Functional Itô calculus and Applications

David-Antoine FOURNIE

Supervised by
Professor Rama CONT

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2010

ABSTRACT

Functional Itô calculus and Applications

David-Antoine FOURNIE

This thesis studies extensions of the Itô calculus to a functional setting, using analytical and probabilistic methods, and applications to the pricing and hedging of derivative securities.

The first chapter develops a non-anticipative pathwise calculus for functionals of two cadlag paths, with a predictable dependence in the second one. This calculus is a functional generalization of Follmer's analytical approach to Itô calculus. An Itô-type change of variable formula is obtained for non-anticipative functionals on the space of right-continuous paths with left limits, using purely analytical methods. The main tool is the Dupire derivative, a Gateaux derivative for non-anticipative functionals on the space of right-continuous paths with left limits. Our framework implies as a special case a pathwise functional Itô calculus for cadlag semimartingales and Dirichlet processes. It is shown how this analytical Itô formula implies a probabilistic Itô formula for general cadlag semimartingales.

In the second chapter, a functional extension of the Itô formula is derived using stochastic analysis tools 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 Itô integral, and derive an integration by parts formula for Itô stochastic integrals. We show that this weak derivative may be viewed as a non-anticipative "lifting" of the Malliavin derivative. Regular functionals of an Itô martingale which have the local martingale property are characterized as solutions of a functional differential equation, for which a uniqueness result is given.

It is also shown how a simple verification theorem based on a functional version of the Hamilton-Jacobi-Bellman equation can be stated for a class of path-dependent stochastic control problems.

In the third chapter, a generalization of the martingale representation theorem is given for functionals satisfying the regularity assumptions for the functional Itô formula only in a local sense, and a sufficient condition taking the form of a functional differential equation is given for such locally regular functional to have the local martingale property. Examples are given to illustrate that the notion of local regularity is necessary to handle processes arising as the prices of financial derivatives in computational finance.

In the final chapter, functional Itô calculus for locally regular functionals is applied to the sensitivity analysis of path-dependent derivative securities, following an idea of Dupire. A general valuation functional differential equation is given, and many examples show that all usual options in local volatility model are actually priced by this equation. A definition is given for the usual sensitivities of a derivative, and a rigorous expression of the concept of $\Gamma - \Theta$ tradeoff is given. This expression is used together with a perturbation result for stochastic differential equations to give an expression for the Vega bucket exposure of a path-dependent derivative, as well as its of Black-Scholes Delta and Delta at a given skew stickiness ratio. An efficient numerical algorithm is proposed to compute these sensitivities in a local volatility model.

Table of Contents

Acknowledgments

I thank first and foremost my parents for their continuous support to pursue my dreams. 9 years of studies would never have been possible without their moral, practical and financial support.

I thank a lot Amal Moussa for her great friendship which supported me throughout all my years of studies in Toulouse, Paris and New York.

I wish to thank deeply my advisor Professor Rama Cont for his constant support and guidance during my PhD studies at Columbia University, and for his deep involvement in our joint work which is compiled in this dissertation.

I am very grateful to Professor Ioannis Karatzas for my recruitment at Columbia University and his support during my PhD studies. Most of my knowledge and interest in Probability and Stochastic Analysis comes from his books and his lectures.

I thank Bruno Dupire whose original ideas and thoughts opened the way to my research work.

I thank Professor Nicole El Karoui from the Ecole Polytechnique who encouraged me to go for a PhD at Columbia University, and my Mathematics and Physics Professors in "Classes Préparatoires" Pierre Gissot, Michel "Jimmy" Gonnord and Jean-Marie Mercier to whom I owe my interest in Mathematics and Science in general.

Last but not least, all my gratitude goes to the administrative team at the Mathematics Department, and especially Terrance Cope, for their great work and constant help in all issues related to my studies here.

To my parents

Chapter 1

Synopsis

Itô's stochastic calculus [36, 37, 18, 49, 45, 54] has proven to be a powerful and useful tool in analyzing phenomena involving random, irregular evolution in time.

Two characteristics distinguish the Itô 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 [21] 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 **Itô formula** [36, 37, 49] 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 Itô 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 Itô 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 naturally led to consider functionals of a semimartingale X and its quadratic variation process $[X]$ such as:

$$\int_0^t g(t, X_t) d[X](t), \qquad G(t, X_t, [X]_t), \quad \text{or} \qquad E[G(T, X(T), [X](T))|\mathcal{F}_t] \tag{1.1}$$

(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 [7, 9, 50, 48, 51, 56, 59] 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 [9, 34, 51], whereas in applications it is more natural to consider non-anticipative, or causal, versions of such representations.

In a recent insightful work, B. Dupire [23] has proposed a method to extend the Itô formula to a functional setting in a *non-anticipative* manner. Building on this insight, we develop hereafter a non-anticipative calculus for a class of functionals -including the above examples- which may be represented as

$$Y(t) = F_t(\{X(u), 0 \le u \le t\}, \{A(u), 0 \le u \le t\}) = F_t(X_t, A_t) \tag{1.2}$$

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

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

represents the dependence of Y on the underlying path and its quadratic variation. For such functionals, we define an appropriate notion of regularity (Section 2.2.2) and a non-anticipative notion of pathwise derivative (Section 2.3). 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 [46]. This allows to consider a wider range of functionals, as in (1.1).

Using these pathwise derivatives, we derive in chapter 2 a purely analytical Itô formula for functionals of two variable (x, v) which are cadlag paths, and the functional has a predictable dependence in the second variable. This formula (Theorems 2.4, 2.5) is preceded by other useful analytical results on the class of functionals that we consider. Our method follows the spirit of H. Föllmer's [29] pathwise approach to Itô calculus, where the term taking the place of the stochastic integral in stochastic calculus is defining as a diagonal limit of discretized integrals with integrand evaluated on discretized paths. It is then show in section 2.6 that this analytical formula allows to define a pathwise notion of stochastic integral for functionals of a cadlag Dirichlet processes, for which a functional Itô formula is derived, and in section 2.7 it is shown that it implies an Itô formula for functionals of cadlag semimartingales. In the case of continuous semimartingales, we provide an alternative derivation (2.4) under different regularity assumptions.

In chapter 3, a direct probabilistic derivation of the functional Itô formula is provided for continuous semimartingales. We then use this functional Itô formula to derive a constructive version of the martingale representation theorem (Section 3.3), which can be seen as a non-anticipative form of the Clark-Haussmann-Ocone formula [9, 33, 34, 51]. The martingale representation formula allows to obtain an integration by parts formula for Itô stochastic integrals (Theorem 3.4), which enables in turn to define a weak functional derivative, for a class of stochastic integrals (Section 3.4). We argue that this weak derivative may be viewed as a non-anticipative "lifting" of the Malliavin derivative (Theorem 3.6). We then show that regular functionals of an Itô martingale which have the local martingale property are characterized as solutions of a functional analogue of Kolmogorov's backward equation (Section 3.5), for which a uniqueness result is given (Theorem 3.8). Finally, in section 3.6, we present as a potential direction for further research a setting of path-dependent stochastic control problem, with dependence of the coefficients of the diffusion as well as the objective and cost functions on the whole path of the controlled process, and eventually on the path of its quadratic variation. We are able to prove two versions of a verification theorem based on a functional version of the Hamilton-Jacobi-Bellman equation, theorems 3.9 and 3.10, depending on whether or not there is explicit dependence on the quadratic variation of the controlled process.

In chapter 4, we present a local version, using optional times, of some results from chapter 3, especially the martingale representation theorem 4.6, in the sense that they apply to functionals defined on *continuous* paths which can be extended only locally to regular functions of cadlag paths. A sufficient condition, taking the form of a functional differential equation, is given on such functionals for defining local martingales (Theorem 4.7). This extension is motivated by examples of functionals which define martingales and do satisfy a functional differential equation, but fail to satisfy the regularity assumptions of chapter 3. The choice of the examples come from processes traditionally encountered in Mathematical Finance, and hence show that chapter 4 is necessary in order to use the functional setting for the sensitivity analysis of path-dependent derivatives (Chapter 5).

In the final chapter, building on Dupire's original insight [23], we show how the setting of functional Itô calculus is a natural formalism for the hedging of path-dependent derivatives, emphasizing the notion of *sensitivity* of the option price to the underlyings and to market variables. A valuation functional differential equation (Theorem 5.1) is derived, and shows that the theoretical replication portfolio of the derivative is the portfolio hedging the directional sensitivity. This theorem also extends the classical relationship between the sensitivites of vanilla options to general path-dependent payoffs, and hence gives a precise meaning to the concept of *Theta - Gamma* tradeoff which is familiar to derivatives traders (Theorem 5.2). We then show that the valuation functional equation applies to most payoffs encountered in the markets; in particular the different classical PDEs satisfied by the prices of Vanilla, Barrier, Asian, Variance Swap options are actually shown to be particular cases of this universal functional equation, which can therefore be seen as a unified description for derivatives pricing. We then build on the functional setting to give expressions for the sensitivities of the derivative to observable market variables, such as the Vega bucket exposure (Section 5.4.2), which are the main tool for a volatility trader to understand his exposure. We also provide with expressions for the Black-Scholes Delta and Delta at a given skew stickiness ratio (Section 5.4.3), which are used by practitioners to hedge their portfolio. We then proceed to suggest an efficient numerical algorithm for the computation of these sensitivities. This final chapter may be seen as an attempt to formalize concepts used in derivatives trading from the point of view of a sell-side traders.

Appendix A contains numerous technical lemmas used in the proofs in chapter 2, more precisely results on the approximation of cadlag functions by piecewise-constant functions and measure-theoretic results. Appendix B is a self-contained digression on strong solutions for stochastic differential equations with functional coefficients, which is the setting we need for the section on stochastic control (Section 3.6). In particular, it defines a concept of strong solution starting from a given initial value, in the case where the coefficient have an explicit dependence on the path of the quadratic variation (Definition B.2). Continuity properties of the solution in the initial value is also investigated (Section B.1.2), as well as perturbation of the coefficient (Section B.1.3). This last perturbation result has application in the computation of Vega bucket exposure and Deltas in sections 5.4.2 and 5.4.3.

Chapter 2

Pathwise calculus for non-anticipative functionals

2.1 Motivation

In his seminal paper *Calcul d'Ito sans probabilités* [29], Hans Föllmer proposed a non-probabilistic version of the Ito formula [36]: Föllmer showed that if a real-valued cadlag (right continuous with left limits) function x has finite quadratic variation along a sequence $\pi_n = (t_k^n)_{k=0..n}$ of subdivisions of $[0,T]$ with step size decreasing to zero, in the sense that the sequence of discrete measures

$$\sum_{k=0}^{n-1} |x(t_{k+1}^n) - x(t_k^n)|^2 \delta_{t_k^n}$$

converges vaguely to a Radon measure with Lebesgue decomposition $\xi + \sum_{t\in[0,T]} |\Delta x(t)|^2 \delta_t$ then for $f \in C^1(\mathbb{R})$ one can define the pathwise integral

$$\int_0^T f(x(t)) d^\pi x = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(x(t_i^n)).(x(t_{i+1}^n) - x(t_i^n)) \tag{2.1}$$

as a limit of Riemann sums along the subdivision $\pi = (\pi_n)_{n\geq 1}$. In particular if $X = (X_t)_{t\in[0,T]}$ is a semimartingale [18, 49, 54], which is the classical setting for stochastic calculus, the paths of X have almost surely finite quadratic variation along such subsequences: when applied to the paths of X, Föllmer's integral (2.1) then coincides, with probability one, with the Ito stochastic integral $\int_0^T f(X)dX$ with respect to the semimartingale X. This construction may in fact be carried out for a more general class of processes, including the class of Dirichlet processes [12, 29, 30, 47].

Of course, the Ito stochastic integral with respect to a semimartingale X may be defined for a much larger class of integrands: in particular, for a cadlag process Y defined as a *non-anticipative functional* $Y(t) = F_t(X(u), 0 \leq u \leq t)$ of X, the stochastic integral $\int_0^T Y dX$ may be defined as a limit of non-anticipative Riemann sums [54].

Using a notion of directional derivative for functionals proposed by Dupire [23], we extend Föllmer's pathwise change of variable formula to non-anticipative functionals on the space $D([0,T],\mathbb{R}^d)$ of cadlag paths (Theorem 2.4). The requirement on the functionals is to possess certain directional derivatives which may be computed pathwise. Our construction allows to define a pathwise integral $\int F_t(x)dx$, defined as a limit of Riemann sums, for a class of functionals F of a cadlag path x with finite quadratic variation. Our results lead to functional extensions of the Ito formula for semimartingales (Section 2.7) and Dirichlet processes (Section 2.6). In particular, we show the stability of the the class of semimartingales under functional transformations verifying a regularity condition. These results yield a non-probabilistic proof for functional Ito formulas obtained in [23] using probabilistic methods and extend them to the case of discontinuous semimartingales.

Notation

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 stochastic 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 $[0,t]$.

2.2 Non-anticipative functionals on spaces of paths

Let $T > 0$, and $U \subset \mathbb{R}^d$ be an open subset of $\mathbb{R}^d$ and $S \subset \mathbb{R}^m$ be a Borel subset of $\mathbb{R}^m$. We call "U-valued cadlag function" a right-continuous function $f : [0,T] \mapsto U$ with left limits such that for each $t \in]0,T]$, $f(t-) \in U$. Denote by $\mathcal{U}_t = D([0,t],U)$ (resp. $\mathcal{S}_t = D([0,t],S)$) the space of U-valued cadlag functions (resp. S), and $C_0([0,t],U)$ the set of continuous functions with values in U.

When dealing with functionals of a path $x(t)$ indexed by time, an important class is formed by those which are *non-anticipative*, in the sense that they only depend on the past values of x. A family $Y : [0,T] \times \mathcal{U}_T \mapsto \mathbb{R}$ of functionals is said to be *non-anticipative* if, for all $(t,x) \in [0,T] \times \mathcal{U}_T$, $Y(t,x) = Y(t,x_t)$ where $x_t = x_{|[0,t]}$ denotes the restriction of the path x to $[0,t]$. A non-anticipative functional may thus be represented as $Y(t,x) = F_t(x_t)$ where $(F_t)_{t\in[0,T]}$ is a family of maps $F_t : \mathcal{U}_t \mapsto \mathbb{R}$. This motivates the following definition:

Definition 2.1 (Non-anticipative functionals on path space)**.** A non-anticipative functional on $\mathcal{U}_T$ is a family $F = (F_t)_{t\in[0,T]}$ of maps

$$F_t : \mathcal{U}_t \to \mathbb{R}$$

Y is said to be *predictable*[1] if, for all $(t,x) \in [0,T] \times \mathcal{U}_T$, $Y(t,x) = Y(t,x_{t-})$ where x_{t-} denotes the function defined on $[0,t]$ by

$$x_{t-}(u) = x(u) \quad u \in [0,t[\qquad x_{t-}(t) = x(t-)$$

Typical examples of predictable functionals are integral functionals, e.g.

$$Y(t,x) = \int_0^t G_s(x_s)ds$$

where G is a non-anticipative, locally integrable, functional.

If Y is predictable then Y is non-anticipative, but predictability is a stronger property. Note that x_{t-} is cadlag and should *not* be confused with the caglad path $u \mapsto x(u-)$.

We consider throughout this work non-anticipative functionals

$$F = (F_t)_{t\in[0,T]} \qquad F_t : \mathcal{U}_t \times \mathcal{S}_t \to \mathbb{R}$$

where F has a predictable dependence with respect to the second argument:

$$\forall t \le T, \quad \forall (x,v) \in \mathcal{U}_t \times \mathcal{S}_t, \quad F_t(x_t, v_t) = F_t(x_t, v_{t-}) \tag{2.2}$$

F can be viewed as a functional on the vector bundle $\Upsilon = \bigcup_{t\in[0,T]} \mathcal{U}_t \times \mathcal{S}_t$. We will also consider non-anticipative functionals $F = (F_t)_{t\in[0,T[}$ indexed by $[0,T[$.

[1] This notion coincides with the usual definition of predictable process when the path space $\mathcal{U}_T$ is endowed with the filtration of the canonical process, see Dellacherie & Meyer [18, Vol. I].

2.2.1 Horizontal and vertical perturbation of a path

Consider a path $x \in D([0,T]),U)$ and denote by $x_t \in \mathcal{U}_t$ 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

$$x_{t,h}(u) = x(u) \qquad u \in [0,t] \ ; \qquad x_{t,h}(u) = x(t) \qquad u \in]t,t+h] \tag{2.3}$$

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

$$x_t^h(u) = x_t(u) \qquad u \in [0,t[\quad x_t^h(t) = x(t) + h \tag{2.4}$$

or in other words $x_t^h(u) = x_t(u) + h1_{t=u}$. By convention, $x_{t,h}^u = (x_t^u)_{t,h}$, ie the vertical perturbation precedes the horizontal extension.

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

$$d_\infty(\ (x,v),(x',v')\ ) = \sup_{u\in[0,t+h]} |x_{t,h}(u) - x'(u)| + \sup_{u\in[0,t+h]} |v_{t,h}(u) - v'(u)| + \ h \tag{2.5}$$

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.

2.2.2 Classes of non-anticipative functionals

Using the distance d_∞ defined above, we now introduce various notions of continuity for non-anticipative functionals.

Definition 2.2 (Continuity at fixed times)**.** A non-anticipative functional $F = (F_t)_{t\in[0,T]}$ is said to be continuous at fixed times if for any $t \leq T$, $F_t : \mathcal{U}_t \times \mathcal{S}_t \mapsto \mathbb{R}$ is continuous for the supremum norm.

Definition 2.3 (Left-continuous functionals)**.** Define $\mathbb{F}_l^\infty$ as the set of functionals $F = (F_t, t \in [0,T])$ which satisfy:

$$\forall t \in [0,T], \quad \forall \epsilon > 0, \forall (x,v) \in \mathcal{U}_t \times \mathcal{S}_t, \quad \exists \eta > 0, \forall h \in [0,t],$$
$$\forall (x',v') \in \mathcal{U}_{t-h} \times \mathcal{S}_{t-h}, \qquad d_\infty((x,v),(x',v')) < \eta \Rightarrow |F_t(x,v) - F_{t-h}(x',v')| < \epsilon \tag{2.6}$$

Definition 2.4 (Right-continuous functionals)**.** Define $\mathbb{F}_r^\infty$ as the set of functionals $F = (F_t, t \in [0,T[)$ which satisfy:

$$\forall t \in [0,T], \quad \forall \epsilon > 0, \forall (x,v) \in \mathcal{U}_t \times \mathcal{S}_t, \quad \exists \eta > 0, \forall h \in [0, T-t],$$
$$\forall (x',v') \in \mathcal{U}_{t+h} \times \mathcal{S}_{t+h}, \qquad d_\infty((x,v),(x',v')) < \eta \Rightarrow |F_t(x,v) - F_{t+h}(x',v')| < \epsilon \quad (2.7)$$

We denote $\mathbb{F}^\infty = \mathbb{F}_r^\infty \cap \mathbb{F}_l^\infty$ the set of continuous non-anticipative functionals.

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

Definition 2.5 (Boundedness-preserving functionals)**.** Define $\mathbb{B}$ as the set of non-anticipative functionals F such that for every compact subset K of U, every $R > 0$, there exists a constant $C_{K,R}$ such that:

$$\forall t \leq T, \forall (x,v) \in D([0,t],K) \times \mathcal{S}_t, \sup_{s\in[0,t]} |v(s)| < R \Rightarrow |F_t(x,v)| < C_{K,R} \quad (2.8)$$

In particular if $F \in \mathbb{B}$, it is "locally" bounded in the neighborhood of any given path i.e.

$$\forall (x,v) \in \mathcal{U}_T \times \mathcal{S}_T, \quad \exists C > 0, \eta > 0, \qquad \forall t \in [0,T], \quad \forall (x',v') \in \mathcal{U}_t \times \mathcal{S}_t,$$
$$d_\infty((x_t,v_t),(x',v')) < \eta \Rightarrow \forall t \in [0,T], |F_t(x',v')| \leq C \quad (2.9)$$

The following lemma shows that a continuous functional also satisfies the local boundedness property (2.9).

Lemma 2.1. *If $F \in \mathbb{F}^\infty$, then it satisfies the property of local boundedness 2.9.*

Proof. Let $F \in \mathbb{F}^\infty$ and $(x,v) \in \mathcal{U}_T \times \mathcal{S}_T$. For each $t < T$, there exists $\eta(t)$ such that, for all $t' < T$, $(x',v') \in \mathcal{U}_{t'} \times \mathcal{S}_{t'}$:

$$d_\infty((x_t,v_t),(x',v')) < \eta(t) \Rightarrow |F_t(x_t,v_t) - F_{t'}(x',v')| < 1$$
$$d_\infty((x_{t-},v_{t-}),(x',v')) < \eta(t) \Rightarrow |F_t(x_{t-},v_{t-}) - F_{t'}(x',v')| < 1 \quad (2.10)$$

Since (x, v) has cadlag trajectories, for each $t < T$, there exists $\epsilon(t)$ such that:

$$\forall t' > t, |t' - t| < \epsilon(t) \Rightarrow d_\infty((x_{t'}, v_{t'}), (x_t, v_t)) < \frac{\eta(t)}{2}$$
$$\forall t' < t, |t' - t| < \epsilon(t) \Rightarrow d_\infty((x_{t'}, v_{t'}), (x_{t-}, v_{t-})) < \frac{\eta(t)}{2} \tag{2.11}$$

Therefore one can extract a finite covering of compact set $[0, T]$ by such intervals

$$[0, T] \subset \bigcup_{j=1}^{N} (x_{t_j} - \epsilon(t_j), x_{t_j} + \epsilon(t_j)) \tag{2.12}$$

Let $t < T$ and $(x', v') \in \mathcal{U}_t \times \mathcal{S}_t$. Assume that:

$$d_\infty((x_t, v_t), (x', v')) < \min_{1 \le j \le N} \frac{\eta(t_j)}{2} \tag{2.13}$$

$t \in (x_{t_j} - \epsilon(t_j), x_{t_j} + \epsilon(t_j))$ for some $j \le N$. If $t < t_j$, then:

$$d_\infty((x_{t_j-}, v_{t_j-}), (x', v')) < d_\infty((x_t, v_t), (x', v')) + d_\infty((x_t, v_t), (x_{t_j-}, v_{t_j-})) \tag{2.14}$$

where both terms in the sum are less than $\frac{\eta(t_j)}{2}$, so that:

$$|F_t(x', v')| < |F_{t_j}(x_{t_j-}, v_{t_j-})| + 1 \tag{2.15}$$

If $t \ge t_j$, then:

$$d_\infty((x_{t_j}, v_{t_j}), (x', v')) < d_\infty((x_t, v_t), (x', v')) + d_\infty((x_t, v_t), (x_{t_j}, v_{t_j})) \tag{2.16}$$

where both terms in the sum are less than $\frac{\eta(t_j)}{2}$, so that:

$$|F_t(x', v')| < |F_{t_j}(x_{t_j}, v_{t_j})| + 1 \tag{2.17}$$

so that in any case:

$$|F_t(x', v')| < \max_{1 \le j \le N} \max(|F_{t_j}(x_{t_j-}, v_{t_j-})|, |F_{t_j}(x_{t_j}, v_{t_j})|) + 1 \tag{2.18}$$

□

The following result describes the behavior of paths generated by the functionals in the above classes:

Proposition 2.1 (Pathwise regularity)**.**

1. *If $F \in \mathbb{F}_l^\infty$ then for any $(x, v) \in \mathcal{U}_T \times \mathcal{S}_T$, the path $t \mapsto F_t(x_{t-}, v_{t-})$ is left-continuous.*

2. *If $F \in \mathbb{F}_r^\infty$ then for any $(x, v) \in \mathcal{U}_T \times \mathcal{S}_T$, the path $t \mapsto F_t(x_t, v_t)$ is right-continuous.*

3. *If $F \in \mathbb{F}^\infty$ then for any $(x, v) \in \mathcal{U}_T \times \mathcal{S}_T$, the path $t \mapsto F_t(x_t, v_t)$ is cadlag and continuous at all points where x and v are continuous.*

4. *If $F \in \mathbb{F}^\infty$ further verifies (2.2) then for any $(x, v) \in \mathcal{U}_T \times \mathcal{S}_T$, the path $t \mapsto F_t(x_t, v_t)$ is cadlag and continuous at all points where x is continuous.*

5. *If $F \in \mathbb{B}$, then for any $(x, v) \in \mathcal{U}_T \times \mathcal{S}_T$, the path $t \mapsto F_t(x_t, v_t)$ is bounded.*

Proof. 1. Let $F \in \mathbb{F}_l^\infty$ and $t \in [0, T)$. For $h > 0$ sufficiently small,

$$d_\infty((x_{t-h}, v_{t-h}), (x_{t-}, v_{t-})) = \sup_{u \in (t-h,t)} |x(u) - x(t-)| + \sup_{u \in (t-h,t)} |v(u) - v(t-)| + h \tag{2.19}$$

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

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

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

2. Let $F \in \mathbb{F}_r^\infty$ and $t \in [0, T)$. For $h > 0$ sufficiently small,

$$d_\infty((x_{t+h}, v_{t+h}), (x_t, v_t)) = \sup_{u \in [t,t+h)} |x(u) - x(t)| + \sup_{u \in [t,t+h)} |v(u) - v(t)| + h \tag{2.20}$$

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

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

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

3. Assume now that F is in $\mathbb{F}^\infty$ and let $t \in]0, T]$. Denote $(\Delta x(t), \Delta v(t))$ the jump of (x, v) at time t. Then

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

and this quantity goes to 0 because x and v have left limits. Hence the path has left limit $F_t(x_t^{-\Delta x(t)}, v_t^{-\Delta v(t)})$ at t. A similar reasoning proves that it has right-limit $F_t(x_t, v_t)$.

