

# Functional Ito calculus and stochastic integral representation of martingales

Rama Cont

David-Antoine Fournié

First version: June 2009. Final revision: August 2011.  
To appear in the **Annals of Probability**.\*

## Abstract

We develop a non-anticipative calculus for functionals of a continuous semimartingale, using an extension of the Ito formula to path-dependent functionals which possess certain directional derivatives. The construction is based on a pathwise derivative, introduced by B Dupire, for functionals on the space of right-continuous functions with left limits. We show that this functional derivative admits a suitable extension to the space of square-integrable martingales. This extension defines a weak derivative which is shown to be the inverse of the Ito integral and which may be viewed as a non-anticipative “lifting” of the Malliavin derivative.

These results lead to a constructive martingale representation formula for Ito processes. By contrast with the Clark-Haussmann-Ocone formula, this representation only involves non-anticipative quantities which may be computed pathwise.

Keywords: stochastic calculus, functional calculus, functional Ito formula, Malliavin derivative, martingale representation, semimartingale, Wiener functionals, Clark-Ocone formula.

---

\*We thank Bruno Dupire for sharing his original ideas with us, Hans-Jürgen Engelbert, Hans Föllmer, Jean Jacod, Shigeo Kusuoka, and an anonymous referee for helpful comments. R Cont is especially grateful to the late Paul Malliavin for encouraging this work.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Functional representation of non-anticipative processes</b>	<b>4</b>
2.1	Horizontal extension and vertical perturbation of a path . . . . .	4
2.2	Adapted processes as non-anticipative functionals . . . . .	4
2.3	Continuity for non-anticipative functionals . . . . .	5
2.4	Measurability properties . . . . .	6
<b>3</b>	<b>Pathwise derivatives of non-anticipative functionals</b>	<b>7</b>
3.1	Horizontal and vertical derivatives . . . . .	7
3.2	Obstructions to regularity . . . . .	10
<b>4</b>	<b>Functional Ito calculus</b>	<b>10</b>
4.1	Functional Ito formula . . . . .	10
4.2	Vertical derivative of an adapted process . . . . .	13
<b>5</b>	<b>Martingale representation formulas</b>	<b>14</b>
5.1	A martingale representation formula . . . . .	14
5.2	Extension to square-integrable functionals . . . . .	15
<b>6</b>	<b>Relation with the Malliavin derivative</b>	<b>18</b>
<b>A</b>	<b>Proof of Theorem 2.7</b>	<b>21</b>

## 1 Introduction

In the analysis of phenomena with stochastic dynamics, Ito's stochastic calculus [15, 16, 8, 23, 19, 28, 29] has proven to be a powerful and useful tool. A central ingredient of this calculus is the *Ito formula* [15, 16, 23], a change of variable formula for functions  $f(X_t)$  of a *semimartingale*  $X$  which allows to represent such quantities in terms of a stochastic integral. Given that in many applications such as statistics of processes, physics or mathematical finance, one is led to consider path-dependent functionals of a semimartingale  $X$  and its quadratic variation process  $[X]$  such as:

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

(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, 24, 22, 25, 30, 31, 32] has proven to be a powerful tool for investigating various properties of Brownian functionals. Since the construction of Malliavin derivative does not refer to an underlying filtration  $\mathcal{F}_t$ , it naturally leads to representations of functionals in terms of *anticipative* processes [4, 14, 25]. However, in most applications it is more natural to consider non-anticipative versions of such representations.

In a recent insightful work, B. Dupire [9] has proposed a method to extend the Ito formula to a functional setting in a *non-anticipative* manner, using a pathwise functional derivative which quantifies the sensitivity of a functional  $F_t : D([0, t], \mathbb{R}) \rightarrow \mathbb{R}$  to a variation in the endpoint of a path  $\omega \in D([0, t], \mathbb{R})$ :

$$\nabla_{\omega} F_t(\omega) = \lim_{\epsilon \rightarrow 0} \frac{F_t(\omega + \epsilon 1_t) - F_t(\omega)}{\epsilon}$$

Building on this insight, we develop hereafter a non-anticipative calculus [6] for a class of processes –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 path  $X_t = \{X(u), 0 \leq u \leq t\}$  of  $X$  and its quadratic variation.

Our first result (Theorem 4.1) is a change of variable formula for path-dependent functionals of the form (2). 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 [20]. This allows our result to cover a wide range of functionals, including the examples in (1).

We then extend this notion of functional derivative to *processes*: we show that for  $Y$  of the form (2) where  $F$  satisfies some regularity conditions, the process  $\nabla_X Y = \nabla_{\omega} F(X_t, A_t)$  may be defined intrinsically, independently of the choice of  $F$  in (2). The operator  $\nabla_X$  is shown to admit an extension to the space of square-integrable martingales, which is the inverse of the Ito integral with respect to  $X$ : for  $\phi \in \mathcal{L}^2(X)$ ,  $\nabla_X (\int \phi \cdot dX) = \phi$  (Theorem 5.8). In particular, we obtain a constructive version of the martingale representation theorem (Theorem 5.9), which states that for any square-integrable  $\mathcal{F}_t^X$ -martingale  $Y$ ,

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

This formula can be seen as a non-anticipative counterpart of the Clark-Haussmann-Ocone formula [4, 13, 14, 18, 25]. The integrand  $\nabla_X Y$  is an adapted process which may be computed pathwise, so this formula is more amenable to numerical computations than those based on Malliavin calculus.

Finally, we show that this functional derivative  $\nabla_X$  may be viewed as a non-anticipative "lifting" of the Malliavin derivative (Theorem 6.1): for square-integrable martingales  $Y$  whose terminal values is differentiable in the sense of Malliavin  $Y(T) \in \mathbf{D}^{1,2}$ , we show that  $\nabla_X Y(t) = E[\mathbb{D}_t H | \mathcal{F}_t]$ .

These results provide a rigorous mathematical framework for developing and extending the ideas proposed by B. Dupire [9] for a large class of functionals. In particular, unlike the results derived from the pathwise approach viewpoint presented in [5, 9], Theorems 5.8 and 5.9 do not require any pathwise regularity of the functionals and hold for non-anticipative square-integrable processes, including stochastic integrals and functionals which may depend on the quadratic variation of the process.

## 2 Functional 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  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  assumed to satisfy the usual hypotheses [8]. Denote by  $\mathcal{P}$  (resp.  $\mathcal{O}$ ) the associated *predictable* (resp. *optional*) sigma-algebra on  $[0, T]$ .  $\mathcal{F}_t^X$  denotes the ( $\mathbb{P}$ -completed) natural filtration of  $X$ . 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  $\mathbb{R}^d$ . We denote by  $[X] = ([X^i, X^j], i, j = 1..d)$  the quadratic (co-)variation process associated to  $X$ , taking values in the set  $S_d^+$  of positive  $d \times d$  matrices. We assume that

$$[X](t) = \int_0^t A(s) ds \quad (3)$$

for some cadlag process  $A$  with values in  $S_d^+$ . Note that  $A$  need not be a semimartingale. The paths of  $A$  lie in  $\mathcal{S}_t = D([0, t], S_d^+)$ , the space of cadlag functions with values  $S_d^+$ .

### 2.1 Horizontal extension and vertical perturbation of a path

Consider a path  $x \in D([0, T], \mathbb{R}^d)$  and denote by  $x_t = (x(u), 0 \leq u \leq t) \in D([0, t], \mathbb{R}^d)$  its restriction to  $[0, t]$  for  $t < T$ . 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  $[0, t]$ .

