D}_s F(X_s, A_s), \nabla_x F_s(X_s, A_s), \nabla_x^2 F_s(X_s, A_s)$  as  $n \rightarrow \infty$  by  $d_\infty$  right-continuity. Since the derivatives are in  $\mathbb{B}$  the integrands in the various above integrals are bounded by a constant dependant only on  $F, K$  and  $R$  and  $t$  hence does not depend on  $s$  nor on  $\omega$ . The dominated convergence and the dominated convergence theorem for the stochastic integrals [28, Ch.IV Theorem 32] then ensure that the integrals converge in probability, uniformly on  $[0, t]$  for each  $t < T$ , to the terms appearing in (62) as  $n \rightarrow \infty$ .

Now we consider the general case where  $X$  and  $A$  may be unbounded. Let  $K_n$  be an increasing sequence of compact sets,  $\bigcup_{n \geq 0} K_n = \mathbb{R}^d$ , and denote  $\tau_n = \inf\{t < T \mid X_s \in \mathbb{R} - K^n \text{ or } |A_s| > n \wedge t\}$ , which are optional times. Applying the previous result to the stopped processes  $(X^{\tau_n}, A^{\tau_n})$  leads to:

$$\begin{aligned} F_t(X_t^{\tau_n}, A_t^{\tau_n}) &= \int_0^{\tau_n} [\mathcal{D}_t Y(u) du + \int_0^t \frac{1}{2} \text{tr}[\nabla_X^2 F_u(X_u, A_u) d[X](u)] + \int_0^{\tau_n} \nabla_X Y(u) \cdot dX(u) \\ &+ \int_{t \wedge \tau_n}^t D_t F_t(X_u^{\tau_n}, A_u^{\tau_n}) du \end{aligned} \quad (127)$$

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