4. If $F \in \mathbb{F}^\infty$ verifies (2.2), for $t \in]0,T]$ the path $t \mapsto F_t(x_t, v_t)$ has left-limit $F_t(x_t^{-\Delta x(t)}, v_t^{-\Delta v(t)})$ at t, but (2.2) implied that this left-limit equals $F_t(x_t^{-\Delta x(t)}, v_t)$.

□

2.2.3 Measurability properties

Consider, on the path space $\mathcal{U}_T \times \mathcal{S}_T$, endowed with the supremum norm and its Borel σ-algebra, the filtration $(\mathcal{F}_t)$ generated by the canonical process

$$\begin{aligned} (X,V) : \mathcal{U}_T \times \mathcal{S}_T \times [0,T] &\to U \times S \\ (x,v), t &\mapsto (X,V)((x,v),t) = (x(t), v(t)) \end{aligned} \tag{2.21}$$

$\mathcal{F}_t$ is the smallest sigma-algebra on $\mathcal{U}_T \times \mathcal{S}_T$ such that all coordinate maps $(X(.,s), V(.,s))$, $s \in [0,t]$ are $\mathcal{F}_t$-measurable.

The *optional* sigma-algebra $\mathcal{O}$ on $\mathcal{U}_T \times \mathcal{S}_T \times [0,T]$ is the sigma-algebra on $\mathcal{U}_T \times \mathcal{S}_T \times [0,T]$ generated by all mappings $f : \mathcal{U}_T \times \mathcal{S}_T \times [0,T] \to \mathbb{R}$ the set into which, for every $\omega \in \mathcal{U}_T \times \mathcal{S}_T$, are right continuous in t, have limits from the left and are adapted to $(\mathcal{F}_t)_{t\in[0,T]}$. The *predictable* sigma-algebra $\mathcal{P}$ is the sigma-algebra on $\mathcal{U}_T \times \mathcal{S}_T \times [0,T]$ generated by all mappings $f : \mathcal{U}_T \times \mathcal{S}_T \times [0,T] \to \mathbb{R}$ the set into which, for every $\omega \in \mathcal{U}_T \times \mathcal{S}_T$, are left-continuous in t and are adapted to $(\mathcal{F}_t)_{t\in[0,T]}$. A positive map $\tau : \mathcal{U}_T \times \mathcal{S}_T \to [0,\infty[$ is called an *optional time* if $\{\omega \in \mathcal{U}_T \times \mathcal{S}_T, \quad \tau(\omega) < t\} \in \mathcal{F}_t$ for every $t \in [0,T]$.

The following result, proved in Appendix A.2, clarifies the measurability properties of processes defined by functionals in $\mathbb{F}_l^\infty, \mathbb{F}_r^\infty$:

Theorem 2.1. *If F is continuous at fixed time, then the process Y defined by $Y((x,v),t) = F_t(x_t, v_t)$ is $\mathcal{F}_t$-adapted. If $F \in \mathbb{F}_l^\infty$ or $F \in \mathbb{F}_r^\infty$, then:*

1. *the process Y defined by $Y((x,v),t) = F_t(x_t, v_t)$ is optional i.e. $\mathcal{O}$-measurable.*

2. *the process Z defined by $Z((x,v),t) = F_t(x_{t-}, v_{t-})$ is predictable i.e. $\mathcal{P}$-measurable.*

2.3 Pathwise derivatives of non-anticipative functionals

2.3.1 Horizontal derivative

We now define a pathwise derivative for a non-anticipative functional $F = (F_t)_{t\in[0,T]}$, which may be seen as a "Lagrangian" derivative along the path x.

Definition 2.6 (Horizontal derivative)**.** The *horizontal derivative* at $(x, v) \in \mathcal{U}_t \times \mathcal{S}_t$ of a non-anticipative functional $F = (F_t)_{t\in[0,T[}$ is defined as

$$\mathcal{D}_t F(x, v) = \lim_{h\to 0^+} \frac{F_{t+h}(x_{t,h}, v_{t,h}) - F_t(x, v)}{h} \tag{2.22}$$

if the corresponding limit exists. If (2.22) is defined for all $(x, v) \in \Upsilon$ the map

$$\begin{aligned} \mathcal{D}_t F : \mathcal{U}_t \times \mathcal{S}_t &\to \mathbb{R}^d \\ (x, v) &\mapsto \mathcal{D}_t F(x, v) \end{aligned} \tag{2.23}$$

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

We will occasionally use the following "local Lipschitz property" that is weaker than horizontal differentiability:

Definition 2.7. A non-anticipative functional F is said to have the horizontal local Lipschitz property if and only if:

$$\forall (x, v) \in \mathcal{U}_T \times \mathcal{S}_T, \exists C > 0, \eta > 0, \forall t_1 < t_2 \le T, \forall (x', v') \in \mathcal{U}_{t_1} \times \mathcal{S}_{t_1},$$

$$d_\infty((x_{t_1}, v_{t_1}), (x', v')) < \eta \Rightarrow |F_{t_2}(x'_{t_1,t_2-t_1}, v'_{t_1,t_2-t_1}) - F_{t_1}((x'_{t_1}, v'_{t_1}))| < C(t_2 - t_1) \tag{2.24}$$

2.3.2 Vertical derivative

Dupire [23] introduced a pathwise spatial derivative for non-anticipative functionals, which we now introduce. Denote $(e_i, i = 1..d)$ the canonical basis in $\mathbb{R}^d$.

Definition 2.8. 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

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

is differentiable at 0. Its gradient at 0

$$\nabla_x F_t\ (x,v) = (\partial_i F_t(x,v),\ i=1..d) \qquad \text{where} \quad \partial_i F_t(x,v) = \lim_{h\to 0} \frac{F_t(x_t^{he_i}, v) - F_t(x,v)}{h} \tag{2.25}$$

is called the *vertical derivative* of F_t at (x,v). If (2.25) is defined for all $(x,v) \in \Upsilon$, the *vertical derivative*

$$\begin{aligned} \nabla_x F : \mathcal{U}_t \times \mathcal{S}_t &\to \mathbb{R}^d \\ (x,v) &\mapsto \nabla_x F_t(x,v) \end{aligned} \tag{2.26}$$

define a non-anticipative functional $\nabla_x F = (\nabla_x F_t)_{t\in[0,T]}$ with values in $\mathbb{R}^d$.

Remark 2.1. If a vertically differentiable functional is predictable with respect to the second variable ($F_t(x_t, v_t) = F_t(x_t, v_{t-})$), so is its vertical derivative.

Remark 2.2. $\partial_i F_t(x,v)$ 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 2.3. 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:

$$\mathcal{D}_t F(x,v) = \partial_t f(t,x(t)) \qquad \nabla_x F_t(x_t, v_t) = \nabla_x f(t, x(t)).$$

Remark 2.4. Note that the assumption (2.2) that F is predictable with respect to the second variable entails that for any $t \in [0,T]$, $F_t(x_t, v_t^e) = F_t(x_t, v_t)$ so an analogous notion of derivative with respect to v would be identically zero under assumption (2.2).

If F admits a horizontal (resp. vertical) derivative $\mathcal{D}F$ (resp. $\nabla_x F$) we may iterate the operations described above and define higher order horizontal and vertical derivatives.

Definition 2.9. Define $\mathbb{C}^{j,k}$ as the set of functionals F which are

- continuous at fixed times,
- admit j horizontal derivatives and k vertical derivatives at all $(x,v) \in \mathcal{U}_t \times \mathcal{S}_t$, $t \in [0,T[$
- $\mathcal{D}^m F, m \le j, \nabla_x^n F, n \le k$ are continuous at fixed times.

2.3.3 Uniqueness results for vertical derivatives

The analytical Itô formula for continuous paths (theorem 2.4), and its probabilistic counterpart (theorem 3.1), refer explicitly to the vertical derivatives of the functional F, which requires the functional to be defined on *cadlag* path although its argument x is *continuous*. Since a functional defined on continuous paths could be extended to Υ in multiple ways, the vertical derivative and therefore Itô's formula seem to depend on the chosen extension. The following two theorems 2.2 and 2.3 show that this is indeed not the case, as the value of the vertical derivatives on continuous paths do not depend on the chosen extension.

Theorem 2.2. *If $F^1, F^2 \in \mathbb{C}^{1,1}$, with $F^i, \nabla_x F^i \in \mathbb{F}_l^\infty$ and $\mathcal{D}F^i$ satisfying the local boundedness assumption 2.9 for $i = 1, 2$, coincide on continuous paths:*

$$\forall t \in]0,T] \quad \forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T, \qquad F_t^1(x_t, v_t) = F_t^2(x,v)$$

$$\text{then} \qquad \forall t \in]0,T], \quad \forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T, \nabla_x F_t^1(x_t, v_{t-}) = \nabla_x F_t^2(x_t, v_{t-})$$

Proof. Let $F = F^1 - F^2 \in \mathbb{C}^{1,1}$ and $(x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T$. Then $F_t(x,v) = 0$ for all $0 < t \le T$. 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 $(x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T$ such that for some $1 \le i \le d$ and $t_0 \in]0,T]$, $\partial_i F_{t_0}(x_{t_0}, v_{t_0-}) > 0$. Let $\alpha = \frac{1}{2}\partial_i F_{t_0}(x_{t_0}, v_{t_0-})$. By the left-continuity of $\partial_i F$ and $\mathcal{D}_t F$ at (x_{t_0}, v_{t_0}), we may choose $l < T - t$ sufficiently small such that, for any $t' \in [0, t_0]$, for any $(x', v') \in \mathcal{U}_{t'} \times \mathcal{S}_{t'}$,

$$d_\infty((x_{t_0}, v_{t_0}), (x', v')) < l \Rightarrow \partial_i F_{t'}(x', v') > \alpha \text{ and } |\mathcal{D}_t F(x', v')| < 1 \tag{2.27}$$

Choose $t < t_0$ such that $d_\infty((x_t, v_t), (x_{t_0}, v_{t_0-})) < \frac{l}{2}$ and define the following extension of x_t to $[0, t+h]$, where $h < \frac{l}{4} \wedge (t_0 - t)$:

$$\begin{aligned} z(u) &= x(u), u \le t \\ z_j(u) &= x_j(t) + 1_{i=j}(u-t), t \le u \le t+h, 1 \le j \le d \end{aligned} \tag{2.28}$$

Define the following sequence of piecewise constant approximations of z:

$$\begin{aligned} z^n(u) &= z(u), u \le t \\ z_j^n(u) &= x_j(t) + 1_{i=j}\frac{h}{n}\sum_{k=1}^{n} 1_{\frac{kh}{n} \le u-t}, t \le u \le t+h, 1 \le j \le d \end{aligned} \tag{2.29}$$

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

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

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

$$F_{t+h}(z^n, v_{t,h}) - F_t(x, v) = \sum_{k=1}^{n} F_{t+\frac{kh}{n}}(z^n_{t+\frac{kh}{n}}, v_{t,\frac{kh}{n}}) - F_{t+\frac{kh}{n}}(z^n_{t+\frac{kh}{n}-}, v_{t,\frac{kh}{n}})$$
$$+ \sum_{k=1}^{n} F_{t+\frac{kh}{n}}(z^n_{t+\frac{kh}{n}-}, v_{t,\frac{kh}{n}}) - F_{t+\frac{(k-1)h}{n}}(z^n_{t+\frac{(k-1)h}{n}}, v_{t,\frac{(k-1)h}{n}}) \tag{2.30}$$

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 $[t + \frac{(k-1)h}{n}, t + \frac{kh}{n}[$.

$$F_{t+\frac{kh}{n}}(z^n_{t+\frac{kh}{n}}, v_{t,\frac{kh}{n}}) - F_{t+\frac{kh}{n}}(z^n_{t+\frac{kh}{n}-}, v_{t,\frac{kh}{n}}) = \phi(\frac{h}{n}) - \phi(0) \tag{2.31}$$

where ϕ is defined as

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

Since F is vertically differentiable, ϕ is differentiable and

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

Since

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

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

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

On the other hand

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

where

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

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

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

Since $F \in \mathbb{F}_l^\infty$, ψ is also left-continous continuous by theorem 4 so

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

Noting that:

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

we obtain that:

$$\mathcal{D}_{t+u}F(z^n_{t+u}, v_{t,u}) \underset{n\to+\infty}{\to} \mathcal{D}_{t+u}F(z_{t+u}, v_{t,u}) \; = 0$$

since the path of z_{t+u} is continuous. Moreover
$|\mathcal{D}_t F_{t+u}(z^n_{t+u}, v_{t,u})| \leq 1$ since $d_\infty((z^n_{t+u}, v_{t,u}), (x_t, v_t) \leq h$, so by dominated convergence the integral goes to 0 as $n \to \infty$. 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 $n \to \infty$ leads to $F_{t+h}(z, v_{t,h}) - F_t(x, v) \geq \alpha h$, a contradiction. □

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 $F^i \in \mathbb{C}^{1,2}$, 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 [36, 37].

Theorem 2.3. *If $F^1, F^2 \in \mathbb{C}^{1,2}$ with $F^i, \nabla_x F^i, \nabla_x^2 F^i \in \mathbb{F}_l^\infty$ and $\mathcal{D}F^i$ satisfying the local boundedness assumption 2.9 for $i = 1, 2$, coincide on continuous paths::*

$$\forall (x, v) \in \mathcal{U}_T^c \times \mathcal{S}_T, \qquad \forall t \in]0, T], \quad F_t^1(x_t, v_t) = F_t^2(x, v) \tag{2.32}$$

then their second vertical derivatives also coincide on continuous paths:

$$\forall (x, v) \in \mathcal{U}_T^c \times \mathcal{S}_T, \quad \forall t \in]0, T], \quad \nabla_x^2 F_t^1(x_t, v_{t-}) = \nabla_x^2 F_t^2(x_t, v_{t-})$$

Proof. Let $F = F^1 - F^2$. Assume now that there exists some $(x, v) \in \mathcal{U}_T^c \times \mathcal{S}_T$ such that for some $1 \leq i \leq d$ and $t_0 \in]0, T]$, and some direction $h \in \mathbb{R}^d$, $\|h\| = 1$, ${}^t h \nabla_x^2 F_{t_0}(x_{t_0}, v_{t_0-}).h > 0$, and denote $\alpha = \frac{1}{2} {}^t h \nabla_x^2 F_{t_0}(x_{t_0}, v_{t_0-}).h$. We will show that this leads to a contradiction. We already know that $\nabla_x F_t(x_t, v_t) = 0$ by theorem 2.2. Let $\eta > 0$ be small enough so that:

$$\forall t' \leq t_0, \forall (x', v') \in \mathcal{U}_{t'} \times \mathcal{S}_{t'},$$
$$d_\infty((x_t, v_t), (x', v')) < \eta \Rightarrow |F_{t'}(x', v')| < |F_{t_0}(x_{t_0}, v_{t_0-})| + 1, |\nabla_x F_{t'}(x', v')| < 1,$$
$$|\mathcal{D}_{t'} F(x', v')| < 1, {}^t h \nabla_x^2 F_{t'}(x', v').h > \alpha \quad (2.33)$$

Choose $t < t_0$ such that $d_\infty((x_t, v_t), (x_{t_0}, v_{t_0-})) < \frac{\eta}{2}$ and denote $\epsilon = \frac{\eta}{2} \wedge (t_0 - t)$. 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

$$\tau = \inf\{s > 0, \qquad |W(s)| = \frac{\epsilon}{2}\} \quad (2.34)$$

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

$$U(t') = x(t') 1_{t' \leq t} + (x(t) + W((t' - t) \wedge \tau) h) 1_{t' > t} \quad (2.35)$$

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

$$d_\infty((U_{t+s}, v_{t,s}), (x_t, v_t)) < \epsilon \quad (2.36)$$

Define the following piecewise constant approximations of the stopped process W^τ:

$$W^n(s) = \sum_{i=0}^{n-1} W(i\frac{\epsilon}{2n} \wedge \tau) 1_{s \in [i\frac{\epsilon}{2n}, (i+1)\frac{\epsilon}{2n}[} + W(\frac{\epsilon}{2} \wedge \tau) 1_{s = \frac{\epsilon}{2}}, 0 \leq s \leq \frac{\epsilon}{2n} \quad (2.37)$$

Denoting

$$Z(s) = F_{t+s}(U_{t+s}, v_{t,s}), \quad s \in [0, T - t] \quad (2.38)$$
$$U^n(t') = x(t') 1_{t' \leq t} + (x(t) + W^n((t' - t) \wedge \tau) h) 1_{t' > t} \qquad Z^n(s) = F_{t+s}(U^n_{t+s}, v_{t,s}) \quad (2.39)$$

we have the following decomposition:

$$\begin{aligned} Z(\frac{\epsilon}{2}) - Z(0) &= Z(\frac{\epsilon}{2}) - Z^n(\frac{\epsilon}{2}) + \sum_{i=1}^{n} Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-) \\ &+ \sum_{i=0}^{n-1} Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) \end{aligned} \quad (2.40)$$

The first term in the right-hand side of (2.40) goes to 0 almost surely since

$$d_\infty((U_{t+\frac{\epsilon}{2}}, v_{t,\frac{\epsilon}{2}}), (U^n_{t+\frac{\epsilon}{2}}, v_{t,\frac{\epsilon}{2}})) \overset{n\to\infty}{\to} 0. \tag{2.41}$$

The second term in (2.40) may be expressed as

$$Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-) = \phi_i(W(i\frac{\epsilon}{2n}) - W((i-1)\frac{\epsilon}{2n})) - \phi_i(0) \tag{2.42}$$

where:

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

Note that $\phi_i(u,\omega)$ is measurable with respect to $\mathcal{B}_{(i-1)\epsilon/2n}$ whereas its argument in (2.42) is independent with respect to $\mathcal{B}_{(i-1)\epsilon/2n}$. Let $\Omega_1 = \{\omega \in \tilde{\Omega}, t \mapsto W(t,\omega) \quad \text{continuous}\}$. Then $\mathbb{P}(\Omega_1) = 1$ and for any $\omega \in \Omega_1$, $\phi_i(.,\omega)$ is $\mathcal{C}^2$ with:

$$\begin{aligned} \phi_i'(u,\omega) &= \nabla_x F_{t+i\frac{\epsilon}{2n}}(U^{n,uh}_{t+i\frac{\epsilon}{2n}-}(\omega), v_{t,i\frac{\epsilon}{2n}})h \\ \phi_i''(u,\omega) &= {}^t h \nabla_x^2 F_{t+i\frac{\epsilon}{2n}}(U^{n,uh}_{t+i\frac{\epsilon}{2n}-}(\omega), v_{t,i\frac{\epsilon}{2n}}).h \end{aligned} \tag{2.43}$$

So, using the above arguments we can apply the Itô formula to (2.42). We therefore obtain, summing on i and denoting $i(s)$ the index such that $s \in [(i-1)\frac{\epsilon}{2n}, i\frac{\epsilon}{2n})$:

$$\begin{aligned} \sum_{i=1}^n Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-) = &\int_0^{\frac{\epsilon}{2}} \nabla_x F_{t+i(s)\frac{\epsilon}{2n}}(U^{n,uh}_{t+i(s)\frac{\epsilon}{2n}-}, v_{t,i(s)\frac{\epsilon}{2n}})h dW(s) \\ &+ \int_0^{\frac{\epsilon}{2}} {}^t h.\nabla_x^2 F_{t+i(s)\frac{\epsilon}{2n}}(U^{n,uh}_{t+i(s)\frac{\epsilon}{2n}-}, v_{t,i(s)\frac{\epsilon}{2n}}).h ds \end{aligned} \tag{2.44}$$

Since the first derivative is bounded by (2.33), the stochastic integral is a martingale, so taking expectation leads to:

$$E[\sum_{i=1}^n Z^n(i\frac{\epsilon}{2n}) - Z^n(i\frac{\epsilon}{2n}-)] > \alpha\frac{\epsilon}{2} \tag{2.45}$$

Now

$$Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) = \psi(\frac{\epsilon}{2n}) - \psi(0) \tag{2.46}$$

where

$$\psi(u) = F_{t+(i-1)\frac{\epsilon}{2n}+u}(U^n_{t+(i-1)\frac{\epsilon}{2n},u}, v_{t,(i-1)\frac{\epsilon}{2n}+u}) \tag{2.47}$$

is right-differentiable with right derivative:

$$\psi'(u) = \mathcal{D}_t F_{t+(i-1)\frac{\epsilon}{2n}+u}(U^n_{(i-1)\frac{\epsilon}{2n},u}, v_{t,(i-1)\frac{\epsilon}{2n}+u}) \tag{2.48}$$

Since $F \in \mathbb{F}^\infty_l([0,T])$, ψ is left-continuous by theorem 4 and the fundamental theorem of calculus yields:

$$\sum_{i=0}^{n-1} Z^n((i+1)\frac{\epsilon}{2n}-) - Z^n(i\frac{\epsilon}{2n}) = \int_0^{\frac{\epsilon}{2}} \mathcal{D}_{t+s}F(U^n_{t+(i(s)-1)\frac{\epsilon}{2n}+u}, v_{t,s})ds \tag{2.49}$$

The integrand converges to $\mathcal{D}_t F_{t+s}(\mathcal{U}_{t+(i(s)-1)\frac{\epsilon}{2n}+u}, v_{t,s}) = 0$ 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 (2.40) are bounded, the dominated convergence theorem allows to take expectations of both sides in (2.40) with respect to the Wiener measure and obtain $\alpha\frac{\epsilon}{2} = 0$, a contradiction. □

Remark 2.5. If a functional is predictable in the second variable, so are its vertical derivatives hence we can state in the setting of theorems 2.2, 2.3 that $\nabla_x F^1_t(x_t, v_t) = \nabla_x F^2_t(x_t, v_t)$, $\nabla^2_x F^1_t(x_t, v_t) = \nabla^2_x F^2_t(x_t, v_t)$.

Remark 2.6. Both results extend (replacing $\forall t \in]0,T]$ by $\forall t \in [0,T[$) if the vertical derivatives (but not the functional itself) are in $\mathbb{F}^\infty_r$ instead of $\mathbb{F}^\infty_l$, following the same proof but extending directly the path of (x,v) from t_0 rather than stepping back in time first.

2.4 Change of variable formula for functionals of a continuous path

We now state our first main result, a functional change of variable formula which extends the Itô formula without probability due to Föllmer [29] to functionals. We denote here S^+_d the set of positive symmetric $d \times d$ matrices.