For  $h \geq 0$ , we define the *horizontal* extension  $x_{t,h} \in D([0, t+h], \mathbb{R}^d)$  of  $x_t$  to  $[0, t+h]$  as

$$x_{t,h}(u) = x(u) \quad u \in [0, t]; \quad x_{t,h}(u) = x(t) \quad u \in ]t, t+h] \quad (4)$$

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  $h$ :

$$x_t^h(u) = x_t(u) \quad u \in [0, t[ \quad x_t^h(t) = x(t) + h \quad (5)$$

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

### 2.2 Adapted processes as non-anticipative functionals

A process  $Y : [0, T] \times \Omega \mapsto \mathbb{R}^d$  adapted to  $\mathcal{F}_t^X$  may be represented as

$$Y(t) = F_t(\{X(u), 0 \leq u \leq t\}, \{A(u), 0 \leq u \leq t\}) = F_t(X_t, A_t) \quad (6)$$

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.

Since  $Y$  is non-anticipative,  $Y(t, \omega)$  only depends on the restriction  $\omega_t$  of  $\omega$  on  $[0, t]$ . This motivates the following definition:

**Definition 2.1** (Non-anticipative functional). A non-anticipative functional on  $\Upsilon$  is a family of functionals  $F = (F_t)_{t \in [0, T]}$  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 canonical filtration on  $D([0, t], \mathbb{R}^d) \times D([0, t], S_d^+)$ .

We can also view  $F = (F_t)_{t \in [0, T]}$  as a map defined on the space  $\Upsilon$  of *stopped paths*:

$$\Upsilon = \{(t, \omega_{t, T-t}), (t, \omega) \in [0, T] \times D([0, T], \mathbb{R}^d \times S_d^+)\} \quad (7)$$

Whenever the context is clear, we will denote a generic element  $(t, \omega) \in \Upsilon$  simply by its second component, the path  $\omega$  stopped at  $t$ .  $\Upsilon$  can also be identified with the ‘vector bundle’

$$\Lambda = \bigcup_{t \in [0, T]} D([0, t], \mathbb{R}^d) \times D([0, t], S_d^+). \quad (8)$$

A natural distance on the space  $\Upsilon$  of stopped paths is given by

$$d_\infty((t, \omega), (t', \omega')) = |t - t'| + \sup_{u \in [0, T]} |\omega_{t, T-t}(u) - \omega'_{t', T-t'}(u)| \quad (9)$$

$(\Upsilon, d_\infty)$  is then a metric space, a closed subspace of  $[0, T] \times D([0, T], \mathbb{R}^d \times S_d^+)$ ,  $\|\cdot\|_\infty$ ) for the product topology.

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 on  $(\Upsilon, d_\infty)$ . Introducing  $A_t$  as a second argument in the functional will allow us to control the regularity of  $Y$  with respect to  $[X]_t = \int_0^t A(u) du$  simply by requiring continuity of  $F_t$  in supremum or  $L^p$  norms with respect to the ‘lifted process’  $(X, A)$  (see Section 2.3). This idea is analogous in some ways to the approach of rough path theory [20], although here we do not resort to p-variation norms.

If  $Y$  is a  $\mathcal{B}_t$ -predictable process, then [8, Vol. I, Par. 97]

$$\forall t \in [0, T], \quad Y(t, \omega) = Y(t, \omega_{t-})$$

where  $\omega_{t-}$  denotes the path defined on  $[0, t]$  by

$$\omega_{t-}(u) = \omega(u) \quad u \in [0, t[ \quad \omega_{t-}(t) = \omega(t-)$$

Note that  $\omega_{t-}$  is cadlag and should *not* be confused with the caglad path  $u \mapsto \omega(u-)$ .

The functionals discussed in the introduction depend on the process  $A$  via  $[X] = \int_0^\cdot A(t) dt$ . In particular, they satisfy the condition  $F_t(X_t, A_t) = F_t(X_t, A_{t-})$ . Accordingly, we will assume throughout the paper that all functionals  $F_t : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t \rightarrow \mathbb{R}$  considered have ‘predictable’ dependence with respect to the second argument:

$$\forall t \in [0, T], \quad \forall (x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \quad F_t(x_t, v_t) = F_t(x_t, v_{t-}) \quad (10)$$

### 2.3 Continuity for non-anticipative functionals

We now define a notion of (left) continuity for non-anticipative functionals.

**Definition 2.2** (Continuity at fixed times). A functional  $F$  defined on  $\Upsilon$  is said to be continuous at fixed times for the  $d_\infty$  metric if and only if:

$$\forall t \in [0, T), \quad \forall \epsilon > 0, \forall (x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \quad \exists \eta > 0, (x', v') \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \\ d_\infty((x, v), (x', v')) < \eta \Rightarrow |F_t(x, v) - F_t(x', v')| < \epsilon \quad (11)$$

We now define a notion of joint continuity with respect to time and the underlying path:

**Definition 2.3** (Continuous functionals). A non-anticipative functional  $F = (F_t)_{t \in [0, T]}$  is said to be continuous at  $(x, v) \in D([0, t], \mathbb{R}^d) \times \mathcal{S}_t$  if

$$\forall \epsilon > 0, \exists \eta > 0, \forall (x', v') \in \Upsilon, \quad d_\infty((x, v), (x', v')) < \eta \Rightarrow |F_t(x, v) - F_t(x', v')| < \epsilon \quad (12)$$

We denote  $\mathbb{C}^{0,0}([0, T])$  the set of non-anticipative functionals continuous on  $\Upsilon$ .

**Definition 2.4** (Left-continuous functionals). A non-anticipative functional  $F = (F_t, t \in [0, T])$  is said to be left-continuous if for each  $t \in [0, T]$ ,  $F_t : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t \rightarrow \mathbb{R}$  in the sup norm and

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

We denote  $\mathbb{C}_l^{0,0}([0, T])$  the set of left-continuous functionals.

We define analogously the class of right continuous functionals  $\mathbb{C}_r^{0,0}([0, T])$ .

We call a functional ‘‘boundedness preserving’’ if it is bounded on each bounded set of paths:

**Definition 2.5** (Boundedness-preserving functionals). Define  $\mathbb{B}([0, T])$  as the set of non-anticipative functionals  $F$  such that for every compact subset  $K$  of  $\mathbb{R}^d$ , every  $R > 0$  and  $t_0 < T$ :

$$\exists C_{K,R,t_0} > 0, \quad \forall t \leq t_0, \forall (x, v) \in D([0, t], K) \times \mathcal{S}_t, \quad \sup_{s \in [0, t]} |v(s)| < R \Rightarrow |F_t(x, v)| < C_{K,R,t_0} \quad (14)$$

## 2.4 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  $F$ .

**Lemma 2.6** (Pathwise regularity). *If  $F \in \mathbb{C}_l^{0,0}$  then for any  $(x, v) \in D([0, T], \mathbb{R}^d) \times \mathcal{S}_T$ , the path  $t \mapsto F_t(x_{t-}, v_{t-})$  is left-continuous.*

*Proof.* Let  $F \in \mathbb{C}_l^{0,0}$  and  $t \in [0, T]$ . For  $h > 0$  sufficiently small,

$$d_\infty((x_{t-h}, v_{t-h}), (x_{t-}, v_{t-})) = \sup_{u \in (t-h, t)} |x(u) - x(t-h)| + \sup_{u \in (t-h, t)} |v(u) - v(t-h)| + h \quad (15)$$

Since  $x$  and  $v$  are cadlag, this quantity converges to 0 as  $h \rightarrow 0+$ , so

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

so  $t \mapsto F_t(x_{t-}, v_{t-})$  is left-continuous.  $\square$

**Theorem 2.7.** (i) *If  $F$  is continuous at fixed times, then the process  $Y$  defined by  $Y((x, v), t) = F_t(x_t, v_t)$  is adapted.*

(ii) *If  $F \in \mathbb{C}_l^{0,0}([0, T])$ , then the process  $Z(t) = F_t(X_t, A_t)$  is optional.*

(iii) If  $F \in \mathbb{C}_l^{0,0}([0, T])$ , and if either  $A$  is continuous or  $F$  verifies (10), then  $Z$  is a predictable process.