Definition 2.10. Let $\pi_n = (t^n_0, \ldots, t^n_{k(n)})$, where $0 = t^n_0 \leq t^n_1 \leq \ldots \leq t^n_{k(n)} = T$, be a sequence of subdivisions of $[0,T]$ with step decreasing to 0 as $n \to \infty$. $f \in C_0([0,T],\mathbb{R})$ is said to have finite quadratic variation along (π_n) if the sequence of discrete measures:

$$\xi^n = \sum_{i=0}^{k(n)-1} (f(t_{i+1}^n) - f(t_i^n))^2 \delta_{t_i^n} \tag{2.50}$$

where δ_t is the Dirac measure at t, converge vaguely to a Radon measure ξ on $[0,T]$ whose atomic part is null. The increasing function $[f]$ defined by

$$[f](t) = \xi([0,t])$$

is then called the quadratic variation of f along the sequence (π_n).
$x \in C_0([0,T],U)$ is said to have finite quadratic variation along the sequence (π_n) if the functions $x_i, 1 \le i \le d$ and $x_i + x_j, 1 \le i < j \le d$ do. The quadratic variation of x along (π_n) is the S_d^+-valued function x defined by:

$$[x]_{ii} = [x_i], [x]_{ij} = \frac{1}{2}([x_i + x_j] - [x_i] - [x_j]), i \neq j \tag{2.51}$$

Theorem 2.4 (Change of variable formula for functionals of continuous paths)**.** *Let* $(x,v) \in C_0([0,T],U) \times \mathcal{S}_T$ *such that* x *has finite quadratic variation along* (π_n) *and verifies* $\sup_{t\in[0,T]-\pi_n} |v(t) - v(t-)| \to 0$*. Denote:*

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

$$v^n(t) = \sum_{i=0}^{k(n)-1} v(t_i) 1_{[t_i,t_{i+1}[}(t) + v(T) 1_{\{T\}}(t), \qquad h_i^n = t_{i+1}^n - t_i^n \tag{2.52}$$

Then for any non-anticipative functional $F \in \mathbb{C}^{1,2}$ *such that:*

1. $F, \nabla_x F, \nabla_x^2 F \in \mathbb{F}_l^\infty$
2. $\nabla_x^2 F, \mathcal{D}F$ *satisfy the local boundedness property* (2.9)

the following limit

$$\lim_{n\to\infty} \sum_{i=0}^{k(n)-1} \nabla_x F_{t_i^n}(x_{t_i^n-}^n, v_{t_i^n-}^n)(x(t_{i+1}^n) - x(t_i^n)) \tag{2.53}$$

exists. Denoting this limit by $\int_0^T \nabla_x F(x_u, v_u) d^\pi x$ *we have*

$$F_T(x_T, v_T) - F_0(x_0, v_0) = \int_0^T \mathcal{D}_t F_t(x_u, v_u) du \tag{2.54}$$

$$+ \int_0^T \frac{1}{2} \mathrm{tr}\left({}^t\nabla_x^2 F_t(x_u, v_u) d[x](u)\right) + \int_0^T \nabla_x F(x_u, v_u) d^\pi x \tag{2.55}$$

Remark 2.7 (Föllmer integral). The limit (2.53), which we call the *Föllmer integral*, was defined in [29] for integrands of the form $f(X(t))$ where $f \in C^1(\mathbb{R}^d)$. It depends a priori on the sequence π of subdivisions, hence the notation $\int_0^T \nabla_x F(x_u, v_u) d^\pi x$. We will see in Section 2.7 that when x is the sample path of a *semimartingale*, the limit is in fact almost-surely independent of the choice of π.

Remark 2.8. The regularity conditions on F are given independently of (x, v) and of the sequence of subdivisions (π_n).

Proof. Denote $\delta x_i^n = x(t_{i+1}^n) - x(t_i^n)$. Since x is continuous hence uniformly continuous on $[0, T]$, and using Lemma A.1 for v, the quantity

$$\eta_n = \sup\{|v(u) - v(t_i^n)| + |x(u) - x(t_i^n)| + |t_{i+1}^n - t_i^n|, 0 \le i \le k(n) - 1, u \in [t_i^n, t_{i+1}^n)\} \quad (2.56)$$

converges to 0 as $n \to \infty$. Since $\nabla_x^2 F, \mathcal{D}F$ satisfy the local boundedness property (2.9), for n sufficiently large there exists $C > 0$ such that

$$\forall t < T, \forall (x', v') \in \mathcal{U}_t \times \mathcal{S}_t, d_\infty((x_t, v_t), (x', v')) < \eta_n \Rightarrow |\mathcal{D}_t F_t(x', v')| \le C, |\nabla_x^2 F_t(x', v')| \le C$$

Denoting $K = \overline{\{x(u), s \le u \le t\}}$ which is a compact subset of U, and $U^c = \mathbb{R} - U$ its complement, one can also assume n sufficiently large so that $d(K, U^c) > \eta_n$.

For $i \le k(n) - 1$, consider the decomposition:

$$\begin{aligned} F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n) - F_{t_i^n}(x_{t_i^n-}^n, v_{t_i^n-}^n) &= F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_i^n, h_i^n}^n) - F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) \\ &+ F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n-}^n) - F_{t_i^n}(x_{t_i^n-}^n, v_{t_i^n-}^n) \end{aligned} \quad (2.57)$$

where we have used property (2.2) to have $F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) = F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n-}^n)$. The first term can be written $\psi(h_i^n) - \psi(0)$ where:

$$\psi(u) = F_{t_i^n+u}(x_{t_i^n,u}^n, v_{t_i^n,u}^n) \quad (2.58)$$

Since $F \in \mathbb{C}^{1,2}([0, T])$, ψ is right-differentiable, and moreover by lemma 4, ψ is left-continuous, so:

$$F_{t_{i+1}^n}(x_{t_i^n, h_i^n}^n, v_{t_i^n, h_i^n}^n) - F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) = \int_0^{t_{i+1}^n - t_i^n} \mathcal{D}_{t_i^n+u} F(x_{t_i^n,u}^n, v_{t_i^n,u}^n) du \quad (2.59)$$

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

$$\phi(u) = F_{t_i^n}(x^{n,u}_{t_i^n-}, v^n_{t_i^n-}) \tag{2.60}$$

Since $F \in \mathbb{C}^{1,2}([0,T])$, ϕ is well-defined and C^2 on the convex set $B(x(t_i^n), \eta_n) \subset U$, with:

$$\begin{aligned}\phi'(u) &= \nabla_x F_{t_i^n}(x^{n,u}_{t_i^n-}, v^n_{t_i^n-})\\ \phi''(u) &= \nabla_x^2 F_{t_i^n}(x^{n,u}_{t_i^n-}, v^n_{t_i^n-})\end{aligned} \tag{2.61}$$

So a second order Taylor expansion of ϕ at $u = 0$ yields:

$$\begin{aligned}F_{t_i^n}(x^n_{t_i^n}, v^n_{t_i^n-}) - F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-}) &= \nabla_x F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-})\delta x_i^n\\ &+\frac{1}{2}\mathrm{tr}\left(\nabla_x^2 F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-}) \quad {}^t\delta x_i^n \delta x_i^n\right) + r_i^n\end{aligned} \tag{2.62}$$

where r_i^n is bounded by

$$K|\delta x_i^n|^2 \sup_{x\in B(x(t_i^n),\eta_n)} |\nabla_x^2 F_{t_i^n}(x^{n,x-x(t_i^n)}_{t_i^n-}, v^n_{t_i^n-}) - \nabla_x^2 F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-})| \tag{2.63}$$

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

- The left-hand side of (2.57) yields $F_T(x^n_{T-}, v^n_{T-}) - F_0(x_0, v_0)$, which converges to $F_T(x_{T-}, v_{T-}) - F_0(x_0, v_0)$ by left-continuity of F, and this quantity equals $F_T(x_T, v_T) - F_0(x_0, v_0)$ since x is continuous and F is predictable in the second variable.

- The first line in the right-hand side can be written:

$$\int_0^T \mathcal{D}_u F(x^n_{t^n_{i^n(u)}, u-t^n_{i^n(u)}}, v^n_{t^n_{i^n(u)}, u-t^n_{i^n(u)}})du \tag{2.64}$$

  where the integrand converges to $\mathcal{D}_u F(x_u, v_{u-})$ and is bounded by C. Hence the dominated convergence theorem applies and (2.64) converges to:

$$\int_0^T \mathcal{D}_u F(x_u, v_{u-})du = \int_0^T \mathcal{D}_u F(x_u, v_u) \tag{2.65}$$

  since $v_u = v_{u-}, du$-almost everywhere.

- The second line can be written:

$$\sum_{i=0}^{k(n)-1} \nabla_x F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-})(x_{t_{i+1}^n} - x_{t_i^n}) + \sum_{i=0}^{k(n)-1} \frac{1}{2} tr[\nabla_x^2 F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-})]^t \delta x_i^n \delta x_i^n] \quad (2.66)$$

$$+ \sum_{i=0}^{k(n)-1} r_i^n \quad (2.67)$$

$[\nabla_x^2 F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-})]1_{t\in]t_i^n, t_{i+1}^n]}$ is bounded by C, and converges to $\nabla_x^2 F_t(x_t, v_{t-})$ by left-continuity of $\nabla_x^2 F$, and the paths of both are left-continuous by lemma 4. Since x and the subdivision (π_n) are as in definition 2.10, lemma A.5 in appendix A.3 applies and gives as limit:

$$\int_0^T \frac{1}{2} \mathrm{tr}[{}^t\nabla_x^2 F_t(x_u, v_{u-})] d[x](u)] = \int_0^T \frac{1}{2} \mathrm{tr}[{}^t\nabla_x^2 F_t(x_u, v_u)] d[x](u)] \quad (2.68)$$

since $\nabla_x^2 F$ is predictable in the second variable i.e. verifies (2.2). Using the same lemma, since $|r_i^n|$ is bounded by $\epsilon_i^n |\delta x_i^n|^2$ where ϵ_i^n converges to 0 and is bounded by $2C$, $\sum_{i=i^n(s)+1}^{i^n(t)-1} r_i^n$ converges to 0.

Since all other terms converge, the limit:

$$\lim_n \sum_{i=0}^{k(n)-1} \nabla_x F_{t_i^n}(x^n_{t_i^n-}, v^n_{t_i^n-})(x(t_{i+1}^n) - x(t_i^n)) \quad (2.69)$$

exists, and the result is established.

□

2.5 Change of variable formula for functionals of a cadlag path

We will now extend the previous result to functionals of cadlag paths. The following definition is taken from Föllmer [29]:

Definition 2.11. Let $\pi_n = (t_0^n, \ldots, t_{k(n)}^n)$, where $0 = t_0^n \le t_1^n \le \ldots \le t_{k(n)}^n = T$ be a sequence of subdivisions of $[0, T]$ with step decreasing to 0 as $n \to \infty$. $f \in D([0, T], \mathbb{R})$ is said to have finite quadratic variation along (π_n) if the sequence of discrete measures:

$$\xi^n = \sum_{i=0}^{k(n)-1} (f(t_{i+1}^n) - f(t_i^n))^2 \delta_{t_i^n} \quad (2.70)$$

where δ_t is the Dirac measure at t, converge vaguely to a Radon measure ξ on $[0,T]$ such that

$$[f](t) = \xi([0,t]) = [f]^c(t) + \sum_{0<s\le t} (\Delta f(s))^2 \tag{2.71}$$

where $[f]^c$ is the continuous part of $[f]$. $[f]$ is called quadratic variation of f along the sequence (π_n). $x \in \mathcal{U}_T$ is said to have finite quadratic variation along the sequence (π_n) if the functions $x_i, 1 \le i \le d$ and $x_i + x_j, 1 \le i < j \le d$ do. The quadratic variation of x along (π_n) is the S_d^+-valued function x defined by:

$$[x]_{ii} = [x_i], \qquad [x]_{ij} = \frac{1}{2}([x_i + x_j] - [x_i] - [x_j]), \quad i \neq j \tag{2.72}$$

Theorem 2.5 (Change of variable formula for functionals of discontinuous paths)**.** *Let $(x,v) \in \mathcal{U}_T \times \mathcal{S}_T$ where x has finite quadratic variation along (π_n) and*

$$\sup_{t\in[0,T]-\pi_n} |x(t) - x(t-)| + |v(t) - v(t-)| \to 0 \tag{2.73}$$

Denote

$$x^n(t) = \sum_{i=0}^{k(n)-1} x(t_{i+1}-)1_{[t_i,t_{i+1})}(t) + x(T)1_{\{T\}}(t)$$
$$v^n(t) = \sum_{i=0}^{k(n)-1} v(t_i)1_{[t_i,t_{i+1})}(t) + v(T)1_{\{T\}}(t), \qquad h_i^n = t_{i+1}^n - t_i^n \tag{2.74}$$

Then for any non-anticipative functional $F \in \mathbb{C}^{1,2}$ such that:

1. *F is predictable in the second variable in the sense of* (2.2)
2. *$\nabla_x^2 F$ and $\mathcal{D}F$ have the local boundedness property* (2.9)
3. *$F, \nabla_x F, \nabla_x^2 F \in \mathbb{F}_l^\infty$*
4. *$\nabla_x F$ has the horizontal local Lipschitz property* (2.24)

the Föllmer integral, defined as the limit

$$\int_{]0,T]} \nabla_x F_t(x_{t-}, v_{t-}) d^\pi x \;:=\; \lim_{n\to\infty} \sum_{i=0}^{k(n)-1} \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)(x(t_{i+1}^n) - x(t_i^n)) \tag{2.75}$$

exists and

$$F_T(x_T, v_T) - F_0(x_0, v_0) = \int_{]0,T]} \mathcal{D}_t F_t(x_{u-}, v_{u-}) du$$
$$+ \int_{]0,T]} \frac{1}{2} \mathrm{tr}\left({}^t\nabla_x^2 F_t(x_{u-}, v_{u-}) d[x]^c(u)\right) + \int_{]0,T]} \nabla_x F_t(x_{t-}, v_{t-}) d^\pi x$$
$$+ \sum_{u \in]0,T]} [F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) - \nabla_x F_u(x_{u-}, v_{u-}).\Delta x(u)] \tag{2.76}$$

Remark 2.9. Condition (2.73) simply means that the subdivision asymptotically contains all discontinuity points of (x, v). Since a cadlag function has at most a countable set of discontinuities, this can always be achieved by adding e.g. the discontinuity points $\{t \in [0, T], \max(|\Delta x(t)|, |\Delta v(t)|) \geq 1/n\}$ to π_n.

Proof. Denote $\delta x_i^n = x(t_{i+1}^n) - x(t_i^n)$. Lemma A.1 implies that

$$\eta_n = \sup\{|v(u) - v(t_i^n)| + |x(u) - x(t_i^n)| + |t_{i+1}^n - t_i^n|, 0 \leq i \leq k(n) - 1, u \in [t_i^n, t_{i+1}^n)\} \overset{n \to \infty}{\to} 0$$

so for n sufficiently large there exists $C > 0$ such that, for any $t < T$, for any $(x', v') \in \mathcal{U}_t \times \mathcal{S}_t$, $d_\infty((x_t, v_t), (x', v')) < \eta_n \Rightarrow |\mathcal{D}_t F_t(x', v')| \leq C, |\nabla_x^2 F_t(x', v')| \leq C$, using the local boundedness property (2.9).

For $\epsilon > 0$, we separate the jump times of x in two sets: a finite set $C_1(\epsilon)$ and a set $C_2(\epsilon)$ such that $\sum_{s \in C_2(\epsilon)} |\Delta x_s|^2 < \epsilon^2$. We also separate the indices $0 \leq i \leq k(n) - 1$ in two sets: a set $I_1^n(\epsilon)$ such that $(t_i, t_{i+1}]$ contains at least a time in $C_1(\epsilon)$, and its complementary $I_2^n(\epsilon)$. Denoting $K = \overline{\{x(u), s \leq u \leq t\}}$ which is a compact subset of U, and $U^c = \mathbb{R} - U$, one may choose ϵ sufficiently small and n sufficiently large so that $d(K, U^c) > \epsilon + \eta_n$.
Denote $i^n(t)$ the index such that $t \in [t_i^n, t_{i+1}^n)$. Property (2.73) implies that for n sufficiently large, $C_1(\epsilon) \subset \{t_{i+1}^n, i = 1..k(n)\}$ so

$$\sum_{0 \leq i \leq k(n)-1, i \in I_1^n(\epsilon)} F_{t_{i+1}^n}(x_{t_{i+1}^n -}^{n, \Delta x(t_{i+1}^n)}, v_{t_{i+1}^n -}^n) - F_{t_i^n}(x_{t_i^n -}^{n, \Delta x(t_n^i)}, v_{t_i^n -}^n) \tag{2.77}$$
$$\longrightarrow_{n \to \infty} \sum_{u \in]0,T] \cup C_1(\epsilon)} F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) \tag{2.78}$$

as $n \to \infty$, by left-continuity of F.

Let us now consider, for $i \in I_2^n(\epsilon), i \leq k(n)-1$, the decomposition:

$$\begin{aligned}F_{t_{i+1}^n}(x_{t_{i+1}^n-}^{n,\Delta x(t_{i+1}^n)}, v_{t_{i+1}^n-}^n) - F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n) = \\ F_{t_{i+1}^n}(x_{t_{i+1}^n-}^{n,\Delta x(t_{i+1}^n)}, v_{t_{i+1}^n-}^n) - F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n) \\ +F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_i^n,h_i^n}^n) - F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) \\ +F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n-}^n) - F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)\end{aligned} \quad (2.79)$$

where we have used the property (2.2) to obtain $F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) = F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n-}^n)$. The third line in (2.79) can be written $\psi(h_i^n) - \psi(0)$ where:

$$\psi(u) = F_{t_i^n+u}(x_{t_i^n,u}^n, v_{t_i^n,u}^n) \quad (2.80)$$

Since $F \in \mathbb{C}^{1,2}([0,T])$, ψ is right-differentiable, and moreover by lemma 4, ψ is continuous, so:

$$F_{t_{i+1}^n}(x_{t_i^n,h_i^n}^n, v_{t_i^n,h_i^n}^n) - F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) = \int_0^{t_{i+1}^n - t_i^n} \mathcal{D}_{t_i^n+u}F(x_{t_i^n,u}^n, v_{t_i^n,u}^n)du \quad (2.81)$$

The fourth line in (2.79) can be written $\phi(x(t_{i+1}^n-) - x(t_i^n)) - \phi(0)$, where:

$$\phi(u) = F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)+u}, v_{t_i^n-}^n) \quad (2.82)$$

Since $F \in \mathbb{C}^{1,2}([0,T])$, ϕ is well-defined and C^2 on the convex set $B(x(t_i^n), \eta_n + \epsilon) \subset U$, with:

$$\phi'(u) = \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)+u}, v_{t_i^n-}^n)\phi''(u) = \nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)+u}, v_{t_i^n-}^n) \quad (2.83)$$

So a second order Taylor expansion of ϕ at $u=0$ yields:

$$\begin{aligned}F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n-}^n) - F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n) = \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)(x(t_{i+1}^n-) - x(t_i^n)) \\ +\frac{1}{2}tr[\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)]^t(x(t_{i+1}^n-) - x(t_i^n))(x(t_{i+1}^n-) - x(t_i^n))] + r_i^n\end{aligned} \quad (2.84)$$

where $r_{i,1}^n$ is bounded by

$$K|(x(t_{i+1}^n-) - x(t_i^n))|^2 \sup_{x \in B(x(t_i^n), \eta_n+\epsilon)} |\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,x-x(t_i^n)}, v_{t_i^n-}^n) - \nabla_x^2 F_{t_i^n}(x_{t_i^n-}^n, v_{t_i^n-}^n)| \quad (2.85)$$

Similarly, the second line in (2.79) can be written $\phi(\Delta x(t_{i+1}^n)) - \phi(0)$ where

$$\phi(u) = F_{t_{i+1}^n}(x_{t_{i+1}^n-}^{n,u}, v_{t_{i+1}^n-}^n)$$

. So, a second order Taylor expansion of ϕ at $u = 0$ yields:

$$\begin{aligned} F_{t_{i+1}^n}(x_{t_{i+1}^n-}^{n,\Delta x(t_{i+1}^n)}, v_{t_{i+1}^n-}^n) - F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n) \\ = \nabla_x F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n)\Delta x(t_{i+1}^n) \\ +\frac{1}{2}tr[\nabla_x^2 F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n)]^t \Delta x(t_{i+1}^n)\Delta x(t_{i+1}^n) + r_{i,2}^n \end{aligned} \tag{2.86}$$

Using the horizontal local Lipschitz property (2.24) for $\nabla_x F$, for n sufficiently large:

$$|\nabla_x F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n) - \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)| < C[(t_{i+1}^n - t_i^n) + |x(t_{i+1}^n-) - x(t_i^n)|]$$

On other hand, since $\nabla_x^2 F$ is bounded by C on all paths considered:

$$\begin{aligned} | \quad \mathrm{tr}\left(\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)^t (x(t_{i+1}^n-) - x(t_i^n))(x(t_{i+1}^n-) - x(t_i^n))\right) \\ +\mathrm{tr}\left(\nabla_x^2 F_{t_{i+1}^n}(x_{t_i^n-,h_i^n}^{n,\Delta x(t_i^n)}, v_{t_i^n,h_i^n}^n)]^t \Delta x(t_{i+1}^n)\Delta x(t_{i+1}^n)\right) \\ -\mathrm{tr}\left(\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)]^t \delta x_i^n \delta x_i^n\right) | < 2C|\Delta x(t_{i+1}^n)|^2 \end{aligned} \tag{2.87}$$

Hence, we have shown that:

$$\begin{aligned} F_{t_{i+1}^n}(x_{t_{i+1}^n-}^{n,\Delta x(t_{i+1}^n)}, v_{t_{i+1}^n-}^n) - F_{t_{i+1}^n}(x_{t_{i+1}^n-}^n, v_{t_{i+1}^n-}^n) + F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n-}^n) - F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n) = \\ \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)\delta x_i^n + \frac{1}{2}tr[\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)]^t \delta x_i^n \delta x_i^n] + r_i^n + q_i^n \end{aligned}$$

where r_i^n is bounded by:

$$4K|\delta x_i^n|^2 \sup_{x \in B(x(t_i^n), \eta_n + \epsilon)} |\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,x-x(t_i^n-)}, v_{t_i^n-}^n) - \nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)| \tag{2.88}$$

and q_i^n is bounded by:

$$C'(h_i^n|\Delta x(t_i^n)| + |\Delta x(t_i^n)|^2) \tag{2.89}$$

Denote $i^n(t)$ the index such that $t \in [t_{i^n(t)}^n, t_{i^n(t)+1}^n[$. Summing all the terms above for $i \in C_2(\epsilon) \cap \{0, 1, ..k(n) - 1\}$:

- The left-hand side of (2.79) yields

$$F_T(x_T^n, v_T^n) - F_0(x_0, v_0) - \sum_{0 \le i \le k(n)-1, i \in I_1^n(\epsilon)} F_{t_{i+1}^n}(x_{t_{i+1}^n}^n, v_{t_{i+1}^n}^n) - F_{t_i^n}(x_{t_i^n}^n, v_{t_i^n}^n) \tag{2.90}$$

which converges to

$$F_T(x_T, v_T) - F_0(x_0, v_0) - \sum_{u \in]0,T] \cup C_1(\epsilon)} F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) \tag{2.91}$$

- The sum of the second and fourth lines of (2.79) can be written:

$$\sum_{0\le i\le k(n)-1, i\in I_2^n(\epsilon)} \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)\delta x_i^n$$
$$+ \sum_{0\le i\le k(n)-1, i\in I_2^n(\epsilon)} \frac{1}{2}\mathrm{tr}\left(\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)^t \delta x_i^n \delta x_i^n\right)$$
$$+ \sum_{0\le i\le k(n)-1, i\in I_2^n(\epsilon)} r_i^n + q_i^n \tag{2.92}$$

Consider the measures $\mu_{ij}^n = \xi_{ij}^n - \sum_{0<s\le T, s\in C_2(\epsilon)} (\Delta f_{ij}(s))^2 \delta_s$, where $f_{ii} = x_i, 1 \le j \le d$ and $f_{ij} = x_i + x_j, 1 \le i < j \le d$ and ξ_{ij}^n is defined in Definition 2.11. The second line of (2.92) can be decomposed as:

$$A_n + \frac{1}{2} \sum_{0<u\le T, u\in C_2(\epsilon)} \mathrm{tr}\left(\nabla_x^2 F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)^t \Delta x(u)\Delta x(u)\right) \tag{2.93}$$

where

$$A_n = \mathrm{tr} \int_{]0,T]} \mu^n(dt) \sum_{0\le i\le k(n)-1, i\in I_2^n(\epsilon)} \nabla_x^2 F_{t_{i^n(t)}^n}(x_{t_{i^n(t)}^n-}^{n,\Delta x(t_{i^n(t)}^n)}, v_{t_{i^n(t)}^n-}^n) \quad 1_{t\in(t_i^n, t_{i+1}^n]}$$

where μ^n denotes the matrix-valued measure with components μ_{ij}^n defined above. μ_{ij}^n converges vaguely to the atomless measure $[f_{ij}]^c$. Since

$$\sum_{0\le i\le k(n)-1, i\in I_2^n(\epsilon)} \nabla_x^2 F_{t_{i^n(t)}^n}(x_{t_{i^n(t)}^n-}^{n,\Delta x(t_{i^n(t)}^n)}, v_{t_{i^n(t)}^n-}^n) \ 1_{t\in(t_i^n, t_{i+1}^n]}$$

is bounded by C and converges to $\nabla_x^2 F_t(x_{t-}, v_{t-}) 1_{t\notin C_1(\epsilon)}$ by left-continuity of $\nabla_x^2 F$, applying Lemma A.5 to A_n and yields that A_n converges to:

$$\int_{]0,T]} \frac{1}{2}\mathrm{tr}\left({}^t\nabla_x^2 F_t(x_{u-}, v_{u-}) d[x]^c(u)\right) \tag{2.94}$$

The second term in (2.93) has the lim sup of its absolute value bounded by $C\epsilon^2$. Using the same argument, since $|r_i^n|$ is bounded by $s_i^n |\delta x_i^n|^2$ for some s_i^n which converges to 0 and is bounded by some constant, $\sum_{i=0}^{k(n)-1} |r_i^n|$ has its lim sup bounded by $2C\epsilon^2$; similarly, the lim sup of $\sum_{i=0}^{k(n)-1} |q_i^n|$ is bounded by $C'(T\epsilon + \epsilon^2)$.

The term in the first line of (2.92) can be written:

$$\sum_{i=0}^{k(n)-1} \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)(x(t_{i+1}^n) - x(t_i^n))$$
$$- \sum_{0\le i\le k(n)-1, i\in I_1^n(\epsilon)} \nabla_x F_{t_i^n}(x_{t_i^n-}^{n,\Delta x(t_i^n)}, v_{t_i^n-}^n)(x_{t_{i+1}^n} - x_{t_i^n}) \tag{2.95}$$

where the second term converges to $\sum_{0<u\leq T, u\in C_1(\epsilon)} \nabla_x F_u(x_{u-}, v_{u-})\Delta x(u)$.

- The third line of (2.79) yields:

$$\int_0^T \mathcal{D}_t F_u(x^n_{t^n_{i^n(u)}, u-t^n_{i^n(u)}}, v^n_{t^n_{i^n(u)}, u-t^n_{i^n(u)}}) 1_{i^n(u)\in I^n_2(\epsilon)} du \tag{2.96}$$

where the integrand converges to $\mathcal{D}_t F_u(x_{u-}, v_{u-})1_{u\notin C_1(\epsilon)}$ and is bounded by C, hence by dominated convergence this term converges to:

$$\int_0^T \mathcal{D}_t F_t(x_{u-}, v_{u-}) du \tag{2.97}$$

Summing up, we have established that the difference between the lim sup and the lim inf of:

$$\sum_{i=0}^{k(n)-1} \nabla_x F_{t^n_i}(x^{n,\Delta x(t^n_i)}_{t^n_i -}, v^n_{t^n_i -})(x(t^n_{i+1}) - x(t^n_i)) \tag{2.98}$$

is bounded by $C''(\epsilon^2 + T\epsilon)$. Since this is true for any ϵ, this term has a limit.

Let us now write the equality we obtained for a fixed ϵ:

$$\begin{aligned} F_T(x_T, v_{tT}) - F_0(x_0, v_0) &= \int_{]0,T]} \mathcal{D}_t F_t(x_{u-}, v_{u-}) du \\ &+ \int_{]0,T]} \frac{1}{2}\mathrm{tr}[{}^t\nabla^2_x F_t(x_{u-}, v_{u-}) d[x]^c(u)] \\ + \lim_n &\sum_{i=0}^{k(n)-1} \nabla_x F_{t^n_i}(x^{n,\Delta x(t^n_i)}_{t^n_i -}, v^n_{t^n_i -})(x(t^n_{i+1}) - x(t^n_i)) \\ + \sum_{u\in]0,T]\cup C_1(\epsilon)} &[F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) - \nabla_x F_u(x_{u-}, v_{u-})\Delta x(u)] + \alpha(\epsilon) \end{aligned}$$

where $\alpha(\epsilon) \leq C''(\epsilon^2 + T\epsilon)$. The only point left to show is that:

$$\sum_{u\in]0,T]\cup C_1(\epsilon)} [F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) - \nabla_x F_u(x_{u-}, v_{u-})\Delta x(u)] \tag{2.99}$$

converges to:

$$\sum_{u\in]0,T]} [F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) - \nabla_x F_u(x_{u-}, v_{u-})\Delta x(u)] \tag{2.100}$$

which is to say that the sum above is absolutely convergent.

We can first choose $d(K, U^c) > \eta > 0$ such that:

$$\forall u \in [0,T], \quad \forall (x', v') \in \mathcal{U}_u \times \mathcal{S}_u, d_\infty((x_t, v_t), (x', v')) \leq \eta \Rightarrow |\nabla^2_x F_u(x(u), v(u))| < C$$

The jumps of x of magnitude greater than η are in finite number. Then, if u is a jump time of x of magnitude less than η, then $x(u-) + h\Delta x(u) \in U$ for $h \in [0,1]$, so that we can write:

$$F_u(x_u, v_u) - F_u(x_{u-}, v_{u-}) - \nabla_x F_u(x_{u-}, v_{u-})\Delta x(u) =$$
$$\int_0^1 (1-v)[{}^t\nabla_x^2 F_u(x_{u-}^{h\Delta x(u)}, v_{u-})^t \Delta x(u)\Delta x(u)] \ \leq \frac{1}{2}C|\Delta x(u)|^2$$

Hence, the theorem is established.

□

Remark 2.10. If the vertical derivatives are right-continuous instead of left-continuous, and without requiring (2.24) for $\nabla_x F$ we can still define:

$$x^n(t) = \sum_{i=0}^{k(n)-1} x(t_i) 1_{[t_i, t_{i+1})}(t) + x(T)1_{\{T\}}(t)$$
$$v^n(t) = \sum_{i=0}^{k(n)-1} v(t_i) 1_{[t_i, t_{i+1})}(t) + v(T)1_{\{T\}}(t) \qquad h_i^n = t_{i+1}^n - t_i^n \tag{2.101}$$

Following the same argument as in the proof with the decomposition:

$$\begin{aligned} F_{t_{i+1}^n}(x^n_{t^n_{i+1}}, v_{t^n_{i+1}}) - F_{t_i^n}(x^n_{t^n_i}, v_{t^n_i}) &= F_{t_{i+1}^n}(x^n_{t^n_{i+1}}, v_{t^n_{i+1}}) - F_{t_{i+1}^n}(x^n_{t^n_{i+1}}, v_{t^n_i, h^n_i}) \\ &+ F_{t_{i+1}^n}(x^n_{t^n_{i+1}}, v_{t^n_i, h^n_i}) - F_{t_{i+1}^n}(x^n_{t^n_i, h^n_i}, v_{t^n_i, h^n_i}) \\ &+ F_{t_{i+1}^n}(x^n_{t^n_i, h^n_i}, v_{t^n_i, h^n_i}) - F_{t_i^n}(x^n_{t^n_i}, v_{t^n_i}) \end{aligned} \tag{2.102}$$

we obtain an analogue of formula (2.76) where the Föllmer integral (2.75) is replaced by

$$\lim_n \sum_{i=0}^{k(n)-1} \nabla_x F_{t_{i+1}^n}(x^n_{t^n_i, h^i}, v^n_{t^n_i, h^n_i})(x(t^n_{i+1}) - x(t^n_i)) \tag{2.103}$$

2.6 Functionals of Dirichlet processes

A Dirichlet process [30, 12], or finite energy process, on a filtered probability space $(\Omega, \mathcal{B}, (\mathcal{B}_t), \mathbb{P})$ is an adapted cadlag process that can be represented as the sum of a semimartingale and an adapted continuous process with zero quadratic variation along dyadic subdivisions.

For continuous Dirichlet processes, a pathwise Itô calculus was introduced by H. Föllmer in [29, 30, 47]. Coquet, Mémin and Slominski [12] extended these results to discontinuous Dirichlet processes [57]. Using Theorem 2.5 we can extend these results to functionals of Dirichlet processes; this yields in particular a pathwise construction of stochastic integrals for functionals of a Dirichlet process.

Let $Y(t) = X(t) + B(t)$ be a U-valued Dirichlet process defined as the sum of a semimartingale X on some filtered probability space $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$ and B an adapted continuous process B with zero quadratic variation along the dyadic subdivision. We denote by $[X]$ the quadratic variation process associated to X, $[X]^c$ the continuous part of $[X]$, and $\mu(dt\ dz)$ the integer-valued random measure describing the jumps of X (see [39] for definitions).

Let A be an adapted process with S-valued cadlag paths. Note that A need not be a semimartingale.

We call $\Pi_n = \{0 = t_0^n < t_1^n < \ldots < t_{k(n)}^n = T\}$ a *random subdivision* if the t_i^n are stopping times with respect to $(\mathcal{B}_t)_{t\in[0,T]}$.

Proposition 2.2 (Change of variable formula for Dirichlet processes)**.** *Let* $\Pi_n = \{0 = t_0^n < t_1^n < \ldots < t_{k(n)}^n = T\}$ *be any sequence of random subdivisions of* $[0, T]$ *such that*
(i) X *has finite quadratic variation along* Π_n *and* B *has zero quadratic variation along* Π_n *almost-surely,*

$$(ii) \qquad \sup_{t\in[0,T]-\Pi_n} |Y(t) - Y(t-)| + |A(t) - A(t-)| \overset{n\to\infty}{\to} 0 \quad \mathbb{P} - a.s.$$

Then there exists $\Omega_1 \subset \Omega$ *with* $\mathbb{P}(\Omega_1) = 1$ *such that for any non-anticipative functional* $F \in \mathbb{C}^{1,2}$ *satisfying*

1. F *is predictable in the second variable in the sense of* (2.2)

2. $\nabla_x^2 F$ *and* $\mathcal{D}F$ *satisfy the local boundedness property* (2.9)

3. $F, \nabla_x F, \nabla_x^2 F \in \mathbb{F}_l^\infty$

4. $\nabla_x F$ *has the horizontal local Lipschitz property* (2.24),

the following equality holds on Ω_1 *for all* $t \leq T$*:*

$$F_t(Y_t, A_t) - F_0(Y_0, A_0) = \int_{]0,t]} \mathcal{D}_u F(Y_{u-}, A_{u-})du + \int_{]0,t]} \frac{1}{2}\mathrm{tr}[{}^t\nabla_x^2 F_u(Y_{u-}, A_{u-})d[X]^c(u)]$$
$$+ \int_{]0,t]} \int_{R^d} [F_u(Y^z_{u-}, A_u-) - F_u(Y_{u-}, A_{u-}) - z\nabla_x F_u(Y_{u-}, A_{u-})]\mu(du, dz)$$
$$+ \int_{]0,t]} \nabla_x F_u(Y_{u-}, A_{u-}).dY(u) \quad (2.104)$$

where the last term is the Föllmer integral (2.75) *along the subdivision* Π_n*, defined for* $\omega \in \Omega_1$ *by:*

$$\int_{]0,t]} \nabla_x F_u(Y_{u-}, A_{u-}).dY(u) :=$$
$$\lim_n \sum_{i=0}^{k(n)-1} \nabla_x F_{t_i^n}(Y^{n,\Delta Y(t_i^n)}_{t_i^n -}, A^n_{t_i^n -})(Y(t^n_{i+1}) - Y(t^n_i))1_{]0,t]}(t^n_i) \quad (2.105)$$

where (Y^n, A^n) *are the piecewise constant approximations along* Π_n*, defined as in* (2.74)*.*

Moreover, the Föllmer integral with respect to any other random subdivision verifying (i)–(ii), is almost-surely equal to (2.105)*.*

Remark 2.11. Note that the convergence of (2.105) holds over a set Ω_1 which may be chosen independently of the choice of $F \in \mathbb{C}^{1,2}$.

Proof. Let (Π_n) be a sequence of random subdivisions verifying (i)–(ii). Then there exists a set Ω_1 with $\mathbb{P}(\Omega_1) = 1$ such that for $\omega \in \Omega_1$ (X, A) is a cadlag function and (i)-(ii) hold pathwise. Applying Theorem 2.5 to $(Y(., \omega), A(., \omega))$ along the subdivision $\Pi_n(\omega)$ shows that (2.104) holds on Ω_1.

To show independence of the limit in (2.105) from the chosen subdivision, we note that if Π^2_n another sequence of random subdivisions satisfies (i)–(ii), there exists $\Omega_2 \subset \Omega$ with $\mathbb{P}(\Omega_2) = 1$ such that one can apply Theorem 2.5 pathwise for $\omega \in \Omega_2$. So we have

$$\int_{]0,t]} \nabla_x F_u(Y_{u-}, A_{u-}).d^{\Pi^2} Y(u) = \int_{]0,t]} \nabla_x F_u(Y_{u-}, A_{u-}).d^{\Pi} Y(u)$$

on $\Omega_1 \cap \Omega_2$. Since $\mathbb{P}(\Omega_1 \cap \Omega_2) = 1$ we obtain the result. □

2.7 Functionals of semimartingales

Proposition 2.2 applies when X is a semimartingale. We will now show that in this case, under an additional assumption, the pathwise Föllmer integral (2.75) coincides almost-surely with the stochastic integral $\int Y dX$. Theorem 2.5 then yields an Itô formula for functionals of a semimartingale X.

2.7.1 Cadlag semimartingales

Let X be a cadlag semimartingale and A an adapted cadlag process on $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$. We use the notations $[X]$, $[X]^c$, $\mu(dt\ dz)$ defined in Section 2.6.

Proposition 2.3 (Functional Itô formula for a semimartingale)**.** *Let $F \in \mathbb{C}^{1,2}$ be a non-anticipative functional satisfying*

1. *F is predictable in the second variable, i.e. verifies* (2.2)*,*
2. *$\nabla_x F,\ \nabla_x^2 F,\ \mathcal{D}F \in \mathbb{B}$,*
3. *$F, \nabla_x F, \nabla_x^2 F \in \mathbb{F}_l^\infty$,*
4. *$\nabla_x F$ has the horizontal local Lipschitz property 2.24.*

Then:

$$F_t(X_t, A_t) - F_0(X_0, A_0) = \int_{]0,t]} \mathcal{D}_u F(X_{u-}, A_{u-}) du +$$
$$\int_{]0,t]} \frac{1}{2} \mathrm{tr}[{}^t\nabla_x^2 F_u(X_{u-}, A_{u-}) d[X]^c(u)] + \int_{]0,t]} \nabla_x F_u(X_{u-}, A_{u-}).dX(u)$$
$$+ \int_{]0,t]} \int_{R^d} [F_u(X_{u-}^z, A_u-) - F_u(X_{u-}, A_{u-}) - z.\nabla_x F_u(X_{u-}, A_{u-})] \mu(du, dz), \mathbb{P}\text{-}a.s. \quad (2.106)$$

where the stochastic integral is the Itô integral with respect to a semimartingale.

In particular, $Y(t) = F_t(X_t, A_t)$ is a semimartingale.

Remark 2.12. These results yield a non-probabilistic proof for functional Ito formulas obtained for continuous semimartingales [23] using probabilistic methods and extend them to the case of discontinous semimartingales.

Proof. Assume first that the process X does not exit a compact set $K \subset U$, and that A is bounded by some constant $R > 0$. We define the following sequence of stopping times:

$$\tau_0^n = 0$$
$$\tau_k^n = \inf\{u > \tau_{k-1}^n | 2^n u \in \mathbb{N} \text{ or } \quad |A(u) - A(u-)| \vee |X(u) - X(u-)| > \frac{1}{n}\} \wedge T \quad (2.107)$$

Then the coordinate processes X_i and their sums $X_i + X_j$ satisfy the property:

$$\sum_{\tau_i < s} (Z(\tau_i) - Z(\tau_{i-1}))^2 \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} [Z](s) \quad (2.108)$$

in probability. There exists a subsequence of subdivisions such that the convergence happens almost surely for all s rational, and hence it happens almost surely for all s because both sides of (2.108) are right-continuous. Let Ω_1 be the set on which this convergence happens, and on which the paths of X and A are U-valued cadlag functions. For $\omega \in \Omega_1$, Theorem 2.5 applies and yields

$$\begin{aligned} F_t(X_t, A_t) - F_0(X_0, A_0) &= \int_{]0,t]} \mathcal{D}_u F(X_{u-}, A_{u-}) du \\ &+ \int_{]0,t]} \frac{1}{2} \mathrm{tr}[{}^t\nabla_x^2 F_u(X_{u-}, A_{u-}) d[X]^c(u)] \quad (2.109) \\ + \int_{]0,t]} \int_{R^d} [F_u(X_{u-}^z, A_u-) &- F_u(X_{u-}, A_{u-}) - z.\nabla_x F_u(X_{u-}, A_{u-})] \mu(du, dz) \\ &+ \lim_{n\to\infty} \sum_{i=0}^{k(n)-1} \nabla_x F_{\tau_i^n}(X_{\tau_i^n -}^{n,\Delta X(\tau_i^n)}, A_{\tau_i^n -}^n)(X(\tau_{i+1}^n) - X(\tau_i^n)) \end{aligned}$$

It remains to show that the last term, which may also be written as

$$\lim_{n\to\infty} \int_{]0,t]} \sum_{i=0}^{k(n)-1} 1_{]\tau_i^n, \tau_{i+1}^n]}(t) \quad \nabla_x F_{\tau_i^n}(X_{\tau_i^n -}^{n,\Delta X(\tau_i^n)}, A_{\tau_i^n -}^n).dX(t) \quad (2.110)$$

coincides with the (Ito) stochastic integral of $\nabla_x F(X_{u-}, A_{u-})$ with respect to the semimartingale X.

First, we note that since X, A are bounded and $\nabla_x F \in \mathbb{B}$, $\nabla_x F(X_{u-}, A_{u-})$ is a bounded predictable process (by Theorem 2.1) hence its stochastic integral $\int_0^{\cdot} \nabla_x F(X_{u-}, A_{u-}).dX(u)$ is well-defined. Since the integrand in (2.110) converges almost surely to $\nabla_x F_t(X_{t-}, A_{t-})$, and is bounded independently of n by a deterministic constant C, the dominated convergence theorem for stochastic integrals [54, Ch.IV Theorem32] ensures that (2.110) converges in probability to $\int_{]0,t]} \nabla_x F_u(X_{u-}, A_{u-}).dX(u)$. Since it converges almost-surely by proposition 2.2, by almost-sure uniqueness of the limit in probability, the limit has to be $\int_{]0,t]} \nabla_x F_u(X_{u-}, A_{u-}).dX(u)$.

Now we consider the general case where X and A may be unbounded. Let $U^c = \mathbb{R}^d - U$ and denote $\tau_n = \inf\{s < t | d(X(s), U^c) \leq \frac{1}{n} \text{ or } |X(s)| \geq n \text{ or } |A(s)| \geq n\} \wedge t$, which are stopping times. Applying the previous result to the stopped processes $(X^{\tau_n -}, A^{\tau_n -}) = (X(t \wedge \tau_n -), A(t \wedge \tau_n -))$ leads to:

$$\begin{aligned} F_t(X_t^{\tau_n -}, A_t^{\tau_n -}) = \int_{]0,\tau^n)} [\mathcal{D}_u F(X_u, A_u)du + \frac{1}{2}\mathrm{tr}[{}^t\nabla_x^2 F_u(X_u, A_u) d[X]^c(u)] \\ + \int_{]0,\tau_n)} \nabla_x F_u(X_u, A_u).dX(u) \\ + \int_{]0,\tau_n)} \int_{R^d} [F_u(X_{u-}^x, A_u -) - F_u(X_{u-}, A_{u-}) - z.\nabla_x F_u(X_{u-}, A_{u-})]\mu(du\ dz) \\ + \int_{(\tau^n, t)} \mathcal{D}_u F(X_u^{\tau_n}, A_u^{\tau_n}) du \end{aligned} \tag{2.111}$$

Since almost surely $t \wedge \tau_n = t$ for n sufficiently large, taking the limit $n \to \infty$ yields:

$$\begin{aligned} F_t(X_{t-}, A_{t-}) = \int_{]0,t)} [\mathcal{D}_u F(X_u, A_u)du + \frac{1}{2}\mathrm{tr}\left({}^t\nabla_x^2 F_u(X_u, A_u) d[X]^c(u)\right) \\ + \int_{]0,t)} \nabla_x F_u(X_u, A_u).dX(u) \\ + \int_{]0,t)} \int_{R^d} [F_u(X_{u-}^x, A_u -) - F_u(X_{u-}, A_{u-}) - z.\nabla_x F_u(X_{u-}, A_{u-})]\mu(du\ dz) \end{aligned}$$

Adding the jump $F_t(X_t, A_t) - F_t(X_{t-}, A_{t-})$ to both the left-hand side and the third line of the right-hand side, and adding $\nabla_x F_t(X_{t-}, A_{t-})\Delta X(t)$ to the second line and subtracting it from the third, leads to the desired result. □

2.7.2 Continuous semimartingales

In the case of a continuous semimartingale X and a continuous adapted process A, an Itô formula may also be obtained for functionals whose vertical derivative is right-continuous rather than left-continuous.

Proposition 2.4 (Functional Itô formula for a continuous semimartingale)**.** *Let X be a continuous semimartingale with quadratic variation process $[X]$, and A a continuous adapted process, on some filtered probability space $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$. Then for any non-anticipative functional $F \in \mathbb{C}^{1,2}$ satisfying*

1. *F has a predictable dependence with respect to the second variable, i.e. verifies* (2.2)*,*

2. $\nabla_x F,\ \nabla_x^2 F,\ \mathcal{D}F \in \mathbb{B}$*,*

3. $F \in \mathbb{F}_l^\infty$

4. $\nabla_x F, \nabla_x^2 F \in \mathbb{F}_r^\infty$

we have

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

where last term is the Itô stochastic integral with respect to the X.

Proof. Assume first that X does not exit a compact set $K \subset U$ and that A is bounded by some constant $R > 0$. Let $0 = t_0^n \leq t_1^n \ldots \leq t_{k(n)}^n = t$ be a deterministic subdivision of $[0, t]$. Define the approximates (X^n, A^n) of (X, A) as in remark 2.10, and notice that, with the same notations:

$$\sum_{i=0}^{k(n)-1} \nabla_x F_{t_{i+1}^n}(X_{t_i^n, h_i^n}^n, A_{t_i^n, h_i^n}^n)(X(t_{i+1}^n) - X(t_i^n)) =$$
$$\int_{]0,t]} \nabla_x F_{t_{i+1}^n}(X_{t_i^n, h_i^n}^n, A_{t_i^n, h_i^n}^n) 1_{]t_i^n, t_{i+1}^n]}(t) dX(t) \tag{2.112}$$

which is a well-defined stochastic integral since the integrand is predictable (left-continuous and adapted by theorem 2.1), since the times t_i^n are *deterministic*; this would not be the case if we had to include jumps of X and/or A in the subdivision as in the case of the proof of proposition 2.3. By right-continuity of $\nabla_x F$, the integrand converges to $\nabla_x F_t(X_t, A_t)$. It is moreover bounded independently of n and ω since $\nabla_x F$ is assumed to be boundedness-preserving. The dominated convergence theorem for the stochastic integrals [54, Ch.IV Theorem32] ensures that it converges in probability to $\int_{]0,t]} \nabla_x F_u(X_{u-}, A_{u-}).dX(u)$. Using remark 2.10 concludes the proof.

Consider now the general case. Let K_n be an increasing sequence of compact sets with $\bigcup_{n \geq 0} K_n = U$ 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:

$$F_t(X_{t\wedge\tau_n}, A_{t\wedge\tau_n}) - F_0(X_0, A_0) = \int_0^{t\wedge\tau_n} \mathcal{D}_u F_u(X_u, A_u) du$$
$$+\frac{1}{2}\int_0^{t\wedge\tau_n} \operatorname{tr}\left({}^t\nabla_x^2 F_u(X_u, A_u) d[X](u)\right) + \int_0^{t\wedge\tau_n} \nabla_x F_u(X_u, A_u).dX$$
$$+\int_{t\wedge\tau^n}^{t} \mathcal{D}_u F(X_{u\wedge\tau_n}, A_{u\wedge\tau_n}) du$$

The terms in the first line converge 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. □

Chapter 3

Functional Itô calculus and applications

3.1 Functionals representation of non-anticipative processes

From this section to the end of this work, the set S is taken to be the set S_d^+ of positive symmetric $d \times d$ matrices.