In particular, any  $F \in \mathbb{C}_l^{0,0}$  is a non-anticipative functional in the sense of Definition 2.1. We propose an easy-to-read proof of points (i) and (iii) in the case where  $A$  is continuous. The (more technical) proof for the cadlag case is given in the Appendix A.

*Continuous case.* Assume that  $F$  is continuous at fixed times and that the paths of  $(X, A)$  are almost-surely continuous. Let us prove that  $Y$  is  $\mathcal{F}_t$ -adapted:  $X(t)$  is  $\mathcal{F}_t$ -measurable. Introduce the partition  $t_n^i = \frac{iT}{2^n}, i = 0..2^n$  of  $[0, T]$ , as well as the following piecewise-constant approximations of  $X$  and  $A$ :

$$\begin{aligned} X^n(t) &= \sum_{k=0}^{2^n} X(t_k^n) 1_{[t_k^n, t_{k+1}^n)}(t) + X_T 1_{\{T\}}(t) \\ A^n(t) &= \sum_{k=0}^{2^n} A(t_k^n) 1_{[t_k^n, t_{k+1}^n)}(t) + A_T 1_{\{T\}}(t) \end{aligned} \quad (16)$$

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.

(i) implies point (iii) since the path of  $Z$  are left-continuous by Lemma 2.6.  $\square$

### 3 Pathwise derivatives of non-anticipative functionals

#### 3.1 Horizontal and vertical derivatives

We now define pathwise derivatives for a functional  $F = (F_t)_{t \in [0, T]} \in \mathbb{C}^{0,0}$ , following Dupire [9].

**Definition 3.1** (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

$$\mathcal{D}_t F(x, v) = \lim_{h \rightarrow 0^+} \frac{F_{t+h}(x_{t,h}, v_{t,h}) - F_t(x_t, v_t)}{h} \quad (17)$$

if the corresponding limit exists. If (17) is defined for all  $(x, v) \in \Upsilon$  the map

$$\begin{aligned} \mathcal{D}_t F : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t &\mapsto \mathbb{R}^d \\ (x, v) &\mapsto \mathcal{D}_t F(x, v) \end{aligned} \quad (18)$$

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

Note that our definition (17) is different from the one in [9] where the case  $F(x, v) = G(x)$  is considered.

Dupire [9] 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 3.2.** A non-anticipative functional  $F = (F_t)_{t \in [0, T]}$  is said to be *vertically differentiable* at  $(x, v) \in D([0, t], \mathbb{R}^d) \times D([0, t], S_d^+)$  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

$$\nabla_x F_t(x, v) = (\partial_i F_t(x, v), i = 1..d) \quad \text{where} \quad \partial_i F_t(x, v) = \lim_{h \rightarrow 0} \frac{F_t(x_t^{he_i}, v) - F_t(x, v)}{h} \quad (19)$$

is called the *vertical derivative* of  $F_t$  at  $(x, v)$ . If (19) is defined for all  $(x, v) \in \Upsilon$ , the maps

$$\begin{aligned} \nabla_x F : D([0, t], \mathbb{R}^d) \times \mathcal{S}_t &\mapsto \mathbb{R}^d \\ (x, v) &\rightarrow \nabla_x F_t(x, v) \end{aligned} \quad (20)$$

define a non-anticipative functional  $\nabla_x F = (\nabla_x F_t)_{t \in [0, T]}$ , the *vertical derivative* of  $F$ .  $F$  is then said to be *vertically differentiable* on  $\Upsilon$ .

*Remark 3.3.*  $\partial_i F_t(x, v)$  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 3.4.* 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:

$$\mathcal{D}_t F(x, v) = \partial_t f(t, X(t)) \quad \nabla_x F_t(X_t, A_t) = \nabla_x f(t, X(t)).$$

*Remark 3.5.* 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. 3.2. This notion, used in [3] to derive an integration by parts formula for pure-jump processes, is natural in the context of discontinuous semimartingales. We will show that the directional derivative (19) also intervenes naturally when the underlying process  $X$  is *continuous*, which is less obvious.

**Definition 3.6** (Regular functionals). Define  $\mathcal{C}^{1,k}([0, T])$  as the set of functionals  $F \in \mathcal{C}_l^{0,0}$  which are

- horizontally differentiable with  $\mathcal{D}_t F$  continuous at fixed times,
- $k$  times vertically differentiable with  $\nabla_x^j F \in \mathcal{C}_l^{0,0}([0, T])$  for  $j = 1..k$ .

Define  $\mathcal{C}_b^{1,k}([0, T])$  as the set of functionals  $F \in \mathcal{C}^{1,2}$  such that  $\mathcal{D}F, \nabla_x F, \dots, \nabla_x^k F \in \mathbb{B}([0, T])$ .

We denote  $\mathcal{C}^{1,\infty}([0, T]) = \cap_{k \geq 1} \mathcal{C}^{1,k}([0, T])$ .

Note that this notion of regularity only involves directional derivatives with respect to *local* perturbations of paths, so  $\nabla_x F$  and  $\mathcal{D}_t F$  seems to contain *less* information on the behavior of  $F$  than, say, the Fréchet derivative which consider perturbations in all directions in  $C_0([0, T], \mathbb{R}^d)$  or the Malliavin derivative [21, 22] which examines perturbations in the direction of all absolutely continuous functions. Nevertheless we will show in Section 4 that knowledge of  $\mathcal{D}F, \nabla_x F, \nabla_x^2 F$  along the paths of  $X$  derivatives are sufficient to reconstitute the path of  $Y(t) = F_t(X_t, A_t)$ .

*Example 1* (Smooth functions). In the case where  $F$  reduces to a smooth *function* of  $X(t)$ ,

$$F_t(x_t, v_t) = f(t, x(t)) \quad (21)$$

where  $f \in C^{1,k}([0, T] \times \mathbb{R}^d)$ , the pathwise derivatives reduces to the usual ones:  $F \in \mathbb{C}_b^{1,k}$  with:

$$\mathcal{D}_t F(x_t, v_t) = \partial_t f(t, x(t)) \quad \nabla_x^j F_t(x_t, v_t) = \partial_x^j f(t, x(t)) \quad (22)$$

In fact to have  $F \in \mathbb{C}^{1,k}$  we just need  $f$  to be right-differentiable in the time variable, with right-derivative  $\partial_t f(t, \cdot)$  which is continuous in the space variable and  $f, \nabla f$  and  $\nabla^2 f$  to be jointly left-continuous in  $t$  and continuous in the space variable.

*Example 2* (Cylindrical functionals). Let  $g \in C^0(\mathbb{R}^d, \mathbb{R}), h \in C^k(\mathbb{R}^d, \mathbb{R})$  with  $h(0) = 0$ . Then

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

is in  $\mathbb{C}_b^{1,k}$  with  $\mathcal{D}_t F(\omega) = 0$  and

$$\forall j = 1..k, \quad \nabla_\omega^j F_t(\omega) = h^{(j)}(\omega(t) - \omega(t_n-)) \quad 1_{t \geq t_n} g(\omega(t_1-), \omega(t_2-), \dots, \omega(t_n-))$$

*Example 3* (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

$$F_t(x_t, v_t) = \int_0^t g(x(u))v(u)du \quad (23)$$

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

$$\mathcal{D}_t F(x_t, v_t) = g(x(t))v(t) \quad \nabla_x^j F_t(x_t, v_t) = 0 \quad (24)$$

*Example 4*. The martingale  $Y(t) = X(t)^2 - [X](t)$  is represented by the functional

$$F_t(x_t, v_t) = x(t)^2 - \int_0^t v(u)du \quad (25)$$

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

$$\begin{aligned} \mathcal{D}_t F(x, v) &= -v(t) & \nabla_x F_t(x_t, v_t) &= 2x(t) \\ \nabla_x^2 F_t(x_t, v_t) &= 2 & \nabla_x^j F_t(x_t, v_t) &= 0, j \geq 3 \end{aligned} \quad (26)$$

*Example 5*.  $Y = \exp(X - [X]/2)$  may be represented as  $Y(t) = F(X_t)$

$$F_t(x_t, v_t) = e^{x(t) - \frac{1}{2} \int_0^t v(u)du} \quad (27)$$

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

$$\mathcal{D}_t F(x, v) = -\frac{1}{2}v(t)F_t(x, v) \quad \nabla_x^j F_t(x_t, v_t) = F_t(x_t, v_t) \quad (28)$$

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 3, 4 and 5 as a left-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 3.6. The examples below illustrate the fundamental obstructions to regularity:

*Example 6* (Delayed functionals). Let  $\epsilon > 0$ .  $F_t(x_t, v_t) = x(t - \epsilon)$  defines a  $\mathbb{C}_b^{0,\infty}$  functional. All vertical derivatives are 0. However,  $F$  fails to be horizontally differentiable.

*Example 7* (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:

$$\mathcal{D}_t F(x_t, v_t) = 0 \quad \nabla_x F_t(x_t, v_t) = 1 \quad (29)$$

However, the functional itself fails to be  $\mathbb{C}_l^{0,0}$ .

*Example 8* (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  $\mathbb{C}_l^{0,0}$  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 either right- or left-continuous so  $F$  is not  $\mathbb{C}^{0,1}$  in the sense of Definition 3.2.

*Example 9* (Maximum).  $F_t(x_t, v_t) = \sup_{s \leq t} x(s)$  is  $\mathbb{C}_l^{0,0}$  but fails to be vertically differentiable on the set

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

## 4 Functional Ito calculus

### 4.1 Functional Ito formula

We are now ready to prove our first main result, which is a change of variable formula for non-anticipative functionals of a semimartingale [6, 9]:

**Theorem 4.1.** *For any non-anticipative functional  $F \in \mathbb{C}_b^{1,2}$  verifying (10) and any  $t \in [0, T)$ ,*

$$\begin{aligned} F_t(X_t, A_t) - F_0(X_0, A_0) &= \int_0^t \mathcal{D}_u F(X_u, A_u) du + \int_0^t \nabla_x F_u(X_u, A_u) \cdot dX(u) \\ &+ \int_0^t \frac{1}{2} \text{tr}(\nabla_x^2 F_u(X_u, A_u) d[X](u)) \quad a.s. \end{aligned} \quad (30)$$

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

(30) shows that, for a regular functional  $F \in \mathbb{C}^{1,2}([0, T])$ , the process  $Y = F(X, A)$  may be reconstructed from the second-order jet  $(\mathcal{D}F, \nabla_x F, \nabla_x^2 F)$  of  $F$  along the paths of  $X$ .

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

$$\tau_0^n = 0 \quad \tau_k^n = \inf\{s > \tau_{k-1}^n | 2^n s \in \mathbb{N} \text{ or } |A(s) - A(s-)| > \frac{1}{n}\} \wedge t \quad (31)$$

The following arguments apply pathwise. Lemma A.3 ensures that

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

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

$$\begin{aligned} F_{\tau_{i+1}^n}({}_nX_{\tau_{i+1}^n-}, {}_nA_{\tau_{i+1}^n-}) - F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) &= F_{\tau_{i+1}^n}({}_nX_{\tau_{i+1}^n-}, {}_nA_{\tau_i^n, h_i^n}) - F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) \\ &+ F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) - F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) \end{aligned} \quad (32)$$

where we have used the fact that  $F$  has predictable dependence in the second variable to have  $F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) = F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-})$ . The first term in (32) can be written  $\psi(h_i^n) - \psi(0)$  where:

$$\psi(u) = F_{\tau_i^n+u}({}_nX_{\tau_i^n, u}, {}_nA_{\tau_i^n, u}) \quad (33)$$

Since  $F \in \mathbb{C}^{1,2}([0, T])$ ,  $\psi$  is right-differentiable and left-continuous by Lemma 2.6, so:

$$F_{\tau_{i+1}^n}({}_nX_{\tau_i^n, h_i^n}, {}_nA_{\tau_i^n, h_i^n}) - F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) = \int_0^{\tau_{i+1}^n - \tau_i^n} \mathcal{D}_{\tau_i^n+u} F({}_nX_{\tau_i^n, u}, {}_nA_{\tau_i^n, u}) du \quad (34)$$

The second term in (32) can be written  $\phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0)$  where  $\phi(u) = F_{\tau_i^n}({}_nX_{\tau_i^n-}^u, {}_nA_{\tau_i^n-})$ . Since  $F \in \mathbb{C}_b^{1,2}$ ,  $\phi$  is a  $C^2$  function and  $\phi'(u) = \nabla_x F_{\tau_i^n}({}_nX_{\tau_i^n-}^u, {}_nA_{\tau_i^n, h_i^n})$ ,  $\phi''(u) = \nabla_x^2 F_{\tau_i^n}({}_nX_{\tau_i^n-}^u, {}_nA_{\tau_i^n, h_i^n})$ . Applying the Ito formula to  $\phi$  between 0 and  $\tau_{i+1}^n - \tau_i^n$  and the  $(\mathcal{F}_{\tau_i^n+s})_{s \geq 0}$  continuous semimartingale  $(X(\tau_i^n + s))_{s \geq 0}$ , yields:

$$\begin{aligned} \phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0) &= \int_{\tau_i^n}^{\tau_{i+1}^n} \nabla_x F_{\tau_i^n}({}_nX_{\tau_i^n-}^{X(s) - X(\tau_i^n)}, {}_nA_{\tau_i^n-}) dX(s) \\ &+ \frac{1}{2} \int_{\tau_i^n}^{\tau_{i+1}^n} \text{tr} \left[ {}^t \nabla_x^2 F_{\tau_i^n}({}_nX_{\tau_i^n-}^{X(s) - X(\tau_i^n)}, {}_nA_{\tau_i^n-}) d[X](s) \right] \end{aligned} \quad (35)$$

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

$$\begin{aligned} F_t({}_nX_t, {}_nA_t) - F_0(X_0, A_0) &= \int_0^t \mathcal{D}_s F({}_nX_{\tau_{i(s)}^n, s - \tau_{i(s)}^n}, {}_nA_{\tau_{i(s)}^n, s - \tau_{i(s)}^n}) ds \\ &+ \int_0^t \nabla_x F_{\tau_{i(s)+1}^n}({}_nX_{\tau_{i(s)}^n-}^{X(s) - X(\tau_{i(s)}^n)}, {}_nA_{\tau_{i(s)}^n, h_{i(s)}^n}) dX(s) \\ &+ \frac{1}{2} \int_0^t \text{tr} \left[ \nabla_x^2 F_{\tau_{i(s)}^n}({}_nX_{\tau_{i(s)}^n-}^{X(s) - X(\tau_{i(s)}^n)}, {}_nA_{\tau_{i(s)}^n-}) \cdot d[X](s) \right] \end{aligned} \quad (36)$$

$F_t({}_nX_t, {}_nA_t)$  converges to  $F_t(X_t, A_t)$  almost surely. Since all approximations of  $(X, A)$  appearing in the various integrals have a  $d_\infty$ -distance from  $(X_s, A_s)$  less than  $\eta_n \rightarrow 0$ , the continuity at fixed times of  $\mathcal{D}F$  and left-continuity  $\nabla_x F, \nabla_x^2 F$  imply that 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  $n \rightarrow \infty$ . 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  $t$  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 Lebesgue-Stieltjes integrals converge almost surely, and the stochastic integral in probability, to the terms appearing in (30) as  $n \rightarrow \infty$ .