Let $X : [0,T] \times \Omega \mapsto U$ be a continuous, $U-$valued cadlag semimartingale defined on a filtered probability space $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$. The paths of X then lie in $\mathcal{U}_T^c$, which we will view as a subspace of the space $\mathcal{U}_T$ of cadlag functions with values in U.

Denote by $\mathcal{F}_t$ he 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

$$[X](t) = \int_0^t A(s)ds \tag{3.1}$$

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 in S_d^+.

A process

$Y : [0,T] \times \Omega \mapsto \mathbb{R}^d$ which is adapted to $\mathcal{F}_t$ may be represented almost surely as

$$Y(t) = F_t(\{X(u), 0 \leq u \leq t\}, \{A(u), 0 \leq u \leq t\}) = F_t(X_t, A_t) \tag{3.2}$$

where $F = (F_t)_{t\in[0,T]}$ is a family of functionals on Υ 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 $\mathcal{U}_t$. Introducing A_t as a second argument in the functional allows 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 the supremum norm and predictability in the second variable (see Section 2.2.2).

Remark 3.1. All results presented in this chapter apply to functionals depending only on their first argument, ie functionals on $\bigcup_{t\in[0,T]} \mathcal{U}_t$, without requiring the assumption that $[X](t)$ can be represented as $[X](t) = \int_0^t A(s)ds$ (except for section 3.2.4). If dealing with such a functional, we will omit the second variable and write simply $F_t(x_t)$, rather than $F_t(x_t, v_t)$. All results presented in this chapter also extend to the case of right-continuous rather than left-continuous vertical derivatives in the case where either A is continuous of F does not depend on the second argument (with sometimes a minor modification of the statement, which will be given as a remark).

3.2 Functional Itô calculus

3.2.1 Space of paths

We shall introduce the following space of paths which is the one required for the probabilistic applications of functional Itô calculus. In chapter 2, all space of functionals introduced are defined up to the time-horizon T. We will here denote, with a slight abuse of notation, for $t_0 < T$, $\mathbb{F}_l^\infty([0,t_0]), \mathbb{B}([0,t_0]), \mathbb{C}^{a,b}([0,t_0])$ as the sets of non-anticipative functionals indexed by $[0,T]$ or $[0,T[$, but whose restriction to $\bigcup_{0\le t\le t_0} \mathcal{U}_t \times \mathcal{S}_t$ belong respectively to $\mathbb{F}_l^\infty, \mathbb{B}, \mathbb{C}^{a,b}$.

Definition 3.1. Define $\mathbb{C}_b^{j,k}([0,T[)$ as the set of non-anticipative functionals F such that:

1. For all $t_0 < T, F \in \mathbb{C}^{j,k}([0,t_0])$

2. For all $t_0 < T, 0 \le i \le k, \nabla_x^i F \in \mathbb{F}_l^\infty([0,t_0])$

3. For all $t_0 < T, 1 \leq i \leq j, 1 \leq l \leq k, \mathcal{D}^i_F, \nabla^l_x F \in \mathbb{B}([0, t_0])$

4. F is predictable in the second variable.

Note that the examples discussed in the synopsis (1.1) with explicit dependence in the quadratic variation are continuous in the quadratic variation for the total variation norm:

$$|[X]|_{TV} = \int_0^T |A(s)| ds \tag{3.3}$$

so they define in particular functionals which are predictable in the second variable.

Example 3.1 (Smooth functions). Let us start by noting that, in the case where F reduces to a smooth *function* of $X(t)$,

$$F_t(x_t, v_t) = f(t, x(t)) \tag{3.4}$$

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

$$\mathcal{D}^i_t F(x_t, v_t) = \partial^i_t f(t, x(t)) \qquad \nabla^m_x F_t(x_t, v_t) = \partial^m_x f(t, x(t)) \tag{3.5}$$

In fact $F \in \mathbb{C}^{j,k}$ simply requires f to be j times right-differentiable in time, that the right-derivatives in time are continuous in space for each fixed time, and that the functional and its derivatives in space are jointly left-continuous in time and continuous in space.

Example 3.2 (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

$$F_t(x_t, v_t) = \int_0^t g(x(u)) v(u) du \tag{3.6}$$

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

$$\mathcal{D}_t F(x_t, v_t) = g(x(t)) v(t) \qquad \nabla^j_x F_t(x_t, v_t) = 0 \tag{3.7}$$

Example 3.3. The martingale $Y(t) = X(t)^2 - [X](t)$ is represented by the functional

$$F_t(x_t, v_t) = x(t)^2 - \int_0^t v(u)du \tag{3.8}$$

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

$$\mathcal{D}_t F(x,v) = -v(t) \qquad \nabla_x F_t(x_t, v_t) = 2x(t)$$
$$\nabla_x^2 F_t(x_t, v_t) = 2 \qquad \nabla_x^j F_t(x_t, v_t) = 0, j \geq 3 \tag{3.9}$$

Example 3.4 (Stochastic exponential). The stochastic exponential $Y = \exp(X - [X]/2)$ may be represented by the functional

$$F_t(x_t, v_t) = e^{x(t) - \frac{1}{2}\int_0^t v(u)du} \tag{3.10}$$

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

$$\mathcal{D}_t F(x,v) = -\frac{1}{2} v(t) F_t(x,v) \qquad \nabla_x^j F_t(x_t, v_t) = F_t(x_t, v_t) \tag{3.11}$$

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 3.2, 3.3 and 3.4 as a left-continuous functional of x alone.

Example 3.5 (Cylindrical functionals). Let $0 = t_0 < t_1 < \ldots < t_n = T$ be a subdivision of $[0, T]$, $\epsilon > 0$ and $f_i : 1 \leq i \leq n$ a family of continuous functions such that:

$$f_i : U^{i+1} \times [t_i, t_{i+1} + \epsilon[\mapsto \mathbb{R}$$

such that:

- For all $1 \leq i \leq n$ and all $(x_0, \ldots, x_{i-1}) \in U^i$), the map

  $$(x, t) \mapsto f_i(x_0, \ldots, x_{i-1}, x, t)$$

  defined on $U \times [t_i, t_{i+1}]$ is $C^{1,2}$
- For all $2 \leq i \leq n$, $f_i(., t_i) = f_{i-1}(., t_i)$

Then the functional

$$F_t(x_t) = \sum_{i=1}^{n} 1_{]t_{i-1},t_i]} f_i(x(t_0), \ldots, x(t_{i-1}), x(t), t)$$

is $\mathbb{C}_b^{1,2}$.

Finally, we conclude by an example showing the non-uniqueness of the functional representation:

Example 3.6 (Non-uniqueness of functional representation). Take $d = 1$. The quadratic variation process $[X]$ may be represented by the following functionals:

$$\begin{aligned} F^0(x_t, v_t) &= \int_0^t v(u)du \\ F^1(x_t, v_t) &= \left(\limsup_n \sum_{i \le t2^n} |x(\frac{i+1}{2^n}) - x(\frac{i}{2^n})|^2 \right) \mathbf{1}_{\limsup_n \sum_{i \le t2^n} (x(\frac{i+1}{2^n}) - x(\frac{i}{2^n}))^2 < \infty} \end{aligned} \tag{3.12}$$

$$F^2(x_t, v_t) =$$

$$\left(\limsup_n \sum_{i \le t2^n} |x(\frac{i+1}{2^n}) - x(\frac{i}{2^n})|^2 - \sum_{0 \le s < t} |\Delta x(s)|^2 \right) \mathbf{1}_{\limsup_n \sum_{i \le t2^n} |x(\frac{i+1}{2^n}) - x(\frac{i}{2^n})|^2 < \infty}$$

If X is a continuous semimartingale, then almost surely:

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

Yet $F^0 \in \mathbb{C}_b^{1,2}([0,T[)$ but F^1, F^2 are not even continuous at fixed time: $F^i \notin \mathbb{F}_l^\infty$ for $i = 1, 2$.

3.2.2 Obstructions to regularity

It is instructive to observe what prevents a functional from being regular in the sense of Definition 3.1. The examples below illustrate the fundamental obstructions to regularity:

Example 3.7 (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 3.8 (Jump of x at the current time). $F_t(x_t, v_t) = \Delta x(t)$ defines a functional which is infinitely differentiable and has regular pathwise derivatives:

$$\mathcal{D}_t F(x_t, v_t) = 0 \qquad \nabla_x F_t(x_t, v_t) = 1 \tag{3.13}$$

However, the functional itself fails to be F_l^∞.

Example 3.9 (Jump of x at a fixed time). $F_t(x_t, v_t) = 1_{t>t_0}\Delta x(t_0)$ defines a functional in $\mathbb{F}_l^\infty$ 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 left continuous so F is not $\mathbb{C}_b^{0,1}$ in the sense of Definition 2.8

Example 3.10 (Maximum). $F_t(x_t, v_t) = \sup_{s\le t} x(s)$ is $\mathbb{F}_l^\infty$ but fails to be vertically differentiable on the set

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

3.2.3 Functional Itô formula

We will here restate the functional Itô formula in the context of continuous semimartingales. This is of course a direct consequence of Theorem 2.4, but we will propose here a direct probabilistic derivation which make use of the standard Itô formula and the dominated convergence theorems for Lebesgue - Stieltjes and Stochastic integrals. Using these tools, we do not have to worry about the measure-theoretic technicalities treated in appendix A.3 that were necessary to derive the functional analytical Itô formula (theorem 2.4).

Theorem 3.1 (Functional Itô formula for continuous semimartingales)**.** *Let $F \in \mathbb{C}_b^{1,2}([0, T[)$. Then for any $t \in [0, T[$:*

$$\begin{aligned} F_t(X_t, A_t) - F_0(X_0, A_0) &= \int_0^t \mathcal{D}_u F(X_u, A_u)du + \int_0^t \nabla_x F_u(X_u, A_u).dX(u) \\ &+ \int_0^t \frac{1}{2}\mathrm{tr}[{}^t\nabla_x^2 F_u(X_u, A_u)\ d[X](u)] \qquad a.s. \end{aligned} \tag{3.14}$$

We note that:

- The dependence of F on the second variable A does not enter the formula (3.14). Indeed, under our regularity assumptions, variations in A lead to “higher order” terms which do not contribute. This is due to F being predictable in the second variable.

- As expected from Theorems 2.2 and 2.3 in the case where X is continuous $Y = F(X, A)$ 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 $\bigcup_{t\in[0,T[} \mathcal{U}_t^c \times \mathcal{S}_t$.

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 A of size greater or equal to $\frac{1}{n}$. We define the following sequence of stopping times:

$$\tau_0^n = 0 \qquad \tau_k^n = \inf\{s > \tau_{k-1}^n | 2^n s \in \mathbb{N} \text{ or } \quad |A(s) - A(s-)| > \frac{1}{n}\} \wedge t \tag{3.15}$$

The following arguments apply pathwise. Lemma A.2 ensures that $\eta_n = \sup\{|A(u) - A(\tau_i^n)| + |X(u) - X(\tau_i^n)| + \frac{t}{2^n}, i \leq 2^n, u \in [\tau_i^n, \tau_{i+1}^n)\} \to_{n\to\infty} 0$.

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 non-adapted 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.

Start with the decomposition:

$$\begin{aligned} F_{\tau_{i+1}^n}({}_nX_{\tau_{i+1}^n-}, {}_nA_{\tau_{i+1}^n-}) - F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) = \\ F_{\tau_{i+1}^n}({}_nX_{\tau_{i+1}^n-}, {}_nA_{\tau_i^n, h_i^n}) - F_{\tau_i^n}({}_nX_{\tau_i^n}, {}_nA_{\tau_i^n}) \\ + F_{\tau_i^n}({}_nX_{\tau_i^n}, {}_nA_{\tau_i^n-}) - F_{\tau_i^n}({}_nX_{\tau_i^n-}, {}_nA_{\tau_i^n-}) \end{aligned} \tag{3.16}$$

where we have used the fact that F is predictable 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 can be written $\psi(h_i^n) - \psi(0)$ where:

$$\psi(u) = F_{\tau_i^n+u}({}_nX_{\tau_i^n,u}, {}_nA_{\tau_i^n,u}) \tag{3.17}$$

Since $F \in \mathbb{C}^{1,2}([0,T[)$, ψ is right-differentiable, and moreover by lemma 4, ψ is left-continuous, so:

$$F_{\tau_{i+1}^n}({}_nX_{\tau_i^n,h_i^n}, {}_nA_{\tau_i^n,h_i^n}) - F_{\tau_i^n}({}_nX_{\tau_i^n}, {}_nA_{\tau_i^n}) = \int_0^{\tau_{i+1}^n - \tau_i^n} \mathcal{D}_{\tau_i^n+u} F({}_nX_{\tau_i^n,u}, {}_nA_{\tau_i^n,u}) du \tag{3.18}$$

The second term in (3.16) can be written $\phi(X(\tau_{i+1}^n) - X(\tau_i^n)) - \phi(0)$ where $\phi(u) = F_{\tau_i^n}({}_nX^u_{\tau_i^n-}, {}_nA_{\tau_i^n})$. Since $F \in \mathbb{C}^{1,2}([0,T])$, ϕ is a C^2 functional parameterized by a $\mathcal{F}_{\tau_i}$-measurable random variable, and $\phi'(u) = \nabla_x F_{\tau_i^n}({}_nX^u_{\tau_i^n-}, {}_nA_{\tau_i^n,h_i})$,

$\phi''(u) = \nabla_x^2 F_{\tau_i^n}({}_nX^u_{\tau_i^n-}, {}_nA_{\tau_i^n,h_i})$. Applying the Itô formula to ϕ between times 0 and $\tau_{i+1}-\tau_i$ and the $(\mathcal{F}_{\tau_i+s})_{s\geq 0})$ continuous semimartingale $(X(\tau_i+s))_{s\geq 0}$, yields:

$$\phi(X(\tau_{i+1}^n) - X(\tau_i^n)\ ) - \phi(0) = \int_{\tau_i^n}^{\tau_{i+1}^n} \nabla_x F_{\tau_i^n}({}_nX^{X(s)-X(\tau_i^n)}_{\tau_i^n-}, {}_nA_{\tau_i^n})dX(s)$$
$$+\frac{1}{2}\int_{\tau_i^n}^{\tau_{i+1}^n} \mathrm{tr}[{}^t\nabla_x^2 F_{\tau_i^n}({}_nX^{X(s)-X(\tau_i^n)}_{\tau_i^n-}, {}_nA_{\tau_i^n})d[X](s) \tag{3.19}$$

Summing over $i=0$ to ∞ and denoting $i(s)$ the index such that $s\in[\tau^n_{i(s)},\tau^n_{i(s)+1})$, we have shown:

$$F_t({}_nX_t, {}_nA_t) - F_0(X_0,A_0) = \int_0^t \mathcal{D}_sF({}_nX_{\tau^n_{i(s)},s-\tau^n_{i(s)}}, {}_nA_{\tau^n_{i(s)},s-\tau^n_{i(s)}})ds$$
$$+\int_0^t \nabla_x F_{\tau^n_{i(s)+1}}({}_nX^{X(s)-X(\tau^n_{i(s)})}_{\tau^n_{i(s)}-}, {}_nA_{\tau^n_{i(s)},h_{i(s)}})dX(s)$$
$$+[\frac{1}{2}\int_0^t \mathrm{tr}\left(\nabla_x^2 F_{\tau^n_{i(s)}}({}_nX^{X(s)-X(\tau^n_{i(s)})}_{\tau^n_{i(s)}-}, {}_nA_{\tau^n_{i(s)}}).d[X](s))\right) \tag{3.20}$$

$F_t({}_nX_t, {}_nA_t)$ converges to $F_t(X_t,A_t)$ almost surely. All the approximations of (X,A) appearing in the various integrals have a d_∞-distance from (X_s,A_s) less than η_n hence all the integrands appearing in the above integrals converge respectively to $\mathcal{D}_sF(X_s,A_s)$,$\nabla_xF_s(X_s,A_s)$, $\nabla_x^2F_s(X_s,A_s)$ as $n\to\infty$ by fixed time continuity for $\mathcal{D}F$ and d_∞ left-continuity for the vertical derivatives. Since the derivatives are in $\mathbb{B}$ the integrands in the various above integrals are bounded by a constant dependent only on F,K and R and t hence does not depend on s nor on ω. The dominated convergence and the dominated convergence theorem for the stochastic integrals [54, 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 (3.14) as $n\to\infty$.

Consider now the general case where X and A may be unbounded. Let $U^c=\mathbb{R}^d-U$ and denote $\tau_n=\inf\{s<t|d(X(s),U^c)\leq\frac{1}{n}$ or $|X(s)|\geq n$ or $|A(s)|\geq n\}\wedge t$, which is a stopping time. Applying the previous result to the stopped process $(X_{t\wedge\tau_n},A_{t\wedge\tau_n})$ leads to:

$$F_t(X_{t\wedge\tau_n},A_{t\wedge\tau_n}) - F_0(Z_0,A_0) = \int_0^{t\wedge\tau_n}\mathcal{D}_uF_u(X_u,A_u)du$$
$$+\frac{1}{2}\int_0^{t\wedge\tau_n}\mathrm{tr}\left({}^t\nabla_x^2F_u(X_u,A_u)d[X](u)\right)$$
$$+\int_0^{t\wedge\tau_n}\nabla_xF_u(X_u,A_u).dX + \int_{t\wedge\tau^n}^t D_uF(X_{u\wedge\tau_n},A_{u\wedge\tau_n})du \tag{3.21}$$

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. $\square$

Example 3.11. If $F_t(x_t, v_t) = f(t, x(t))$ where $f \in C^{1,2}([0,T] \times \mathbb{R}^d)$, (3.14) reduces to the standard Itô formula.

Example 3.12. For integral functionals of the form

$$F_t(x_t, v_t) = \int_0^t g(x(u))v(u)du \tag{3.22}$$

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

$$F_t(X_t, A_t) = \int_0^t g(X(u))A(u)du \tag{3.23}$$

since the vertical derivatives are zero in this case.

Example 3.13. 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 Itô product formula:

$$X(t)^2 - [X](t) = \int_0^t 2X.dX \tag{3.24}$$

Example 3.14. For the stochastic exponential functional (Ex. 3.4)

$$F_t(x_t, v_t) = e^{x(t) - \frac{1}{2}\int_0^t v(u)du} \tag{3.25}$$

the formula (3.14) yields the well-known integral representation

$$\exp(X(t) - \frac{1}{2}[X](t)\ ) = \int_0^t e^{X(u) - \frac{1}{2}[X](u)} dX(u) \tag{3.26}$$

3.2.4 Intrinsic nature of the vertical derivative

Whereas the functional representation (3.2) of a ($\mathcal{F}_t$−adapted) process Y is not unique, Theorem 3.1 implies that the process $\nabla_x F_t(X_t, A_t)$ has an intrinsic character i.e. independent of the chosen representation:

Corollary 3.1. *Let F^1, F^2 be two functionals in $\mathbb{C}_b^{1,2}([0,T[)$, such that:*

$$\forall t < T, \qquad F_t^1(X_t, A_t) = F_t^2(X_t, A_t) \qquad \mathbb{P} - a.s. \tag{3.27}$$

Then, outside an evanescent set:

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

Proof. Let $X(t) = B(t) + M(t)$ where B is a continuous process with finite variation and M is a continuous local martingale. Theorem 3.1 implies that the local martingale part of the null process $0 = F^1(X_t,A_t) - F^2(X_t,A_t)$ can be written:

$$0 = \int_0^t \left[\nabla_x F_u^1(X_u,A_u) - \nabla_x F_u^2(X_u,A_u)\right] dM(u) \tag{3.28}$$

Considering its quadratic variation, we have almost surely:

$$0 = \int_0^t {}^t[\nabla_x F_u^1(X_u,A_u) - \nabla_x F_u^2(X_u,A_u)]A(u-)[\nabla_x F_u^1(X_u,A_u) - \nabla_x F_u^2(X_u,A_u)]du \tag{3.29}$$

Let $\Omega_1 \subset \Omega$ a set of probability 1, in which the above integral is 0 and in which the path of X is continuous and the path of A is right-continuous. For $\omega \in \Omega_1$, the integrand in (3.29) is left-continuous by proposition 4 ($\nabla_x F^1(X_t,A_t) = \nabla_x F^1(X_{t-},A_{t-})$ because X is continuous and F is predictable in the second variable), this yields that, for all $t < T$ 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 $0 < t < T$, $A(t-)$ is almost surely positive definite, Corollary 3.1 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 3.2 (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 $\mathbb{C}_b^{1,2}([0,T[) \cap \mathbb{F}_l^\infty([0,T])$:

$$\mathcal{C}_b^{1,2}(X) = \{Y, \quad \exists F \in \mathbb{C}_b^{1,2}([0,T[) \cap \mathbb{F}_l^\infty([0,T]), \quad Y(t) = F_t(X_t,A_t) \quad \mathbb{P}-\text{ a.s.}\} \tag{3.30}$$

If $A(t-)$ is almost-surely non-singular 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 $F \in \mathbb{C}_b^{1,2}([0,T[) \cap \mathbb{F}_l^\infty([0,T])$ in the representation (3.2). 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 3.3 (Vertical derivative of non-anticipative Brownian functionals)**.** Let W be a standard d-dimensional Brownian motion. For any $Y \in \mathcal{C}_b^{1,2}(W)$ with representation $Y(t) = F_t(W_t)$, the predictable process

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

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

3.3 Martingale representation formula

The functional Itô formula (Theorem 3.1) then leads to an explicit martingale representation formula for $\mathcal{F}_t$-martingales in $\mathcal{C}_b^{1,2}(X)$. This result may be seen as a non-anticipative counterpart of the Clark-Haussmann-Ocone formula [9, 51, 34] and generalizes explicit martingale representation formulas previously obtained in a Markovian context by Elliott and Kohlmann [27] and Jacod et al. [38].

3.3.1 Martingale representation theorem

Assume in this section that X is a local martingale. 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 admits a representation $Y(t) = F_t(X_t, A_t)$ where $F \in \mathbb{C}_b^{1,2}([0,T[) \times \mathbb{F}_l^\infty([0,T])$, we obtain the following explicit martingale representation theorem:

Theorem 3.2. *If $Y(t) = F_t(X_t, A_t)$ for some functional $F \in \mathbb{C}_b^{1,2}([0,T[) \times \mathbb{F}_l^\infty([0,T])$, then:*

$$Y(T) = E[Y(T)] + \int_0^T \nabla_x F_t(X_t, A_t) dX(t) \tag{3.31}$$

Note that regularity assumptions are given 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 3.1 implies that for $t \in [0, T[$:

$$Y(t) = \int_0^t \mathcal{D}_u F(X_u, A_u) du + \frac{1}{2} \int_0^t \mathrm{tr}[{}^t\nabla_x^2 F_u(X_u, A_u) d[X](u)] + \int_0^t \nabla_x F_u(X_u, A_u) dX(u) \tag{3.32}$$

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 $F \in \mathbb{F}_l^\infty([0,T])$ $Y(t)$ has limit $F_T(X_T, A_T)$ as $t \to T$, and on the other hand since $\int_0^t |\partial_i F_u(X_u, A_u)|^2 d[X_i](u) = [Y_i(t)] \to [Y_i(T)] < \infty$, the stochastic integral also converges. □

Example 3.15.

If the stochastic exponential $e^{X(t)-\frac{1}{2}[X](t)}$ is a martingale, applying Theorem 3.2 to the functional $F_t(x_t, v_t) = e^{x(t)-\int_0^t v(u)du}$ yields the familiar formula:

$$e^{X(t)-\frac{1}{2}[X](t)} = 1 + \int_0^t e^{X(s)-\frac{1}{2}[X](s)} dX(s) \tag{3.33}$$

If $X(t)^2$ is integrable, applying Theorem 3.2 to the functional $F_t(x(t), v(t)) = x(t)^2 - \int_0^t v(u)du$, we obtain the well-known Itô product formula

$$X(t)^2 - [X](t) = \int_0^t 2X(s)dX(s) \tag{3.34}$$

3.3.2 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 [48] and developed by Bismut [6, 7], Stroock [59], Shigekawa [56], Watanabe [64] 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 Ω_0 the canonical Wiener space $(C_0([0,T], \mathbb{R}^d), \|.\|_\infty, \mathbb{P})$ endowed with its Borel σ-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 $E[|H|] < \infty$ and define the martingale $Y(t) = E[H|\mathcal{F}_t]$. If H is differentiable in the Malliavin sense

[48, 50, 59] i.e. $H \in \mathbf{D}^{1,1}$ with Malliavin derivative $\mathbb{D}_t H$, then the Clark-Haussmann-Ocone formula [40, 51, 50] gives a stochastic integral representation of the martingale Y in terms of the Malliavin derivative of H:

$$H = E[H] + \int_0^T {}^p E[\mathbb{D}_t H|\mathcal{F}_t] dW_t \tag{3.35}$$

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 [6, 16, 27, 1].