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

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

Applying the previous result to the stopped process  $(X_{t \wedge \tau_n}, A_{t \wedge \tau_n})$  and noting that, by (10),  $F_t(X_t, A_t) = F_t(X_t, A_{t-})$  leads to:

$$\begin{aligned} F_t(X_{t \wedge \tau_n}, A_{t \wedge \tau_n}) - F_0(Z_0, A_0) &= \int_0^{t \wedge \tau_n} \mathcal{D}_u F_u(X_u, A_u) du + \frac{1}{2} \int_0^{t \wedge \tau_n} \text{tr}({}^t \nabla_x^2 F_u(X_u, A_u) d[X](u)) \\ &\quad + \int_0^{t \wedge \tau_n} \nabla_x F_u(X_u, A_u) \cdot dX + \int_{t \wedge \tau_n}^t D_u F(X_{u \wedge \tau_n}, A_{u \wedge \tau_n}) du \end{aligned}$$

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

*Remark 4.2.* The above proof is probabilistic and makes use of the (classical) Ito formula [15]. In the companion paper [5] we give a non-probabilistic proof of Theorem 4.1, using the analytical approach of Föllmer [12], which allows  $X$  to have discontinuous (cadlag) trajectories.

*Example 10.* If  $F_t(x_t, v_t) = f(t, x(t))$  where  $f \in C^{1,2}([0, T] \times \mathbb{R}^d)$ , (30) reduces to the standard Itô formula.

*Example 11.* For the functional in Example 5)  $F_t(x_t, v_t) = e^{x(t) - \frac{1}{2} \int_0^t v(u) du}$ , the formula (30) yields the well-known integral representation

$$\exp(X(t) - \frac{1}{2}[X](t)) = \int_0^t e^{X(u) - \frac{1}{2}[X](u)} dX(u) \quad (37)$$

An immediate corollary of Theorem 4.1 is that, if  $X$  is a local martingale, any  $\mathbb{C}_b^{1,2}$  functional of  $X$  which has finite variation is equal to the integral of its horizontal derivative:

**Corollary 4.3.** *If  $X$  is a local martingale and  $F \in \mathbb{C}_b^{1,2}$ , the process  $Y(t) = F_t(X_t, A_t)$  has finite variation if only if  $\nabla_x F_t(X_t, A_t) = 0$   $d[X] \times d\mathbb{P}$ -almost everywhere.*

*Proof.*  $Y(t)$  is a continuous semimartingale by Theorem 4.1, with semimartingale decomposition given by (30). If  $Y$  has finite variation, then by formula (30), its continuous martingale component should be zero i.e.  $\int_0^t \nabla_x F_t(X_t, A_t) \cdot dX(t) = 0$  a.s. Computing its quadratic variation, we obtain

$$\int_0^T \text{tr}({}^t \nabla_x F_t(X_t, A_t) \cdot \nabla_x F_t(X_t, A_t) \cdot d[X]) = 0$$

which implies in particular that  $\|\partial_i F_t(X_t, A_t)\|^2 = 0$   $d[X^i] \times d\mathbb{P}$ -almost everywhere for  $i = 1..d$ . Thus,  $\nabla_x F_t(X_t, A_t) = 0$  for  $(t, \omega) \notin A \subset [0, T] \times \Omega$  where  $\int_A d[X^i] \times d\mathbb{P} = 0$  for  $i = 1..d$ .  $\square$

## 4.2 Vertical derivative of an adapted process

For a  $(\mathcal{F}_t$ -adapted) process  $Y$ , the the functional representation (42) is not unique, and the vertical  $\nabla_x F$  depends on the choice of representation  $F$ . However, Theorem 4.1 implies that the *process*  $\nabla_x F_t(X_t, A_t)$  has an intrinsic character i.e. independent of the chosen representation:

**Corollary 4.4.** *Let  $F^1, F^2 \in \mathbb{C}_b^{1,2}([0, T])$ , such that:*

$$\forall t \in [0, T], \quad F_t^1(X_t, A_t) = F_t^2(X_t, A_t) \quad \mathbb{P} - a.s. \quad (38)$$

*Then, outside an evanescent set:*

$${}^t[\nabla_x F_t^1(X_t, A_t) - \nabla_x F_t^2(X_t, A_t)]A(t-)[\nabla_x F_t^1(X_t, A_t) - \nabla_x F_t^2(X_t, A_t)] = 0 \quad (39)$$

*Proof.* Let  $X(t) = B(t) + M(t)$  where  $B$  is a continuous process with finite variation and  $M$  is a continuous local martingale. There exists  $\Omega_1 \subset \Omega$  such that  $\mathbb{P}(\Omega_1) = 1$  and for  $\omega \in \Omega$  the path of  $t \mapsto X(t, \omega)$  is continuous and  $t \mapsto A(t, \omega)$  is cadlag. Theorem 4.1 implies that the local martingale part of  $0 = F^1(X_t, A_t) - F^2(X_t, A_t)$  can be written:

$$0 = \int_0^t [\nabla_x F_u^1(X_u, A_u) - \nabla_x F_u^2(X_u, A_u)] dM(u) \quad (40)$$

Considering its quadratic variation, we have, on  $\Omega_1$

$$0 = \int_0^t \frac{1}{2} {}^t[\nabla_x F_u^1(X_u, A_u) - \nabla_x F_u^2(X_u, A_u)]A(u-)[\nabla_x F_u^1(X_u, A_u) - \nabla_x F_u^2(X_u, A_u)] du \quad (41)$$

By Lemma 2.6 ( $\nabla_x F^1(X_t, A_t) = \nabla_x F^1(X_{t-}, A_{t-})$ ) since  $X$  is continuous and  $F$  verifies (10). So on  $\Omega_1$  the integrand in (41) is left-continuous; therefore (41) implies that for  $t < T$  and  $\omega \in \Omega_1$ ,

$${}^t[\nabla_x F_u^1(X_u, A_u) - \nabla_x F_u^2(X_u, A_u)]A(u-)[\nabla_x F_u^1(X_u, A_u) - \nabla_x F_u^2(X_u, A_u)] = 0$$

□

In the case where for all  $t < T$ ,  $A(t-)$  is almost surely positive definite, Corollary 4.4 allows to define intrinsically the pathwise derivative of a process  $Y$  which admits a functional representation  $Y(t) = F_t(X_t, A_t)$ :

**Definition 4.5** (Vertical derivative of a process). Define  $\mathcal{C}_b^{1,2}(X)$  the set of  $\mathcal{F}_t$ -adapted processes  $Y$  which admit a functional representation in  $\mathbb{C}_b^{1,2}$ :

$$\mathcal{C}_b^{1,2}(X) = \{Y, \exists F \in \mathbb{C}_b^{1,2} \quad Y(t) = F_t(X_t, A_t) \quad \mathbb{P} - a.s.\} \quad (42)$$

If  $A(t)$  is non-singular i.e.  $\det(A(t)) \neq 0$   $dt \times d\mathbb{P}$  almost-everywhere then for any  $Y \in \mathcal{C}_b^{1,2}(X)$ , the predictable process:

$$\nabla_X Y(t) = \nabla_x F_t(X_t, A_t)$$

is uniquely defined up to an evanescent set, independently of the choice of  $F \in \mathbb{C}_b^{1,2}$  in the representation (42). We will call  $\nabla_X Y$  the *vertical derivative* of  $Y$  with respect to  $X$ .

In particular this construction applies to the case where  $X$  is a standard Brownian motion, where  $A = I_d$ , so we obtain the existence of a vertical derivative process for  $\mathbb{C}_b^{1,2}$  Brownian functionals:

**Definition 4.6** (Vertical derivative of non-anticipative Brownian functionals). Let  $W$  be a standard  $d$ -dimensional Brownian motion. For any  $Y \in \mathbb{C}_b^{1,2}(W)$  with representation  $Y(t) = F_t(W_t, t)$ , the predictable process

$$\nabla_W Y(t) = \nabla_x F_t(W_t, t)$$

is uniquely defined up to an evanescent set, independently of the choice of  $F \in \mathbb{C}_b^{1,2}$ .

## 5 Martingale representation formulas

Consider now the case where  $X$  is a Brownian martingale:

**Assumption 5.1.**  $X(t) = X(0) + \int_0^t \sigma(u).dW(u)$  where  $\sigma$  is a process adapted to  $\mathcal{F}_t^W$  verifying

$$\det(\sigma(t)) \neq 0 \quad dt \times d\mathbb{P} - a.e. \quad (43)$$

The functional Ito formula (Theorem 4.1) then leads to an explicit martingale representation formula for  $\mathcal{F}_t$ -martingales in  $\mathbb{C}_b^{1,2}(X)$ . This result may be seen as a non-anticipative counterpart of the Clark-Haussmann-Ocone formula [4, 25, 14] and generalizes other constructive martingale representation formulas previously obtained using Markovian functionals [7, 10, 11, 17, 26], Malliavin calculus [2, 18, 14, 25, 24] or other techniques [1, 27].

Consider an  $\mathcal{F}_T$  measurable random variable  $H$  with  $E|H| < \infty$  and consider the martingale  $Y(t) = E[H|\mathcal{F}_t]$ .

### 5.1 A martingale representation formula

If  $Y$  admits a representation  $Y(t) = F_t(X_t, A_t)$  where  $F \in \mathbb{C}_b^{1,2}$ , we obtain the following stochastic integral representation for  $Y$  in terms of its derivative  $\nabla_X Y$  with respect to  $X$ :

**Theorem 5.2.** *If  $Y(t) = F_t(X_t, A_t)$  for some functional  $F \in \mathbb{C}_b^{1,2}$ , then:*

$$Y(T) = Y(0) + \int_0^T \nabla_x F_t(X_t, A_t) dX(t) = Y(0) + \int_0^T \nabla_X Y.dX \quad (44)$$

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

*Proof.* Theorem 4.1 implies that for  $t \in [0, T)$ :

$$\begin{aligned} Y(t) = & \left[ \int_0^t \mathcal{D}_u F(X_u, A_u) du + \frac{1}{2} \int_0^t \text{tr} [{}^t \nabla_x^2 F_u(X_u, A_u) d[X](u)] \right. \\ & \left. + \int_0^t \nabla_x F_u(X_u, A_u) dX(u) \right] \end{aligned} \quad (45)$$

Given the regularity assumptions on  $F$ , the first term in this sum is a continuous process with finite variation while the second is a continuous local martingale. However,  $Y$  is a martingale and its

decomposition as sum of a finite variation process and a local martingale is unique [29]. Hence the first term is 0 and:  $Y(t) = \int_0^t F_u(X_u, A_u) dX_u$ . Since  $F \in \mathcal{C}_l^{0,0}([0, T])$   $Y(t)$  has limit  $F_T(X_T, A_T)$  as  $t \rightarrow T$ , so the stochastic integral also converges.  $\square$

*Example 12.*

If  $e^{X(t) - \frac{1}{2}[X](t)}$  is a martingale, applying Theorem 5.2 to the functional  $F_t(x_t, v_t) = e^{x(t) - \int_0^t v(u) du}$  yields the familiar formula:

$$e^{X(t) - \frac{1}{2}[X](t)} = 1 + \int_0^t e^{X(s) - \frac{1}{2}[X](s)} dX(s) \quad (46)$$

## 5.2 Extension to square-integrable functionals

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

$$\|\phi\|_{\mathcal{L}^2(X)}^2 = E \left[ \int_0^t \phi_s^2 d[X](s) \right] < \infty \quad (47)$$

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

$$\mathcal{I}^2(X) = \left\{ \int_0^\cdot \phi(t) dX(t), \phi \in \mathcal{L}^2(X) \right\} \quad (48)$$

endowed with the norm  $\|Y\|_2^2 = E[Y(T)^2]$ . The Ito integral  $I_X : \phi \mapsto \int_0^\cdot \phi_s dX(s)$  is then a bijective isometry from  $\mathcal{L}^2(X)$  to  $\mathcal{I}^2(X)$ .

We will now show that the operator  $\nabla_X : \mathcal{L}^2(X) \rightarrow \mathcal{I}^2(X)$  admits a suitable extension to  $\mathcal{I}^2(X)$  which verifies

$$\forall \phi \in \mathcal{L}^2(X), \quad \nabla_X \left( \int \phi \cdot dX \right) = \phi, \quad dt \times d\mathbb{P} - a.s. \quad (49)$$

i.e.  $\nabla_X$  is the inverse of the Ito stochastic integral with respect to  $X$ .

**Definition 5.3** (Space of test processes). The space of *test processes*  $D(X)$  is defined as

$$D(X) = \mathcal{C}_b^{1,2}(X) \cap \mathcal{I}^2(X) \quad (50)$$

Theorem 5.2 allows to define intrinsically the vertical derivative of a process in  $D(X)$  as an element of  $\mathcal{L}^2(X)$ .

**Definition 5.4.** Let  $Y \in D(X)$ , define the process  $\nabla_X Y \in \mathcal{L}^2(X)$  as the equivalence class of  $\nabla_x F_t(X_t, A_t)$ , which does not depend on the choice of the representation functional  $Y(t) = F_t(X_t, A_t)$

**Proposition 5.5** (Integration by parts on  $D(X)$ ). *Let  $Y, Z \in D(X)$ . Then:*

$$E[Y(T)Z(T)] = E \left[ \int_0^T \nabla_X Y(t) \nabla_X Z(t) d[X](t) \right] \quad (51)$$

*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 5.2 to  $Y$  and  $Z$ , we obtain

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

Applying the Ito isometry formula yields the 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 (51) 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 (51). This definition may be also viewed as a non-anticipative counterpart of Wiener-Sobolev spaces in the Malliavin calculus [22, 30].

**Definition 5.6** (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 5.7.**  $\{\nabla_X Y, Y \in D(X)\}$  is dense in  $\mathcal{L}^2(X)$  and

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

*Proof.* We first observe that the set  $U$  of ‘‘cylindrical’’ processes of the form

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

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

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

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

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

so that:

$$\nabla_x F_t(x_t, v_t) = f(x_{t_1-}, \dots, x_{t_n-})1_{t>t_n}, \nabla_x^2 F_t(x_t, v_t) = 0, \mathcal{D}_t F(x_t, v_t) = 0$$

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

Since  $I_X$  is a bijective isometry from  $\mathcal{L}^2(X)$  to  $\mathcal{I}^2(X)$ , 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 5.8** (Extension of  $\nabla_X$  to  $\mathcal{W}^{1,2}(X)$ ). *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*

$$\begin{aligned} \nabla_X : \mathcal{W}^{1,2}(X) &\mapsto \mathcal{L}^2(X) \\ \int_0^\cdot \phi.dX &\mapsto \phi \end{aligned} \quad (52)$$

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

$$\forall Z \in D(X), \quad E[Y(T)Z(T)] = E \left[ \int_0^T \nabla_X Y(t) \nabla_X Z(t) d[X](t) \right]. \quad (53)$$