However, as shown by Pardoux and Peng [52, 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 [52, 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 3.2 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 $Y \in \mathcal{C}_b^{1,2}(X)$.

The uniqueness of the integrand in the martingale representation (3.31) leads, with a slight abuse of notations, to:

$$E[\mathbb{D}_t H|\mathcal{F}_t] = \nabla_W \left(E[H|\mathcal{F}_t]\right), \qquad dt \times d\mathbb{P} - a.s. \tag{3.36}$$

Theorem 3.3. *Denote by*

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

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

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

$$\begin{array}{ccc} \mathbb{H} & \stackrel{\mathbb{D}}{\to} & L^1([0,T]\times\Omega) \\ \downarrow_{({}^pE[.|\mathcal{F}_t])_{t\in[0,T]}} & & \downarrow_{({}^pE[.|\mathcal{F}_t])_{t\in[0,T]}} \\ \mathcal{C}_b^{1,2}(W) & \stackrel{\nabla_W}{\to} & \mathcal{P} \end{array}$$

Proof. The Clark-Haussmann-Ocone formula extended to $\mathbf{D}^{1,1}$ in [40] gives

$$H = E[H] + \int_0^T {}^pE[\mathbb{D}_t H|\mathcal{F}_t]dW_t \tag{3.37}$$

where ${}^pE[\mathbb{D}_t H|\mathcal{F}_t]$ denotes the predictable projection of the Malliavin derivative. On other hand theorem 3.2 gives:

$$H = E[H] + \int_0^T \nabla_W E[H|\mathcal{F}_t]dW(t) \tag{3.38}$$

Hence: ${}^pE[\mathbb{D}_t H|\mathcal{F}_t] = \nabla_W E[H|\mathcal{F}_t]$, $dt \times d\mathbb{P}$ almost everywhere. □

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 in a pathwise manner. This implies the usefulness of (3.31) for the numerical computation of martingale representations (see remark 3.5).

3.4 Weak derivatives and integration by parts for stochastic integrals

Assume now that X is a continuous, square-integrable real-valued martingale.

Several authors [35, 65, 14] gave a meaning to the notion of *stochastic derivative* of the stochastic integral $\int_0^t \phi_s dX(s)$ along the path of X, in the Brownian and general continuous semimartingale case, and, with some regularity assumptions on the path of the integrand ϕ, showed that the stochastic derivative of $\int_0^t \phi_s dX(s)$ along the path of X is indeed ϕ_t. The notion of stochastic derivative they defined is a *strong derivative*, in the sense of limit in probability.

In the setting of theorem 3.2, where $Y(t) = F_t(X_t, A_t) = \int_0^t \nabla_x F_s(X_s, A_s) dX(s)$, the stochastic integral $\int_0^t \nabla_x F_s(X_s, A_s) dX(s)$ admits as pathwise *vertical derivative* $\nabla_X Y(t) = \nabla_x F_t(X_t, A_t)$. Both notions of derivatives are *strong derivatives*, but a vertical derivative is a derivative with respect to an instantaneous perturbation of the path of X, while the stochastic derivative is a derivative going forward in time along the path of X; however they coincide since in both cases the derivative of the stochastic integral is the integrand. In this section, we will show that ∇_X may be extended to a *weak derivative* which acts as the inverse of the Itô stochastic integrals, that is, an operator which satisfies

$$\nabla_X \left(\int \phi . dX \right) = \phi, \qquad dt \times d\mathbb{P} - a.s. \tag{3.39}$$

for square-integrable stochastic integrals of the form:

$$Y(t) = \int_0^t \phi_s dX(s) \qquad \text{where} \qquad E\left[\int_0^t \phi_s^2 d[X](s)\right] < \infty \tag{3.40}$$

Remark 3.2. The construction in this section does not require the assumption of absolute continuity for $[X]$, since the functionals used to prove lemma 3.1 do not depend on A. This construction also easily extends to multidimensional case with heavier notations.

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

$$||\phi||^2_{\mathcal{L}^2(X)} = E\left[\int_0^t \phi_s^2 d[X](s)\right] < \infty \tag{3.41}$$

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

$$\mathcal{I}^2(X) = \{\int_0^{\cdot} \phi(t) dX(t), \phi \in \mathcal{L}^2(X)\} \tag{3.42}$$

endowed with the norm

$$||Y||_2^2 = E[Y(T)^2] \tag{3.43}$$

The Itô integral $\phi \mapsto \int_0^{\cdot} \phi_s dX(s)$ is then a bijective isometry from $\mathcal{L}^2(X)$ to $\mathcal{I}^2(X)$ [54].

Definition 3.4 (Space of test processes)**.** The space of *test processes* $D(X)$ is defined as

$$D(X) = \mathcal{C}_b^{1,2}(X) \cap \mathcal{I}^2(X) \tag{3.44}$$

Martingale representation theorem 3.2 allows to define intrinsically the vertical derivative of a process in $D(X)$ as an element of $\mathcal{L}^2(X)$.

Definition 3.5. 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 F.

Theorem 3.4 (Integration by parts on $D(X)$)**.** *Let $Y, Z \in D(X)$. Then:*

$$E\left[Y(T)Z(T)\right] = E\left[\int_0^T \nabla_X Y(t)\nabla_X Z(t)d[X](t)\right] \tag{3.45}$$

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 3.2 to Y and Z, we obtain

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

Applying the Itô isometry formula yields the result. □

Using this result, we can extend the operator ∇_X in a weak sense to a suitable space of the space of (square-integrable) stochastic integrals, where $\nabla_X Y$ is characterized by (3.45) being satisfied against all test processes.

The following definition introduces the Hilbert space $\mathcal{W}^{1,2}(X)$ of martingales on which ∇_X acts as a weak derivative, characterized by integration-by-part formula (3.45). This definition may be also viewed as a non-anticipative counterpart of Wiener-Sobolev spaces in the Malliavin calculus [48, 56].

Definition 3.6 (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 3.1. $\{\nabla_X Y, Y \in D(X)\}$ *is dense in* $\mathcal{L}^2(X)$ *and*

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

Proof. We first observe that the set $\mathfrak{C}$ of "cylindrical" integrands of the form

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

where $n \geq 1$, $0 \leq t_1 < .. < t_n \leq T$ and $f \in C^\infty(\mathbb{R}^n \to \mathbb{R})$, bounded, is a total set in $\mathcal{L}^2(X)$ i.e. the linear span of $\mathfrak{C}$ is dense in $\mathcal{L}^2(X)$.

For such an integrand $\phi_{n,f,(t1,..,t_n)}$, the stochastic integral with respect to X is given by the martingale

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

where the non-anticipative functional F is defined on Υ as:

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

so that:

$$\nabla_x F_t(x_t) = f(x_{t_1-}, ..., x_{t_n-})1_{t>t_n} \in F_l^\infty \cap \mathbb{B}$$

$$\nabla_x^2 F_t(x_t) = 0, \mathcal{D}_t F(x_t, v_t) = 0$$

which prove that $F \in \mathbb{C}_b^{1,2}([0,T[)$. Hence, $Y \in \mathcal{C}_b^{1,2}(X)$. Since f is bounded, Y is obviously square integrable so $Y \in D(X)$. Hence $I_X(\mathfrak{C}) \subset D(X)$.

Since I_X is a bijective isometry from $\mathcal{L}^2(X)$ to $\mathcal{I}^2(X)$, the density of $\mathfrak{C}$ in $\mathcal{L}^2(X)$ entails the density of $I_X(\mathfrak{C})$ in $\mathcal{I}^2(X)$, so $W^{1,2}(X) = \mathcal{I}^2(X)$. □

Remark 3.3. To obtain the result for right-continuous derivatives rather than left-continuous, the functional F in the above proof has to be defined as:

$$F_t(x_t) = f(x(t_1-), ..., x(t_n-))(x(t) - x(t_n-))1_{t \geq t_n}$$

Theorem 3.5 (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*

$$\begin{aligned} \nabla_X : \quad \mathcal{W}^{1,2}(X) &\mapsto \mathcal{L}^2(X) \\ \int_0^T \phi . dX &\mapsto \phi \end{aligned} \tag{3.46}$$

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*

$$\forall Z \in D(X), \qquad E[Y(T)Z(T)] = E\left[\int_0^T \nabla_X Y(t) \nabla_X Z(t) d[X](t)\right]. \tag{3.47}$$

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

$$
\begin{aligned}
I_X : \mathcal{L}^2(X) &\mapsto \mathcal{W}^{1,2}(X) \\
\phi &\mapsto \int_0^{\cdot} \phi.dX
\end{aligned}
\tag{3.48}
$$

in the following sense:

$$\forall \phi \in \mathcal{L}^2(X), \quad \forall Y \in \mathcal{W}^{1,2}(X), \quad < Y, I_X(\phi) >_{\mathcal{W}^{1,2}(X)} = < \nabla_X Y, \phi >_{\mathcal{L}^2(X)} \tag{3.49}$$

$$i.e. \qquad E\left[Y(T)\int_0^T \phi.dX\right] = E\left[\int_0^T \nabla_X Y \ \phi.d[X]\right] \tag{3.50}$$

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 Itô isometry formula then guarantees that (3.47) holds for ϕ. One still needs to prove that (3.47) uniquely characterizes ϕ. If some process ψ also satisfies (3.47), then, denoting $Y' = \mathcal{I}_X(\psi)$ its stochastic integral with respect to X, (3.47) then implies that $U = Y' - Y$ verifies

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

which implies $U = 0$ in $\mathcal{W}^{1,2}(X)$ since by definition $D(X)$ 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 Itô integral on $\mathcal{W}^{1,2}(X)$. □

Thus, Itô'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 ∇_X introduced in Definition 2.8.

Remark 3.4. In other words, we have established that for any $\phi \in \mathcal{L}^2(X)$ the relation

$$\nabla_X \int_0^T \phi(t)dX(t) = \phi(t) \tag{3.51}$$

holds in a weak sense.

Remark 3.5. This result has implications for the numerical computation of the stochastic integral representation of a martingale. Assume that one can simulate simple paths of the martingale X and for each sample path the terminal value Y_T of some $Y \in \mathcal{I}^2(X)$. One can directly evaluate $Y(0)$ by running a Monte-Carlo. This result implies that he can also evaluate its martingale representation $\nabla_X Y(0)$: the terminal value Y_T can be approximated by the terminal value of a sequence of test processes $(F^n_T(X_T, A_T))_{n\geq 0}$ in L^2, and the process $\nabla_x F^n_t(X_t, A_t)$ converges in $\mathcal{L}^2(X)$ to $\nabla_X Y(t)$. Since $\nabla_x F^n$ is the vertical derivative of a $\mathbb{C}^{1,2}_b$ functional it can be estimated by finite differences. Hence combining the approximation of Y_T by terminal value of test processes and finite differences allows numerical computation of a martingale representation.

In particular these results hold when $X = W$ is a Brownian motion. We can now restate a square-integrable version of theorem 3.3, which holds on $\mathbf{D}^{1,2}$, and where the operator ∇_W is defined in the weak sense of theorem 3.5.

Theorem 3.6 (Lifting theorem)**.** *Consider $\Omega_0 = \mathcal{U}^c_T$ endowed with its Borel σ-algebra, the filtration of the canonical process and the Wiener measure $\mathbb{P}$. Then the following diagram is commutative is the sense of $dt \times d\mathbb{P}$ equality:*

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

Remark 3.6. With a slight abuse of notation, the above result can be also written as

$$\forall H \in \mathbf{D}^{1,2}, \qquad \nabla_W \left(E[H|\mathcal{F}_t]\right) = E[\mathbb{D}_t H|\mathcal{F}_t] \tag{3.52}$$

In other words, the conditional expectation operator intertwines ∇_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 ∇_W and the Itô stochastic integral. Obviously, making this last statement precise is a whole research program, beyond the scope of this work.

3.5 Functional equations for martingales

Consider now a semimartingale X satisfying a stochastic differential equation with functional coefficients:

$$dX(t) = b_t(X_t, A_t)dt + \sigma_t(X_t, A_t)dW(t) \tag{3.53}$$

where b, σ are non-anticipative functionals on Υ with values in $\mathbb{R}^d$-valued (resp. $\mathbb{R}^{d\times n}$), whose coordinates are in $\mathbb{F}_l^\infty$. The *topological support* in $(\mathcal{U}_T^c \times \mathcal{S}_T, \|.\|_\infty)$ of the law of (X, A) is defined as the subset $\text{supp}(X, A)$ of all paths $(x, v) \in \mathcal{U}_T^c \times \mathcal{S}_T$ for which every neighborhood has positive measure:

$$\text{supp}(X, A) = \tag{3.54}$$

$$\{(x, v) \in \mathcal{U}_T^c \times \mathcal{S}_T |\text{ for any Borel neighborhood } V \text{ of } (x, v), \mathbb{P}((X, A) \in V) > 0\}$$

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 $F \in \mathbb{C}_b^{1,2}([0, T[) \times \mathbb{F}_l^\infty([0, T])$ which define a *local martingale* as the solution to a functional version of the Kolmogorov backward equation:

Theorem 3.7 (Functional equation for $\mathcal{C}^{1,2}$ martingales)**.** *If $F \in \mathbb{C}_b^{1,2}([0, T[) \times \mathbb{F}_l^\infty([0, T])$, then $Y(t) = F_t(X_t, A_t), t \le T$ is a local martingale if and only if F satisfies the functional differential equation for all $t \in]0, T[$:*

$$\mathcal{D}_t F(x_t, v_t) + b_t(x_t, v_t)\nabla_x F_t(x_t, v_t) + \frac{1}{2}\text{tr}[\nabla_x^2 F(x_t, v_t)\sigma_t{}^t\sigma_t(x_t, v_t)] = 0, \tag{3.55}$$

on the topological support of (X, A) in $(\mathcal{U}_T^c \times \mathcal{S}_T, \|.\|_\infty)$.

Proof. If $F \in \mathbb{C}_b^{1,2}([0, T[) \cap \mathbb{F}_l^\infty([0, T])$, then applying Theorem 3.1 to $Y(t) = F_t(X_t, A_t)$, (3.55) implies that the finite variation term in (3.14) is almost-surely zero:
$Y(t) = \int_0^t \nabla_x F_t(X_t, A_t)dX(t)$, and also Y is continuous up to time T by left-continuity of F. Hence Y is a local martingale.

Conversely, assume that Y is a local martingale. Note that Y is left-continuous by Theorem 4. Suppose the functional relation (3.55) is not satisfied at some (x, v) belongs to the $\text{supp}(X, A) \subset \mathcal{U}_T^c \times \mathcal{S}_T$. Then there exists $t_0 < T$, $\eta > 0$ and $\epsilon > 0$ such that

$$|\mathcal{D}_t F(x'_t, v'_t) + b_t(x'_t, v'_t)\nabla_x F_t(x'_t, v'_t) + \frac{1}{2}\text{tr}[\nabla_x^2 F(x'_t, v'_t)\sigma_t{}^t\sigma_t(x'_t, v'_t)]| > \epsilon \tag{3.56}$$

for $t \leq t_0$ and $d_\infty((x_{t_0}, v_{t_0}), (x'_t, v'_t)) < \eta$, by left-continuity of the expression. It is in particular true for all $t \in [t_0 - \frac{\eta}{2}, t_0]$ and all (x', v') belonging to the following neighborhood of (x, v) in $\mathcal{U}_T^c \times \mathcal{S}_T$:

$$\{(x', v') \in \mathcal{U}_T^c \times \mathcal{S}_T | d_\infty((x, v), (x', v')) < \frac{\eta}{2}\} \tag{3.57}$$

Since (X_T, A_T) belongs to this neighborhood with non-zero probability, it proves that:

$$\mathcal{D}_t F(X_t, A_t) + b_t(X_t, A_t)\nabla_x F_t(x_t, v_t) + \frac{1}{2} tr[\nabla_x^2 F(X_t, A_t)\sigma_t{}^t\sigma_t(X_t, A_t)]| > \frac{\epsilon}{2} \tag{3.58}$$

with non-zero $dt \times d\mathbb{P}$ measure. Applying theorem 3.1 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. □

Remark 3.7. If the vertical derivatives (but not the functional itself) and the coefficients b and σ are right-continuous rather than left-continuous, the theorem is the same with the functional differential equation being satisfied for all $t \in [0, T[$ rather than $]0, T[$; the proof is the same but going forward rather than backward in time from t_0.

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 Υ in general. However, the following result gives a condition for uniqueness of a solution of (3.55) on $\text{supp}(X, A)$:

Theorem 3.8 (Uniqueness of solutions)**.** *Let h be a continuous functional on $(C_0([0,T]) \times \mathcal{S}_T, \|.\|_\infty)$. Any solution $F \in \mathbb{C}_b^{1,2}([0,T[) \times \mathbb{F}_l^\infty([0,T])$ of the functional equation* (3.55)*, verifying*

$$F_T(x, v) = h(x, v) \tag{3.59}$$

$$E[\sup_{t\in[0,T]} |F_t(X_t, A_t)|] < \infty \tag{3.60}$$

is uniquely defined on the topological support $supp(X, A)$ of (X, A) in $(\mathcal{U}_T^c \times \mathcal{S}_T, \|.\|)$: if $F^1, F^2 \in \mathbb{C}_b^{1,2}([0,T])$ verify (3.55)-(3.59)-(3.60) *then*

$$\forall (x, v) \in \text{supp}(X, A), \quad \forall t \in [0, T] \qquad F_t^1(x_t, v_t) = F_t^2(x_t, v_t). \tag{3.61}$$

Proof. Let F^1 and F^2 be two such solutions. Theorem 3.7 shows that they are local martingales. The integrability condition (3.60) 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 3.7 shows that $F_t^1(x_t, v_t) = F_t^2(x_t, v_t)$ if $(x, v) \in \text{supp}(X, A)$. □

Example 3.16. Consider a scalar diffusion

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t) \qquad X(0) = x_0 \tag{3.62}$$

whose law $\mathbb{P}^{x_0}$ is defined as the solution of the martingale problem [61] 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 σ are continuous and bounded functions, with σ bounded away from zero. We are interested in computing the martingale

$$Y(t) = E[\int_0^T g(t, X(t))d[X](t)|\mathcal{F}_t] \tag{3.63}$$

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 [60, Theorem 3.1.] which yields:

$$\{(x, (\sigma^2(t, x(t)))_{t\in[0,T]}) \quad |x \in C_0([0,T], \mathbb{R}^d), x(0) = x_0\}, \tag{3.64}$$

From theorem 3.7 a necessary condition for Y to have a a functional representation $Y = F(X, A)$ with $F \in \mathbb{C}_b^{1,2}([0,T[) \cap \mathbb{F}_l^\infty([0,T])$ is that F verifies

$$\begin{aligned}\mathcal{D}_t F(x_t, (\sigma^2(u, x(u)))_{u\le t}) + b(t, x(t))\nabla_x F_t(x_t, (\sigma^2(u, x(u)))_{u\in[0,t]}) \\ + \frac{1}{2}\sigma^2(t, x(t))\nabla_x^2 F_t(x_t, (\sigma^2(u, x(u)))_{u\in[0,t]}) = 0\end{aligned} \tag{3.65}$$

together with the terminal condition:

$$F_T(x_T, (\sigma^2(u, x(u))_{u\in[0,T]}) = \int_0^T g(t, x(t))\sigma^2(t, x(t))dt \tag{3.66}$$

for all $x \in C_0(\mathbb{R}^d), x(0) = x_0$. Moreover, from theorem 3.8, we know that there any solution satisfying the integrability condition:

$$E[\sup_{t\in[0,T]} |F_t(X_t, A_t)|] < \infty \tag{3.67}$$

is unique on 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 $F \in \mathbb{C}^{1,2}([0,T[) \times \mathbb{F}_l^\infty([0,T])$; so F is solution of the functional equation (3.65) if and only if f satisfies the Partial Differential Equation with source term:

$$\frac{1}{2}\sigma^2(t,x)\partial_x^2 f(t,x) + b(t,x)\partial_x f(t,x) + \partial_t f(t,x) = -g(t,x)\sigma^2(t,x) \tag{3.68}$$
$$\text{with terminal condition} \qquad f(T,x) = 0$$

The existence of a solution f with at most exponential growth is then guaranteed by standard results on parabolic PDEs [44]. In particular, theorem 3.8 guarantees that there is at most one solution such that:

$$E[\sup_{t\in[0,T]} |f(t, X(t))|] < \infty \tag{3.69}$$

Hence the martingale Y in (3.63) 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 (3.68).

3.6 Functional verification theorem for a non-markovian stochastic control problem

3.6.1 Control problem

In this section, we will take the open set $U = \mathbb{R}^d$.

We consider here a simple non-markovian case of stochastic control problem. A good survey paper on stochastic control in the usual Markovian setting is [53], while standard textbook references are [28], [43]. The more general framework of controlled semimartingales has been studied by probabilistic methods in [24]. We will specify here a non-markovian framework for stochastic control based on strong solutions of stochastic differential equations whose coefficients are functionals of the whole trajectory up to date, and show how the functional Itô formula 3.14 links the optimal control problem to a functional version of the Hamilton-Jacobi-Bellman equation. We refer the reader to appendix B for an introduction to stochastic differential equations with functional coefficients that we will consider in this section. Let $(\Omega, \mathbb{B}, \mathbb{P})$ be a probability space, $\mathcal{N}$ the set of null sets, W an n-dimensional Brownian motion on that probability space and $(\mathcal{B}_t)_{t\geq 0}$ the natural left-continuous filtration of the Brownian W, augmented by the null sets.

Definition 3.7. Let A be a subset of $\mathbb{R}^m$, for a filtration $\mathcal{G}$ we define

$$\mathcal{A}_{\mathcal{G}} = \{\alpha = (\alpha(s))_{0\leq s\leq T}, \mathcal{G}\text{-progressively measurable and } A\text{-valued}, E[\int_0^T |\alpha(t)|^2 dt] < \infty\}$$

. $\mathcal{A}_{\mathcal{G}}$ is called the set of A-valued admissible controls for the filtration $\mathcal{G}$.

Let b_t and σ_t be functionals on $C_0([0,t], U) \times C_0([0,t], S_d^+) \times A$,respectively $\mathbb{R}^d$ and $\mathcal{M}^{d,n}$-valued. Conditions will be imposed later on b and σ. Fixing an initial value $(x_0, 0)$, we suppose that for each $\alpha \in \mathcal{A}_{\mathcal{B}}$, the stochastic differential equation:

$$dX(t) = b_t(X_t, [X]_t, \alpha(t))dt + \sigma_t(X_t, [X]_t, \alpha(t))dt \tag{3.70}$$

has a unique strong solution in the sense of definition B.2 denoted X^α. We are also given a real-valued functional g defined on $C_0([0,T], \mathbb{R}^d) \times C_0([0,T], S_d^+)$, and a real-valued functional L defined on $\bigcup_{0\leq t<T} C^0([0,t], \mathbb{R}^d) \times C_0([0,t], S_d^+) \times A$, which are respectively called the terminal value and the penalty function of the control problem. The control problem is finding:

$$\inf_{\alpha\in\mathcal{A}_{\mathcal{B}}} E\left[g(X_T^\alpha, [X^\alpha]_T) + \int_0^T L_t(X_t^\alpha, [X^\alpha]_t, \alpha(t))dt\right] \tag{3.71}$$

as well as an optimal control α^* attaining this infimum. This control setting models a situation where an operator can specify at each time t a control $\alpha(t)$ on the process X in order to minimize an objective function g of this process, and that imposing this control has the infinitesimal cost $L_t(X_t^\alpha, [X]_t^\alpha, \alpha(t))dt$ to the operator.

We will present two twin versions of a functional verification theorem, theorems 3.9 and

3.10. The first one will be given in the case where the coefficients of the stochastic differential equations, as well as the objective and cost function do not depend on the quadratic variation. The second one will allow for such dependence, in the case where the control acts only on the drift term of the diffusion (3.70).