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

$$\begin{aligned} I_X : \mathcal{L}^2(X) &\mapsto \mathcal{W}^{1,2}(X) \\ \phi &\mapsto \int_0^\cdot \phi.dX \end{aligned} \quad (54)$$

in the following sense:

$$\forall \phi \in \mathcal{L}^2(X), \quad \forall Y \in \mathcal{W}^{1,2}(X), \quad E[Y(T) \int_0^T \phi.dX] = E \left[ \int_0^T \nabla_X Y \phi d[X] \right] \quad (55)$$

*Proof.* Any  $Y \in \mathcal{W}^{1,2}(X)$  may be written as  $Y(t) = \int_0^t \phi(s) dX(s)$  with  $\phi \in \mathcal{L}^2(X)$ , which is uniquely defined  $d[X] \times d\mathbb{P}$  a.e. The Ito isometry formula then guarantees that (53) holds for  $\phi$ . To show that (53) uniquely characterizes  $\phi$ , consider  $\psi \in \mathcal{L}^2(X)$  which also satisfies (53), then, denoting  $I_X(\psi) = \int_0^\cdot \psi dX$  its stochastic integral with respect to  $X$ , (53) then implies that

$$\forall Z \in D(X), \quad \langle I_X(\psi) - Y, Z \rangle_{\mathcal{W}^{1,2}(X)} = E \left[ (Y(T) - \int_0^T \psi dX) Z(T) \right] = 0$$

which implies  $I_X(\psi) = Y$   $d[X] \times d\mathbb{P}$  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, the Ito integral  $I_X$  with respect to  $X$

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

admits an inverse on  $\mathcal{W}^{1,2}(X)$  which is an extension of the (pathwise) vertical derivative  $\nabla_X$  operator introduced in Definition 3.2, and

$$\forall \phi \in \mathcal{L}^2(X), \quad \nabla_X \left( \int_0^\cdot \phi dX \right) = \phi \quad (56)$$

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

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

**Theorem 5.9** (Martingale representation formula: general case). *For any square-integrable  $\mathcal{F}_t^X$ -martingale  $Y$ ,*

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

## 6 Relation with the Malliavin derivative

The above results hold in particular in the case where  $X = W$  is a Brownian motion. In this case, the vertical derivative  $\nabla_W$  may be related to the *Malliavin derivative* [22, 2, 3, 31] as follows.

Consider the canonical Wiener space  $(\Omega_0 = C_0([0, T], \mathbb{R}^d), \|\cdot\|_\infty, \mathbb{P})$  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  $E[|H|^2] < \infty$ . If  $H$  is differentiable in the Malliavin sense [2, 22, 24, 31] e.g.  $H \in \mathbf{D}^{1,2}$  with Malliavin derivative  $\mathbb{D}_t H$ , then the Clark-Haussmann-Ocone formula [25, 24] gives a stochastic integral representation of  $H$  in terms of the Malliavin derivative of  $H$ :

$$H = E[H] + \int_0^T {}^p E[\mathbb{D}_t H | \mathcal{F}_t] dW_t \quad (57)$$

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

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

Related martingale representations have been obtained under a variety of conditions [2, 7, 11, 18, 26, 24].

Denote by

- $L^2([0, T] \times \Omega)$  the set of (anticipative) processes  $\phi$  on  $[0, T]$  with  $E \int_0^T \|\phi(t)\|^2 dt < \infty$ .
- $\mathbb{D}$  the Malliavin derivative operator, which associates to a random variable  $H \in \mathbf{D}^{1,2}(0, T)$  the (anticipative) process  $(\mathbb{D}_t H)_{t \in [0, T]} \in L^2([0, T] \times \Omega)$ .

**Theorem 6.1** (Lifting theorem). *The following diagram is commutative in the sense of  $dt \times d\mathbb{P}$  equality:*

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

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

$$\forall H \in L^2(\Omega_0, \mathcal{F}_T, \mathbb{P}), \quad \nabla_W (E[H | \mathcal{F}_t]) = E[\mathbb{D}_t H | \mathcal{F}_t] \quad (58)$$

*Proof.* The Clark-Haussmann-Ocone formula [25] gives

$$\forall H \in \mathbf{D}^{1,2}, \quad H = E[H] + \int_0^T {}^p E[\mathbb{D}_t H | \mathcal{F}_t] dW_t \quad (59)$$

where  ${}^pE[\mathbb{D}_t H|\mathcal{F}_t]$  denotes the predictable projection of the Malliavin derivative. On other hand theorem 5.2 gives:

$$\forall H \in L^2(\Omega_0, \mathcal{F}_T, \mathbb{P}), \quad H = E[H] + \int_0^T \nabla_W Y(t) dW(t) \quad (60)$$

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

Thus, the conditional expectation operator (more precisely: the *predictable* projection on  $\mathcal{F}_t$  [8, Vol. I]) can be viewed as a morphism which “lifts” relations obtained in the framework of Malliavin calculus into relations between non-anticipative quantities, where the Malliavin derivative and the Skorokhod integral are replaced, respectively, by the vertical derivative  $\nabla_W$  and the Ito stochastic integral.

From a computational viewpoint, unlike the Clark-Haussmann-Ocone representation which requires to simulate the *anticipative* process  $\mathbb{D}_t H$  and compute conditional expectations,  $\nabla_X Y$  only involves non-anticipative quantities which can be computed path by path. It is thus more amenable to numerical computations. This topic is further explored in a forthcoming work.

## 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] ———, *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*, Journal of Functional Analysis, 259 (2010), pp. 1043–1072.
- [6] ———, *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] B. DUPIRE, *Functional Itô calculus*, Portfolio Research Paper 2009-04, Bloomberg, 2009.
- [10] R. J. ELLIOTT AND M. KOHLMANN, *A short proof of a martingale representation result*, Statistics & Probability Letters, 6 (1988), pp. 327–329.
- [11] P. FITZSIMMONS AND B. RAJEEV, *A new approach to the martingale representation theorem*, Stochastics, 81 (2009), pp. 467–476.

- [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] ———, *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] ———, *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] H. KUNITA AND S. WATANABE, *On square integrable martingales*, Nagoya Math. J., 30 (1967), pp. 209–245.
- [20] T. J. LYONS, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana, 14 (1998), pp. 215–310.
- [21] 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.
- [22] ———, *Stochastic analysis*, Springer, 1997.
- [23] P. MEYER, *Un cours sur les intégrales stochastiques*. Semin. Probab. X, Univ. Strasbourg 1974/75, Lect. Notes Math. 511, 245–400 (1976)., 1976.
- [24] D. NUALART, *Malliavin calculus and its applications*, vol. 110 of CBMS Regional Conference Series in Mathematics, CBMS, Washington, DC, 2009.
- [25] D. L. OCONE, *Malliavin's calculus and stochastic integral representations of functionals of diffusion processes*, Stochastics, 12 (1984), pp. 161–185.
- [26] 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, pp. 200–217.
- [27] J. PICARD, *Excursions, stochastic integrals and representation of wiener functionals*, Electronic Journal of Probability, 11 (2006), pp. 199–248.
- [28] P. E. PROTTER, *Stochastic integration and differential equations*, Springer-Verlag, Berlin, 2005. Second edition.

- [29] D. REVUZ AND M. YOR, *Continuous martingales and Brownian motion*, vol. 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, third ed., 1999.
- [30] I. SHIGEKAWA, *Derivatives of Wiener functionals and absolute continuity of induced measures*, J. Math. Kyoto Univ., 20 (1980), pp. 263–289.
- [31] D. W. STROOCK, *The Malliavin calculus, a functional analytic approach*, J. Funct. Anal., 44 (1981), pp. 212–257.
- [32] S. WATANABE, *Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels*, Ann. Probab., 15 (1987), pp. 1–39.

## A Proof of Theorem 2.7

In order to prove theorem 2.7 in the general case where  $A$  is only required to be cadlag, we need the following three lemmas. The first lemma states a property analogous to 'uniform continuity' for cadlag functions:

**Lemma A.1.** *Let  $f$  be a cadlag function on  $[0, T]$  and define  $\Delta f(t) = f(t) - f(t-)$ . Then*

$$\forall \epsilon > 0, \quad \exists \eta(\epsilon) > 0, \quad |x - y| \leq \eta \Rightarrow |f(x) - f(y)| \leq \epsilon + \sup_{t \in (x, y]} \{|\Delta f(t)|\} \quad (61)$$

*Proof.* If (61) 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,  $\sup_{t \in [u_n, v_n]} \{|\Delta f(t)|\} \geq |\Delta f(x)|$  but  $|f(u_n) - f(v_n)| \rightarrow |\Delta f(x)|$ , which results in a contradiction as well. Therefore (61) must hold.  $\square$

**Lemma A.2.** *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:*

$$\tau = \inf\{t > \sigma, \quad |V(t) - V(t-)| > \alpha\} \quad (62)$$

*is a stopping time.*

*Proof.* We can write that:

$$\{\tau \leq t\} = \bigcup_{q \in \mathbb{Q} \cap [0, t)} (\{\sigma \leq t - q\} \cap \{\sup_{t \in (t-q, t]} |V(u) - V(u-)| > \alpha\}) \quad (63)$$

and, using Lemma A.1,

$$\{\sup_{u \in (t-q, t]} |V(u) - V(u-)| > \alpha\} = \bigcup_{n_0 > 1} \bigcap_{n > n_0} \bigcup_{m \geq 1} \{\sup_{1 \leq i \leq 2^n} |V(t - q \frac{i-1}{2^n}) - V(t - q \frac{i}{2^n})| > \alpha + \frac{1}{m}\}. \quad (64)$$

$\square$

**Lemma A.3** (Uniform approximation of cadlag functions by step functions).

Let  $f \in D([0, T], \mathbb{R}^d)$  and  $\pi^n = (t_i^n)_{n \geq 1, i=0..k_n}$  a sequence of partitions ( $0 = t_0^n < t_1 < \dots < t_{k_n}^n = T$ ) of  $[0, T]$  such that:

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

$$\text{then} \quad \sup_{u \in [0, T]} |f(u) - \sum_{i=0}^{k_n-1} f(t_i^n) 1_{[t_i^n, t_{i+1}^n)}(u) + f(t_{k_n}^n) 1_{\{t_{k_n}^n\}}(u)| \xrightarrow{n \rightarrow \infty} 0 \quad (65)$$

*Proof.* Denote  $h^n = f - \sum_{i=0}^{k_n-1} f(t_i^n) 1_{[t_i^n, t_{i+1}^n)} + f(t_{k_n}^n) 1_{\{t_{k_n}^n\}}$ . Since  $f - h^n$  is piecewise constant on  $\pi^n$  and  $h^n(t_i^n) = 0$  by definition,

$$\sup_{t \in [0, T]} |h^n(t)| = \sup_{i=0..k_n-1} \sup_{[t_i^n, t_{i+1}^n)} |h^n(t)| = \sup_{t_i^n < t < t_{i+1}^n} |f(t) - f(t_i^n)|$$

Let  $\epsilon > 0$ . For  $n \geq N$  sufficiently large,  $\sup_{u \in [0, T] \setminus \pi^n} |\Delta f(u)| \leq \epsilon/2$  and  $\sup_i |t_{i+1}^n - t_i^n| \leq \eta(\epsilon/2)$  using the notation of Lemma A.1. Then, applying Lemma A.1 to  $f$  we obtain, for  $n \geq N$ ,

$$\sup_{t \in [t_i^n, t_{i+1}^n)} |f(t) - f(t_i^n)| \leq \frac{\epsilon}{2} + \sup_{t_i^n < t < t_{i+1}^n} |\Delta f(u)| \leq \epsilon.$$

□

We can now prove Theorem 2.7 in the case where  $A$  is a cadlag adapted process.

**Proof of Theorem 2.7:** Let us first show that  $F_t(X_t, A_t)$  is adapted. Define:

$$\tau_0^N = 0 \quad \tau_k^N = \inf \{ t > \tau_{k-1}^N \mid 2^N t \in \mathbb{N} \text{ or } |A(t) - A(t-)| > \frac{1}{N} \} \wedge t \quad (66)$$

From lemma A.2,  $\tau_k^N$  are stopping times. Define the following piecewise constant approximations of  $X_t$  and  $A_t$  along the partition  $(\tau_k^N, k \geq 0)$ :

$$\begin{aligned} X^N(s) &= \sum_{k \geq 0} X_{\tau_k^N} 1_{[\tau_k^N, \tau_{k+1}^N)}(s) + X(t) 1_{\{t\}}(s) \\ A^N(s) &= \sum_{k=0} A_{\tau_k^N} 1_{[\tau_k^N, \tau_{k+1}^N)}(s) + A(t) 1_{\{t\}}(s) \end{aligned} \quad (67)$$

as well as their truncations of rank  $K$ :

$${}_K X^N(s) = \sum_{k=0}^K X_{\tau_k^N} 1_{[\tau_k^N, \tau_{k+1}^N)}(s) \quad {}_K A^N(t) = \sum_{k=0}^K A_{\tau_k^N} 1_{[\tau_k^N, \tau_{k+1}^N)}(t) \quad (68)$$

Since  $({}_K X_t^N, {}_K A_t^N)$  coincides with  $(X_t^N, A_t^N)$  for  $K$  sufficiently large,

$$F_t(X_t^N, A_t^N) = \lim_{K \rightarrow \infty} F_t({}_K X_t^N, {}_K A_t^N). \quad (69)$$

The approximations  $F_t({}_K X_t^N, {}_K A_t^N)$  are  $\mathcal{F}_t$ -measurable as they are continuous functions of the random variables:

$$\{(X(\tau_k^N) 1_{\tau_k^N \leq t}, A(\tau_k^N) 1_{\tau_k^N \leq t}), k \leq K\}$$

so their limit  $F_t(X_t^N, A_t^N)$  is also  $\mathcal{F}_t$ -measurable. Thanks to Lemma A.3,  $X_t^N$  and  $A_t^N$  converge uniformly to  $X_t$  and  $A_t$ , hence  $F_t(X_t^N, A_t^N)$  converges to  $F_t(X_t, A_t)$  since  $F_t : (D([0, t], \mathbb{R}^d) \times \mathcal{S}_t, \|\cdot\|_\infty) \rightarrow \mathbb{R}$  is continuous.

To show the optionality of  $Z$  in point (ii), we will show that  $Z$  is as limit of right-continuous adapted processes. For  $t \in [0, T]$ , define  $i^n(t)$  to be the integer such that  $t \in [\frac{i^n T}{n}, \frac{(i^n+1)T}{n})$ . Define the process:  $Z_t^n = F_{\frac{(i^n(t)+1)T}{n}}(X_{\frac{i^n(t)T}{n}}, A_{\frac{i^n(t)T}{n}})$ , which is piecewise-constant and has right-continuous trajectories, and is also adapted by the first part of the theorem. Since  $F \in \mathbb{C}_l^{0,0}$ ,  $Z^n(t) \rightarrow Z(t)$  almost surely, which proves that  $Z$  is optional. Point (iii) follows from (i) and lemma 2.6, since in both cases  $F_t(X_t, A_t) = F_t(X_{t-}, A_{t-})$  hence  $Z$  has left-continuous trajectories.