3.6.2 Optimal control and functional Hamilton-Jacobi-Bellman equation, first version

We suppose now that b,σ are as in theorem B.1. For any $\alpha \in \mathcal{A}_{\mathcal{B}}$ and $x \in C_0([0,t_0],\mathbb{R}^d)$, denote $X^{x,\alpha}$ the strong solution of

$$dX(t_0+t) = b_{t_0+t}(X_{t_0+t},\alpha(t))dt + \sigma_{t_0+t}(X_{t_0+t},\alpha(t))dW(t) \tag{3.72}$$

with initial value x, in the sense of definition B.1. Let $x_0 \in \mathbb{R}$, viewed as element of $C_0(\{0\},\mathbb{R}^d)$, we consider the control problem:

$$\inf_{\alpha\in\mathcal{A}_{\mathcal{B}}} E\left[g(X_T^{x_0,\alpha}) + \int_0^T L_t(X_t^{x_0,\alpha},\alpha(t))dt\right] \tag{3.73}$$

where we make the following assumptions on :

1. g is continuous for the sup norm in $\mathcal{U}_T^c$
2. $-K \leq g(x) \leq K(1+\sup_{s\in[0,T]}|x(s)|^2)$ for some constant K
3. $-K' \leq L_t(x,u) \leq K'(1+\sup_{s\in[0,T]}|x(s)|^2$ for some constant K'

The cost functional of the control problem (3.73) is a functional on $\bigcup_{t\leq T} C_0([0,t],\mathbb{R}^d)\times\mathcal{A}_{\mathcal{B}}$ as:

$$J_t(x,\alpha) = E[g(X_T^{x,\alpha}) + \int_t^T L_s(X_s^{x,\alpha},\alpha(s-t))ds] \tag{3.74}$$

It is obvious from assumptions 1. to 3. that $J_t(x,\alpha)$ is finite for any admissible control α, thanks to (B.32).

We define the value functional of the problem on $\bigcup_{t\leq T} C_0([0,t],\mathbb{R}^d)$ as:

$$V_t(x) = \inf_{\alpha\in\mathcal{A}_{\mathcal{B}}} J_t(x,\alpha) \tag{3.75}$$

It is obvious from assumptions 1. to 4. and the L^2 bound (B.32) that $V_t(x)$ is finite and satisfies for some constant K'':

$$-K'' \leq V_t(x) \leq K''(1 + \sup_{s\in[0,t]} |x(s)|^2) \tag{3.76}$$

It is readily observed that the value process

$$U^{\alpha}(t) = V_t(X_t^{x_0,\alpha}) + \int_0^t L_s(X_s^{x_0,\alpha}, \alpha(s))ds$$

has the submartingale property. The martingale approach to stochastic optimal control then characterizes an optimal control $\alpha*$ by the property that the value process $U^{\alpha*}(t) = V_t(X_t^{x_0,\alpha*}) + \int_0^t L_s(X_s^{x_0,\alpha*}, \alpha*(s))ds$ has the (local) martingale property [55, 17, 15, 24, 58]. We can therefore use the functional Itô formula (3.14) to give a sufficient condition for a functional W to be equal to the value functional V and for a control α^* to be optimal. This necessary condition takes the form of a functional Hamilton-Jacobi-Bellman equation.
The Hamiltonian of the control problem is a functional on $\bigcup_{t\leq T} C_0([0,T],\mathbb{R}^d) \times \mathbb{R}^d \times \mathcal{S}^d$, defined as:

$$H_t(x,\rho,M) = \inf_{u\in A} \frac{1}{2}\mathrm{tr}[M^t\sigma_t(x,u)\sigma_t(x,u)] + \rho b_t(x,u) + L_t(x,u) \tag{3.77}$$

The following theorem is a functional version of the Hamilton-Jacobi-Bellman equation. It links the solution of the optimal control problem to a functional differential equation.

Theorem 3.9 (Verification Theorem, first version)**.** *Let W be a functional in $\mathbb{C}_b^{1,2}([0,T[) \times \mathbb{F}_l^{\infty}([0,T])$, depending on the first argument only. Assume that W solves the functional Hamilton-Jacobi-Bellman equation on $C_0([0,T],\mathbb{R}^d)$, ie for any $x \in C_0([0,T],\mathbb{R}^d)$:*

$$\mathcal{D}_t W_t(x_t) + H_t(x_t, \nabla_x W_t(x_t), \nabla_x^2 W_t(x_t)) = 0 \tag{3.78}$$

as well as the terminal condition:

$$W_T(x) = g(x) \tag{3.79}$$

and the quadratic growth boundedness:

$$W_t(x_t) \leq C \sup_{s\leq t} |x(s)|^2 \tag{3.80}$$

Then, for every $x \in C_0([0,t],\mathbb{R}^d)$ and every admissible control α:

$$W_t(x) \leq J_t(x,\alpha) \tag{3.81}$$

If furthermore for $x \in C_0([0,t],\mathbb{R}^d)$ there exists an admissible control α^* such that:

$$\begin{aligned} H_{t+s}(X_{t+s}^{x,\alpha^*}, \nabla_x W(X_{t+s}^{x,\alpha^*}), \nabla_x^2 W(X_{t+s}^{x,\alpha^*})) = \\ \frac{1}{2} tr[\nabla_x^2 W(X_{t+s}^{x,\alpha^*})^t \sigma_{t+s}(X_{t+s}^{x,\alpha^*}, \alpha^*(s))\sigma_{t+s}(X_{t+s}^{x,\alpha^*}, \alpha^*(s))] \\ +\nabla_x W(X_{t+s}^{x,\alpha^*}) b_{t+s}(X_{t+s}^{x,\alpha^*}, \alpha^*(s)) + L_{t+s}(X_{t+s}^{x,\alpha^*}, \alpha^*(s)) \end{aligned} \tag{3.82}$$

for $0 \leq s < T$, $ds \times d\mathbb{P}$ almost surely, then:

$$W_t(x) = V_t(x) \tag{3.83}$$

Proof. Let α be an admissible control, $t < T$ and $x \in C_0([0,t],\mathbb{R}^d)$, applying functional Itô formula to the functional F defined on $\bigcup_{s \leq T-t} D([0,s],\mathbb{R}^d)$ by:

$$F_s(y) = W_{t+s}((x(u)1_{u<t} + [x(t) + y(u-t)]1_{u \geq t})_{u \leq t+s}) \tag{3.84}$$

yields:

$$\begin{aligned} W_{t+s}(X_{t+s}^{x,\alpha} - W_t(x) &= \int_0^s \nabla_x W_{t+u}(X_{t+u}^{x,\alpha}) dW(u) \\ &+ \int_0^s \mathcal{D}_{t+u} W(X_{t+u}^{x,\alpha}) + \nabla_x W_{t+u}(X_{t+u}^{x,\alpha}) b_{t+u}(X_{t+u}^{x,\alpha}, \alpha(t+u)) du \\ &+ \frac{1}{2} \mathrm{tr}[^t \nabla_x^2 F_u(X_u^{x,\alpha})^t \sigma_u \sigma_u(X_u^{x,\alpha})] du \end{aligned} \tag{3.85}$$

Since W solves the Hamilton-Jacobi-Bellman equation, it implies that:

$$\begin{aligned} W_{t+s}(X_T^{x,\alpha}) - W_t(x) \geq \int_0^s \nabla_x W_{t+u}(X_{t+u}^{x,\alpha}) dW(u) \\ - \int_0^s L_{t+u}(X_u^{x,\alpha}, \alpha(u)) du \end{aligned} \tag{3.86}$$

In other words, $W_{t+s}(X_T^{x,\alpha}) - W_t(x) + \int_0^s L_{t+u}(X_u, \alpha(u)) du$ is a local submartingale. Growth bound (3.80) and L^2-boundedness of the solution of stochastic differential equation (B.32) guarantee that it is actually a true submartingale, hence, taking $s \to T-t$, the left-continuity of W yields:

$$E\left[g(X_T^{x,\alpha}) + \int_0^{T-t} L_{t+u}(X_u^{x,\alpha}, \alpha(u)) du\right] \geq W_t(x) \tag{3.87}$$

This being true for any admissible control α proves that $W_t(x) \leq J_t(x)$.Taking $\alpha = \alpha^*$ transforms inequalities to equalities, submartingale to martingale, hence establishes the second part of the theorem.

□

3.6.3 Optimal control and functional Hamilton-Jacobi-Bellman equation, second version

We suppose now that b,σ are as in corollary B.1, with σ not depending on the control, and that furthermore the path of $t \mapsto \sigma_t(x_t, \int_0^t v(s)ds)$ are right-continuous for all $(x,v) \in C_0([0,t_0],\mathbb{R}^d) \times \mathcal{S}_t$. For any $\alpha \in \mathcal{A}_\mathcal{B}$ and $(x,v) \in C_0([0,t_0],\mathbb{R}^d) \times \mathcal{S}_t$, denote $X^{x,v,\alpha}$ the strong solution of

$$dX(t_0+t) = b_{t_0+t}(X_{t_0+t}, [X]_{t_0+t}, \alpha(t))dt + \sigma_{t_0+t}(X_{t_0+t}, [X]_{t_0+t})dW(t) \tag{3.88}$$

with initial value $\left(x, \left(\int_0^t v(s)ds\right)_{t\leq t_0}\right)$, in the sense of definition B.2. Let $(x_0, 0) \in \mathbb{R}$, viewed as element of $C_0(0,\mathbb{R}^d) \times \mathcal{S}_0$, we consider the control problem:

$$\inf_{\alpha\in\mathcal{A}_\mathcal{B}} E\left[g(X_T^{x_0,0,\alpha}, [X^{(x_0,0,\alpha)}]_T) + \int_0^T L_t(X_t^{x_0,0,\alpha}, [X^{x_0,0,\alpha}]_t, \alpha(t))dt\right] \tag{3.89}$$

where we make the following assumptions on :

1. g is continuous for the sup norm in $\mathcal{U}_T^c \times \mathcal{S}_T$
2. $-K \leq g(x,y) \leq K(1+\sup_{s\in[0,T]}|x(s)|^2 + \sup_{s\in[0,T]}|y(s)|^2)$ for some constant K
3. $-K \leq L_t(x,y,u) \leq K(1+\sup_{s\in[0,T]}|x(s)|^2 + \sup_{s\in[0,T]}|y(s)|^2)$ for some constant K'

The cost functional of the control problem (3.89) is a functional on $\bigcup_{t\leq T} C_0([0,t],\mathbb{R}^d) \times \mathcal{S}_t \times \mathcal{A}_\mathcal{B}$ as:

$$J_t(x,y,\alpha) = E\left[g(X_T^{x,y,\alpha}, [X^{x,y,\alpha}]_T) + \int_t^T L_s(X_s^{x,y,\alpha}, [X^{x,y,\alpha}]_s, \alpha(s))ds\right] \tag{3.90}$$

It is obvious from assumptions 1. to 3. that $J_t(x,v,\alpha)$ is finite for any admissible control α, thanks to the L^2 bound (B.32).

We define the value functional of the problem on $\bigcup_{t\leq T}\mathcal{U}_t^c \times \mathcal{S}_t$ as:

$$V_t(x,v) = \inf_{\alpha\in\mathcal{A}_\mathcal{B}} J_t(x,v,\alpha) \tag{3.91}$$

It is obvious from assumptions 1. to 4. and the L^2 bound (B.32) that $V_t(x,v)$ is finite and satisfies for some constant K":

$$-K'' \leq V_t(x,v) \leq K''(1 + \sup_{s\in[0,t]} |x(s)|^2 + \int_0^t |v(s)|^2) \tag{3.92}$$

It is readily observed that the value process

$$U^\alpha(t) = V_t(X_t^{x_0,\alpha}, [X^{x_0,\alpha}]_t) + \int_0^t L_s(X_s^{x_0,\alpha}, [X^{x_0,\alpha}]_s, \alpha(s))ds$$

has the submartingale property. The martingale approach to stochastic optimal control then characterizes an optimal control $\alpha*$ by the property that the value process $U^{\alpha*}(t) = V_t(X_t^{x_0,\alpha*}, [X^{x_0,\alpha*}]_t) + \int_0^t L_s(X_s^{x_0,\alpha*}, [X^{x_0,\alpha*}]_s, \alpha*(s))ds$ has the (local) martingale property [55, 17, 15, 24, 58]. We can therefore use the functional Itô formula (3.14) to give a sufficient condition for a functional W to be equal to the value functional V and for a control α^* to be optimal. This necessary condition takes the form of a functional Hamilton-Jacobi-Bellman equation.

The Hamiltonian of the control problem is a functional on $\bigcup_{t\leq T} \mathcal{U}_t^c \times \mathcal{S}_t \times \mathbb{R}^d \times \mathcal{S}^d$, defined as:

$$H_t(x,v,\rho,M) = \inf_{u\in A} \frac{1}{2} tr[M^t \sigma_t(x,v)\sigma_t(x,v)] + \rho b_t(x,v,u) + L_t(x,v,u) \tag{3.93}$$

The following theorem is a functional version of the Hamilton-Jacobi-Bellman equation. It links the solution of the optimal control problem to a functional differential equation.

Theorem 3.10 (Verification Theorem, second version)**.** *Let W be a functional in $\mathbb{C}_b^{1,2}([0,T[) \times \mathbb{F}_l^\infty([0,T])$. Assume that W solves the functional Hamilton-Jacobi-Bellman equation on $C_0([0,T],\mathbb{R}^d)$, ie for any $(x,v) \in C_0([0,T],\mathbb{R}^d) \times D([0,T],S_d^+)$:*

$$\mathcal{D}_t W_t(x_t,v_t) + H_t\left(x_t, \left(\int_0^s v(u)du\right)_{s\leq t}, \nabla_x W_t(x_t,v_t), \nabla_x^2 W_t(x_t,v_t)\right) = 0 \tag{3.94}$$

as well as the terminal condition:

$$W_T(x,v) = g\left(x, \left(\int_0^t v(s)ds\right)_{t\leq T}\right) \tag{3.95}$$

Then, for every $x \in C_0([0,t],\mathbb{R}^d)$ and every admissible control α:

$$W_t(x,v) \leq J_t(x,v,\alpha) \tag{3.96}$$

If furthermore for $x \in C_0([0,t],\mathbb{R}^d)$ *there exists an admissible control* α^* *such that, for* $0 \leq s < T$, $ds \times d\mathbb{P}$ *almost surely:*

$$H_{t+s}(X_{t+s}^{x,v,\alpha^*},[X_{t+s}^{x,v,\alpha^*}],\nabla_x W(X_{t+s}^{x,v,\alpha^*},(\sigma_u)_{u\leq t+s}),\nabla_x^2 W(X_{t+s}^{x,v,\alpha^*},(\sigma_u)_{u\leq t+s})) = \frac{1}{2}tr[\nabla_x^2 W(X_{t+s}^{x,v,\alpha^*},(\sigma_u)_{u\leq t+s})^t \sigma_{t+s}\sigma_{t+s}] + \nabla_x W(X_{t+s}^{x,v,\alpha^*},(\sigma_u)_{u\leq t+s}) b_{t+s}(X_{t+s}^{x,v,\alpha^*},[X_{t+s}^{x,v,\alpha^*}],\alpha^*(s)) + L_{t+s}(X_{t+s}^{x,v,\alpha^*},[X_{t+s}^{x,v,\alpha^*}],\alpha^*(s))$$

where $\sigma_u := \sigma_u(X_u^{x,v,\alpha^*},[X_u^{x,v,\alpha^*}]))$*. Then:*

$$W_t(x,v) = V_t(x,v) \tag{3.97}$$

The proof of this theorem goes exactly as the proof of theorem 3.9. The subtle point is that if σ were dependent on the control, it would be impossible to apply the functional Itô formula to $W_s(X_s^{x,v,\alpha},[X^{x,v,\alpha}]_s)$ because $\frac{d[X^{x,v,\alpha}]_s}{ds}$ would not necessarily admit a right-continuous representative for any admissible control α.

Chapter 4

Localization

4.1 Motivation

The regularity assumption needed for functionals in theorems 3.1, 3.2 and 3.7 is strong in the sense that it excludes many interesting examples of functionals representing conditional expectations and satisfying the functional differential equation in theorem 3.7, but yet failing to satisfy its regularity assumptions. The tfollowing two examples show the main issues that we will be addressing by defining a notion of local regularity and proving that some of our main theorems from chapter 3 still hold with this weaker notion.

Example 4.1 (Non-continuous functional). One-dimensional standard Brownian motion W, $b > 0$, $M_t = \sup_{0\leq s\leq t} W(s)$ and the process:

$$Y(t) = E[1_{M_T \geq b}|\mathcal{F}_t] \tag{4.1}$$

This process admits the functional representation $Y(t) = F_t(W(t))$ with F_t defined as:

$$F_t(x_t) = 1_{\sup_{0\leq s\leq t} x(s)\geq b} + 1_{\sup_{0\leq s\leq t} x(s)<b}\left[2 - 2\Phi\left(\frac{b - x(t)}{\sqrt{T-t}}\right)\right] \tag{4.2}$$

where Φ is the cumulative distribution function of the standard Normal random variable. This functional is not even continuous at fixed times because a path x_t where $x(t) < b$ but $\sup_{0\leq s\leq t} x(s) = b$ can be approximated in the sup norm by paths where $\sup_{0\leq s\leq t} x(s) < b$. Also, the path $t \mapsto F_t(x_t)$ is not always continuous at T (take a suitable path that hit b for the first time at time T). However, one can easily check that $\nabla_x F$, $\nabla_x^2 F$ and $\mathcal{D}F$ exist and satisfy the functional differential equation in theorem 3.7 on the set $C_0([0,T],\mathbb{R}^d) \times \mathcal{S}_T$.

Example 4.2 (Exploding derivatives). Take a geometric Brownian motion $S(t) = e^{\sigma W(t) - \frac{\sigma^2}{2} t}$, and times $0 \le s_1 < s_2 < T$

$$Y(t) = E\left[\left(\frac{S(s_2)}{S(s_1)} - K\right)^+ |\mathcal{F}_t\right] \tag{4.3}$$

$BS(x, K, \sigma, r, \delta)$ the Black-Scholes price of a call with underlying price x, strike K, implied volatility σ, interest rate r and time to maturity δ, the process Y admits the functional representation $Y(t) = F_t(S_t)$, where:

$$F_t(x_t) = 1_{t<s_1} BS(1, K, \sigma, 0, s_2 - s_1) + 1_{s_1 \le t < s_2} BS(\frac{x(t)}{x(s_1-)}, K, \sigma, 0, s_2 - t)$$
$$+ 1_{t>s_2} \left(\frac{x(s_2-)}{x(s_1)} - K\right)^+ \tag{4.4}$$

$\nabla_x^2 F$ and $\mathcal{D}F$ fail to be boundedness-preserving since they explode as $t \to s_2-$ on paths where $x(s_2) = Kx(s_1)$, since the Black-Scholes Γ and Θ explode as the underlying price equals the strike and time-to-maturity goes to 0 (see for example [25]). However, the functional does satisfy the functional differential equation in theorem 3.7 on the set $\{x \in C_0([0,T], \mathbb{R}), x(s_2) \neq Kx(s_1)\}$.

In this chapter, we will consider functionals defined on $\Upsilon^c := \bigcup_{t \le T} \mathcal{U}_T^c \times \mathcal{S}_T$, which can be extended to cadlag paths only in a local sense that will be made precise (definition 4.10), allowing us to define their vertical derivatives (definition 4.12). As pointed out by the two examples above, the notion of local regularity is necessary in order for the prices of most usual exotic derivatives to satisfy a valuation functional differential equation (see chapter 5).

4.2 A local version of the functional Itô formula

4.2.1 Spaces of continuous and differentiable functionals on optional intervals

We will introduce in this section spaces of non-anticipative functionals defined on the bundle of cadlag paths Υ.

Definition 4.1. Let $t < T$, define the following equivalence relation on $\mathcal{U}_T \times \mathcal{S}_T$:

$$(x,v)\mathcal{R}_t(x',v') \text{ if and only if } \forall s \leq t, (x(s),v(s)) = (x'(s),v'(s)) \tag{4.5}$$

Denote $\overline{(x_t,v_t)}$ the equivalence class of (x,v) for $\mathcal{R}_t$ which depends only on the restriction (x_t,v_t).

In the following definitions, optional times are defined on the canonical space $\mathcal{U}_T^c \times \mathcal{S}_T$, filtered by the natural filtration of the canonical process $(X,V)((x,v),t) = (x(t),v(t))$. Let $\tau_1 \leq \tau_2$ be two optional times. We introduce the following definitions:

Definition 4.2. Let Ψ be an application from Υ to the set of open subsets of U. Ψ is said to be containing the paths on $[\tau_1,\tau_2[$ if and only if:

$$\forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T, \forall t \in [\tau_1(x,v),\tau_2(x,v)[, x(t) \in \Psi(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}) \tag{4.6}$$

Definition 4.3. Let V be a subset of U and $t_1 < t_2 \leq T$. Define $\Xi_V(t_1,t_2) \subset \mathcal{U}_T \times \mathcal{S}_T$ as the set of paths which are in V between time t_1 and t_2, that is:

$$\Xi_V(t_1,t_2) = \{(x,v) \in \mathcal{U}_T \times \mathcal{S}_T | \forall t \in [t_1,t_2[, x(t) \in V\} \tag{4.7}$$

We will define notions of continuity on the optional interval $[\tau_1,\tau_2[$, which only considers paths coinciding up to time τ_1. Note that contrary to the equivalent spaces in section 2.2.2, the spaces introduced below must be *defined* as subspaces of the space of adapted functionals. The reason is that, as the notion of continuity developed here does not control the dependence of the functional in the paths of its arguments up to time τ_1, it is insufficient to imply measurability.

Definition 4.4 (Adapted functionals)**.** A non-anticipative functional is said to be adapted if and only if the process $(x,v,t) \mapsto F_t(x_t,v_t)$ is adapted to the natural filtration of the canonical process $(x,v,t) \mapsto (x(t),v(t))$ on the space $\mathcal{U}_T \times \mathcal{S}_T$ endowed with the supremum norm and its Borel σ-algebra.

Definition 4.5 (Continuous at fixed times on optional intervals)**.** An adapted functional F defined on Υ_- is said to be continuous at fixed times on the interval $[\tau_1,\tau_2[$ if and only if there exists an application Ψ containing the paths on $[\tau_1, \tau_2[$ such that:

$$\forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T \forall t \in [\tau_1(x,v), \tau_2(x,v)[,$$
$$\forall (x',v') \in \Xi_{\Psi(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}))}(\tau_1(x,v), \tau_2(x,v)) \cap \overline{(x_{\tau_1(x,v)}, v_{\tau_1(x,v)})}, \forall \epsilon > 0, \exists \eta > 0,$$
$$\forall (x'',v'') \in \Xi_{\Psi(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}))}(\tau_1(x,v), \tau_2(x,v)) \cap \overline{(x_{\tau_1(x,v)}, v_{\tau_1(x,v)})},$$
$$d_\infty((x'_t, v'_t), (x''_t, v''_t) < \eta \Rightarrow |F_t(x'_t, v'_y) - F_{t'}(x''_{t'}, v''_{t'})| < \epsilon \quad (4.8)$$

Definition 4.6 (Space of continuous functionals on optional intervals)**.** Define $\mathbb{F}^\infty([\tau_1, \tau_2[)$ as the set of adapted functionals $F = (F_t, t \in [0,T[)$, for which there exists an application Ψ containing the paths on $[\tau_1, \tau_2[$ such that:

$$\forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T \forall t \in [\tau_1(x,v), \tau_2(x,v)[,$$
$$\forall (x',v') \in \Xi_{\Psi(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}))}(\tau_1(x,v), \tau_2(x,v)) \cap \overline{(x_{\tau_1(x,v)}, v_{\tau_1(x,v)})}, \forall \epsilon > 0, \exists \eta > 0,$$
$$\forall t' \in [\tau_1(x,v), \tau_2(x,v)[, \forall (x'',v'') \in \Xi_{\Psi(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}))}(\tau_1(x,v), \tau_2(x,v)) \cap \overline{(x_{\tau_1(x,v)}, v_{\tau_1(x,v)})},$$
$$d_\infty((x'_t, v'_t), (x''_{t'}, v''_{t'}) < \eta \Rightarrow |F_t(x'_t, v'_y) - F_{t'}(x''_{t'}, v''_{t'})| < \epsilon$$

Definition 4.7 (Space of boundedness-preserving functionals on optional interval)**.** Define $\mathbb{B}([\tau_1, \tau_2[)$ as the set of adapted functional F, such that there exists an increasing sequence of stopping times $\sigma_n \geq \tau_1, \lim_n \sigma_n = \tau_2$, a sequence of applications Ψ_n containing the paths on $[\tau_1, \sigma_n[$, such that for every compact subset $K \subset U, R > 0, n \geq 0$ there exists a constant $C_{K,R,n} > 0$ such that:

$$\forall (x,v) \in C_0([0,T], K) \times \mathcal{S}_T \text{ s.t. } \sup_{s \in [0,T]} |v(s)| \leq R$$
$$\forall (x',v') \in \Xi_{\Psi_n(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}))}(\tau_1(x,v), \tau_2(x,v)) \cap \overline{(x_{\tau_1(x,v)}, v_{\tau_1(x,v)})}$$
$$\forall t \in [\tau_1(x,v), \sigma_n(x,v)[, \sup_{s \in [0, \sigma_n(x,v)]} |v'(s)| \leq R \Rightarrow |F_t(x'_t, v'_t)| \leq C_{K,R,n} \quad (4.9)$$

Definition 4.8 (Spaces of differentiable functions on an optional interval)**.** Define $\mathbb{C}_b^{j,k}([\tau_1, \tau_2[)$ as the set of functionals $F \in \mathbb{F}^\infty([\tau_1, \tau_2[)$, predictable in the second variable, for which there exists an application Ψ containing the paths, such that for all $(x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T$, the function F is differentiable j times horizontally and k times vertically at each points (x'_t, v'_t) for $t \in [\tau_1(x,v), \tau_2(x,v)[$ and $(x',v') \in \Xi_{\Psi(x_{\tau_1(x,v)}, v_{\tau_1(x,v)}))}(\tau_1(x,v), \tau_2(x,v))$ in the sense of definitions 2.6, 2.8, such that the horizontal derivatives define functionals continuous at fixed times on $[\tau_1, \tau_2[$ and the vertical derivatives define elements of $\mathbb{F}^\infty([\tau_1, \tau_2[)$.

Remark 4.1. If a functional F belongs to $\mathbb{C}_b^{j,k}([\tau_1,\tau_2[)$, the same application Ψ can be chosen in the definitions of $\mathbb{C}_b^{j,k}([\tau_1,\tau_2[), \mathbb{F}^\infty([\tau_1,\tau_2[)$ for the functional F and its derivatives, the same sequence (σ_n) can be chosen in the definition of $\mathbb{B}([\tau_1,\tau_2[)$ for all derivatives, the same sequence Ψ_n in the definition of $\mathbb{B}([\tau_1,\tau_2[)$ can be chosen for all derivatives, and it can be chosen such that $\Psi_n \subset \Psi$.

The role of the application Ψ in those definitions is to reduce the regularity requirement of the functionals to an open set in which continuous paths evolve on the interval $[\tau_1,\tau_2[$, which can be dependent on the history of the path up to time τ_1. The role of the sequences σ_n and Ψ_n are to allow the derivatives of the functional to explode as time τ_2 is approached.

4.2.2 A local version of functional Itô formula

Theorem 4.1. *Let $\tau_1 \le \sigma \le \tau_2$ be three optional times, such that $\sigma < \tau_2$ on the event $\tau_1 < \tau_2$. If $F \in \mathbb{C}_b^{1,2}([\tau_1,\tau_2[)$, then:*

$$\begin{aligned} F_\sigma(X_\sigma, A_\sigma) - F_{\tau_1}(X_{\tau_1}, A_{\tau_1}) &= \int_{\tau_1}^{\sigma} \mathcal{D}_u F(X_u, A_u) du + \int_{\tau_1}^{\sigma} \nabla_x F_u(X_u, A_u).dX(u) \\ &+ \int_{\tau_1}^{\sigma} \frac{1}{2} \mathrm{tr}[{}^t\nabla_x^2 F_u(X_u, A_u)\ d[X](u)] \qquad a.s. \end{aligned} \tag{4.10}$$

where, with a slight abuse of notation, $\sigma = \sigma(X_T, A_T)$ and $\tau_1 = \tau_1(X_T, A_T)$.

Proof. Let σ_n be as in definition, define a random subdivision of $[\tau_1, \sigma \wedge \sigma_n]$ as follows:

$$\tau_0^n = \tau_1 \qquad \tau_k^n = \inf\{u > \tau_{k-1}^n | 2^n u \in \mathbb{N} \text{ or } |\Delta A(u)| \ge \frac{1}{n}\} \wedge \sigma \wedge \sigma_n \tag{4.11}$$

and the following approximations of the path of X and A:

$$\begin{aligned} {}_nX &= 1_{[0,\tau_1[} X + \sum_{i=0}^{\infty} X(\tau_{i+1}^n) 1_{[\tau_i^n, \tau_{i+1}^n)} + X(\sigma \wedge \sigma_n) 1_{\{\sigma\wedge\sigma_n\}} \\ {}_nA &= 1_{[0,\tau_1[} A + \sum_{i=0}^{\infty} A(\tau_i^n) 1_{[\tau_i^n, \tau_{i+1}^n)} + A(\sigma \wedge \sigma_n) 1_{\{\sigma\wedge\sigma_n\}} \end{aligned} \tag{4.12}$$

which coincides with (X, A) up to time τ_1 , then are pathwise constant between times τ_1 and $\sigma \wedge \sigma_n$.

With this setup, following the proof of theorem 3.1 establishes the formula between times τ_1 and $\sigma_n \wedge \sigma$. The point is that all approximations of the paths of X and A in the proof take values that are actually taken by X and A themselves, and coincide with X and A up to time τ_1, so all of them belong to $\Xi_{\Psi_n(x_{\tau_1},v_{\tau_1})}(\tau_1,\tau_2)\cap\overline{(x_{\tau_1},v_{\tau_1})}$. The observation that $\mathbb{P}\{\exists n,\sigma_n\geq\sigma\}=1$ then concludes the proof. □

4.3 Locally regular functionals

4.3.1 Spaces of locally regular functionals

Definition 4.9 (Non-anticipative functional of continuous paths)**.** A non-anticipative functional on continuous paths is a real-valued functional defined on the vector bundle:

$$\Upsilon_c=\bigcup_{t\in[0,T]} C_0([0,t],U)\times D([0,t],S_d^+) \tag{4.13}$$

We will introduce spaces of functionals defined on Υ_c, which can locally be extended to functionals on Υ in a sense that will be made precise in definition 4.10, which will allow to define their vertical derivatives (definition 4.12).

Definition 4.10 (Space of locally regular functionals)**.** Define the space of locally regular functionals, denoted $\mathcal{R}_{loc}$, as the space of functionals F defined on Υ_c such that there exists an increasing sequence $(\tau_i)_{i\geq0}$ of optional times , satisfying $\tau_0=0,\lim_i\tau_i=T$, and a sequence $(F^i)_{i\geq0}$ of functionals, such that $F^i\in\mathbb{C}_b^{1,2}([\tau_i,\tau_{i+1}[)$, and:

$$\forall(x,v)\in C_0([0,T],\mathbb{R}^d)\times\mathcal{S}_T,\forall t\in[0,T],\forall i\geq0$$
$$F_t(x_t,v_t)1_{t\in[\tau_i(x,v),\tau_{i+1}(x,v)[}=F_t^i(x_t,v_t)1_{t\in[\tau_i(x,v),\tau_{i+1}(x,v)[} \tag{4.14}$$

Definition 4.11 (Space of functionals having continuous path on a set)**.** Let $\mathfrak{D}$ be a Borel subset of the canonical space. A functional F on Υ^c is said to have continuous path on the set $\mathfrak{D}$ if and only if, for all $(x,v)\in\mathfrak{D}$, the path $t\mapsto F_t(x_t,v_t)$ is continuous on $[0,T]$.

Example 4.3. A functional $F\in\mathbb{C}_b^{1,2}([0,T[)\cap\mathbb{F}^\infty([0,T])$, whose vertical derivatives are in $\mathbb{F}^\infty([0,T[)$, is of course in $\mathcal{R}_{loc}$, taking the following constant sequences:

$$\tau_i(x,v)=T,i\geq1;t_i=T,F^i=F,i\geq0,\sigma_n=T-\frac{1}{n} \tag{4.15}$$

It has continuous path on $\mathfrak{D}=\mathcal{U}_T^c\times\mathcal{S}_T$.

Example 4.4. The functional F defined as in example 4.1 belongs to $\mathcal{R}_{loc}$, defining the following sequences:

$$\tau_1(x,v) = \inf\{t \geq 0 | x(t) = b\} \wedge T, \tau_i(x,v) = T, i \geq 2$$
$$F_t^0(x_t, v_t) = 2 - 2\Phi\left(\frac{b - x(t)}{\sqrt{T-t}}\right), F_t^1(x_t, v_t) = 1, i \geq 1 \tag{4.16}$$

The functional F^0 is in $\mathbb{C}_b^{1,2}([\tau_0, \tau_1[)$ taking $\Psi(x_t, v_t) =]-\infty, b[$, $\Psi_n(x_t, v_t) =]-\infty, b[$ and $\sigma_n(x,v) = \tau_1(x,v) \wedge \left(T - \frac{1}{n}\right)$. F has continuous path on the set $\mathfrak{D} = \{(x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T | x(T) \neq b\}$. Note that $\mathbb{P}(W_T \in \mathfrak{D}) = 1$.

Example 4.5. The functional F defined as in example 4.2 belongs to $\mathcal{R}_{loc}$, defining the following sequences:

$$\tau_1(x,v) = s_1, \tau_2(x,v) = s_2, \tau_i(x,v) = T, i \geq 3$$
$$F_t^0(x_t, v_t) = BS(1, K, \sigma, 0, s_2 - s_1)$$
$$F_t^1(x,v) = 1_{t \geq s_1} BS(\frac{x(t)}{x(s_1-)}, K, \sigma, 0, s_2 - t)$$
$$F_t^2(x_t, v_t) = 1_{t \geq s_2}\left(\frac{x(s_2-)}{x(s_1)} - K\right)^+, i \geq 1 \tag{4.17}$$

The functional F^1 is in $\mathbb{C}_b^{1,2}([\tau_1, \tau_2[)$ with the trivial choice $\Psi = U$ and with $\sigma_n = s_2 - \frac{1}{n}$. It has continuous path on the set $\mathcal{U}_T^c \times \mathcal{S}_T$.

The following example shows a case where the application Ψ is non-constant, ie a case in which the open set where the paths on which the functionals F^i are required to be regular, has to be chosen dependent on history.

Example 4.6. Consider a one-dimensional standard Brownian motion W, $b > 0$ take $0 < t_1 < T$, and define for $t \geq t_1$: $M_t = \sup_{t_1 \leq s \leq t}[W(s) - W(t_1)]$ and the process:

$$Y(t) = E[1_{M_T \geq b} | \mathcal{F}_t] \tag{4.18}$$

This process admits the functional representation $Y(t) = F_t(W(t))$ with F_t defined as:

$$F_t(x_t) = 1_{t<t_1}\left[2 - 2\Phi\left(\frac{b}{\sqrt{T-t_1}}\right)\right] +$$
$$1_{t \geq t_1}\left(1_{\sup_{t_1 \leq s \leq t} x(s) - x(t_1) \geq b} + 1_{\sup_{t_1 \leq s \leq t} x(s) - x(t_1) < b}\left[2 - 2\Phi\left(\frac{b + x(t_1) - x(t)}{\sqrt{T-t}}\right)\right]\right) \tag{4.19}$$

It belongs to $\mathcal{R}_{loc}$, defining the following sequences:

$$\tau_1(x,v) = t_1, \tau_2(x,v) = \inf\{t \geq t_1 | x(t) - x(t_1) = b\} \wedge T, \tau_i(x,v) = T, i \geq 3$$
$$F_t^0(x_t,v_t) = 2 - 2\Phi\left(\frac{b}{\sqrt{T-t_1}}\right), F_t^1(x_t,v_t) = 2 - 2\Phi\left(\frac{b + x(t_1) - x(t)}{\sqrt{T-t}}\right),$$
$$F_t^i(x_t,v_t) = 1, i \geq 2 \qquad (4.20)$$

The functional F^0 is in $\mathcal{C}_b^{1,2}([0,\tau_1[)$ taking $\Psi = \Psi_n = \mathbb{R}$ and $\sigma_n(x,v) = t_1 - \frac{1}{n}$.
The functional F^1 is in $\mathcal{C}_b^{1,2}([\tau_1,\tau_2[)$ taking $\Psi(x_{t_1},v_{t_1}) =]-\infty, b + x(t_1)[$, $\Psi_n(x_{t_1},v_{t_1}) =]-\infty, b + x(t_1) - \frac{1}{n}[$ and $\sigma_n(x,v) = \inf\{t \geq 0 | x(t) - x(t_1) \geq b - \frac{1}{n}\}$. F has continuous path on the set $\mathfrak{D} = \{(x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T | x(T) - x(t_1) \neq b\}$. Note that $\mathbb{P}(W_T \in \mathfrak{D}) = 1$.

4.3.2 A local uniqueness result on vertical derivatives

The following theorem are local versions of theorems 2.2 and 2.3. We will only give the proof for theorem 4.3 which needs to be modified in the local case. The adaptation of theorem 2.2 to obtain its local version 4.2 is pretty straightforward and hence will not be given. These theorems will allow us to define the vertical derivatives of functionals defined on Υ_c and satisfying a local regularity assumption (see definition 4.12).

Theorem 4.2. *Let $\tau_1 \leq \tau_2 \leq T$ be optional times defined on the canonical space $\mathcal{U}_T^c \times \mathcal{S}_T$, endowed with the filtration of the canonical process $(X,V)((x,v),t) = (x(t),v(t))$. Assume $F^1, F^2 \in \mathbb{C}_b^{1,1}([\tau_1,\tau_2[)$. If F^1 and F^2 coincide on continuous paths on $[\tau_1,\tau_2[$:*

$$\forall t < T, \quad \forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T,$$
$$F_t^1(x_t,v_t) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t) = F_t^2(x,v) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t)$$
$$\text{then} \qquad \forall t < T, \quad \forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T,$$
$$\nabla_x F_t^1(x_t,v_t) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t) = \nabla_x F_t^2(x_t,v_t) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t)$$

Theorem 4.3. *Let $\tau_1 \leq \tau_2 \leq T$ be two optional times defined on the canonical space $\mathcal{U}_T^c \times \mathcal{S}_T$, endowed with the filtration of the canonical process $(X,V)((x,v),t) = (x(t),v(t))$. Assume $F^1, F^2 \in \mathbb{C}_b^{1,2}([\tau_1,\tau_2[)$. If F^1 and F^2 coincide on continuous paths on $[\tau_1,\tau_2[$:*

$$\forall t < T, \quad \forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T,$$
$$F_t^1(x_t,v_t) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t) = F_t^2(x,v) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t) \qquad (4.21)$$

Then:

$$\forall t < T, \quad \forall (x,v) \in \mathcal{U}_T^c \times \mathcal{S}_T,$$
$$\nabla_x^2 F_t^1(x_t, v_t) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t) = \nabla_x^2 F_t^2(x_t, v_t) 1_{[\tau_1(x,v),\tau_2(x,v)[}(t)$$

Proof. Let $F = F^1 - F^2$. Assume that there exists some $(x,v) \in C_0([0,T],\mathbb{R}^d) \times \mathcal{S}_T$ such that for some $t \in [\tau_1(x,v), \tau_2(x,v))$, and some direction $h \in \mathbb{R}^d$, $\|h\| = 1$, ${}^t h \nabla_x^2 F_t(x_t, v_t).h > 0$, and denote $\alpha = \frac{1}{2} {}^t h \nabla_x^2 F_t(x_t, v_t).h$. We will show that this leads to a contradiction. We already know that $\nabla_x F_t(x_t, v_t) = 0$ by theorem 2.2. Let $\epsilon > 0$ be small enough so that:

$$\begin{aligned} &\forall t' > t, \forall (x', v') \in \mathcal{U}_{t'} \times \mathcal{S}_{t'}, d_\infty((x_t, v_t), (x', v')) < \epsilon \\ &\Rightarrow |F_{t'}(x', v')| < |F_t(x_t, v_t)| + 1, |\nabla_x F_{t'}(x', v')| < 1, \\ &\qquad |\mathcal{D}_t F(x', v')| < C, {}^t h \nabla_x^2 F_{t'}(x', v').h > \alpha \end{aligned} \tag{4.22}$$

where $C > 0$ is some positive constant. Let W be a one dimensional Brownian motion on some probability space $(\Omega, \mathcal{B}, \mathbb{P})$, $(\mathcal{B}_s)$ its natural filtration, and let

$$\tau = \inf\{s > 0 || W(s)| = \frac{\epsilon}{2}\} \tag{4.23}$$

Define the process:

$$Z(s) = F_{t+s}(U_{t+s}, v_{t,s}), \in [0, T-t] \tag{4.24}$$

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

$$U(t') = x(t') 1_{t' \le t} + (x(t) + W((t'-t) \wedge \tau) h) 1_{t' > t} \tag{4.25}$$

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

$$d_\infty((U_{t+s}, v_{t,s}), (x_t, v_t)) < \epsilon \tag{4.26}$$

Note that since $t \in [\tau_1(x,v), \tau_2(x,v)[$ the path $(U, v_{t,T-t})$ coincides with (x,v) on $[0,t]$. Therefore $t \in [\tau_1(U, v_{t,T-t}), \tau_2(U, v_{t,T-t})[$ almost surely since τ_1 and τ_2 are optional times for the canonical filtration. Since τ_2 is an optional time, $\tau_2(x,v) > t$ implies that $\tau_2(x', v') > t$ for any path (x', v') which coincides with (x,v) on $[0,t]$. In particular $\mathbb{P}(\tau_2(U, v_{t,T-t}) > t) = 1$. Let the sequence σ_n be as in definition 4.7 for the derivatives of the functional F, since

$\sigma_n \to \tau_2$, there exists n_0 such that $\mathbb{P}(\sigma_{n_0} > t) > 0$. Define the following optional time on $(\Omega, \mathcal{B}, (\mathcal{B}_s), \mathbb{P})$:

$$\sigma = (\sigma_{n_0}(U, v_{t,T-t}) - t) \wedge \frac{\epsilon}{2} \wedge \tau \tag{4.27}$$

so that we have:

$$\mathbb{P}(\sigma > 0) > 0. \tag{4.28}$$

Define the following random subdivision of $[0, \sigma]$:

$$\upsilon_0^n = 0; \upsilon_i^n = \inf_{s > \upsilon_{i-1}^n 2^n s \in \mathbb{N}} \wedge \sigma \tag{4.29}$$

Define the following sequence of non-adapted piecewise constant approximations of the process W:

$$W^n(s) = \sum_{i=0}^{n\infty} W(\upsilon_{i+1}) 1_{s \in [\upsilon_i, \upsilon_{i+1}[} + W(\sigma) 1_{s \geq \sigma}, 0 \leq s \tag{4.30}$$

Denote:

$$U^n(t') = x(t') 1_{t' \leq t} + (x(t) + W^n((t'-t))h) 1_{t' > t} \qquad Z^n(s) = F_{t+s}(U^n_{t+s}, v_{t,s}) \tag{4.31}$$

Note first that as $n \to \infty$, $Z^n(\sigma-) - Z^n(0)$ converges to $Z(\sigma) - Z(0)$ because $F \in \mathbb{F}^\infty([\tau_1, \tau_2[)$, and that therefore by bounded convergence $E[Z^n(\sigma-) - Z^n(0)]$ converges to $E[Z(\sigma) - Z(0)]$. We then have the following decomposition:

$$\begin{aligned} Z^n(\upsilon_i^n -) - Z^n(\upsilon_{i-1}^n -) &= F_{t+\upsilon_i^n}(U^n_{t+\upsilon_i^n -}, v_{t,\upsilon_i^n}) - F_{t+\upsilon_{i-1}^n}(U^n_{t+\upsilon_{i-1}^n}, v_{t,\upsilon_{i-1}^n}) \\ &+ F_{t+\upsilon_{i-1}^n}(U^n_{t+\upsilon_{i-1}^n}, v_{t,\upsilon_{i-1}^n}) - F_{t+\upsilon_{i-1}^n}(U^n_{t+\upsilon_{i-1}^n -}, v_{t,\upsilon_{i-1}^n}) \end{aligned} \tag{4.32}$$

The first line in the right-hand side of (4.32) can be written: $\psi(\upsilon_i^n - \upsilon_{i-1}^n) - \psi(0)$ where:

$$\psi(h) = F_{t+\upsilon_{i-1}^n + h}(U^n_{t+\upsilon_{i-1}^n, h}, v_{t,\upsilon_{i-1}^n + h}) \tag{4.33}$$

ψ is continuous and right-differentiable with right derivative:

$$\psi'(h) = \mathcal{D}_{t+\upsilon_{i-1}^n + h}(U^n_{t+\upsilon_{i-1}^n, h}, v_{t,\upsilon_{i-1}^n + h}) \tag{4.34}$$

so:

$$\psi(\upsilon_i^n - \upsilon_{i-1}^n) - \psi(0) = \int_{\upsilon_{i-1}^n}^{\upsilon_i^n} \mathcal{D}_{t+u}(U^n_{t+\upsilon_{i-1}^n, u-\upsilon_{i-1}^n}, v_{t,u}) du \tag{4.35}$$

Summing all of the above from $i = 1$ to ∞ (the sum is finite) and taking the limit by dominated convergence theorem, (almost surely U^n converges to U for the supremum norm and $\mathcal{D}F$ is continuous at fixed times on $[\tau_1, \tau_2[$), and the integrand is bounded since $t+\sigma \leq \sigma_{n_0}$, one gets as almost-sure limit as $n \to \infty$:

$$\int_0^\sigma \mathcal{D}_{t+u} F(U_{t+u}, v_{t,u}) du = 0 \tag{4.36}$$

because $\mathcal{D}F$ is 0 on continuous first-argument. By bounded convergence theorem, one finally obtains:

$$E[\sum_{i=0}^{\infty} F_{t+v_i^n}(U^n_{t+v_i^n-}, v_{t,v_i^n}) - F_{t+v_{i-1}^n}(U^n_{t+v_{i-1}^n}, v_{t,v_{i-1}^n})] \to 0 \tag{4.37}$$

as $n \to \infty$.

The second line can be written:

$$\phi(W(v_i) - W(v_{i-1})) - \phi(0) \tag{4.38}$$

where:

$$\phi(u) = F_{t+v_{i-1}^n}(U^{n,u}_{t+v_{i-1}^n-}, v_{t,v_{i-1}^n}) \tag{4.39}$$

so that ϕ is a function parameterized by an $\mathcal{F}_{v_{i-1}^n}$ measurable vector that is almost surely C^2. Applying Itô's formula to ϕ yields:

$$\begin{aligned}\phi(W(v_i) - W(v_{i-1})) - \phi(0) = \int_{v_{i-1}}^{v^i} & \nabla_x F_{t+vi-1}(U^{n,W(u)-W(v_{i-1})}_{t+v_{i-1}^n-}, v_{t,v_{i-1}^n}) h dW(u) \\ & \frac{1}{2} {}^t h \nabla_x F_{t+vi-1}(U^{n,W(u)-W(v_{i-1})}_{t+v_{i-1}^n-}, v_{t,v_{i-1}^n}) h du\end{aligned} \tag{4.40}$$

Summing all of the above from $i = 1$ to ∞ (the sum is finite) yields, denoting $i(s)$ the index such that $s \in [v_i^n, v_{i+1}^n[$:

$$\begin{aligned}\sum_{i=0}^{\infty} F_{t+v_{i-1}^n}(U^n_{t+v_{i-1}^n}, v_{t,v_{i-1}^n}) - F_{t+v_{i-1}^n}(U^n_{t+v_{i-1}^n-}, v_{t,v_{i-1}^n}) = \\ \int_0^\sigma \nabla_x F_{t+v_{i(s)}}(U^{n,W(u)-W(v_{i(s)})}_{t+v_{i(s)}^n-}, v_{t,v_{i(s)}^n}) h dW(u) \\ + \int_0^\sigma {}^t h \nabla_x^2 F_{t+v_{i(s)}}(U^{n,W(u)-W(v_{i(s)})}_{t+v_{i(s)}^n-}, v_{t,v_{i(s)}^n}) h du\end{aligned}$$

Taking the expectation of the above leads to, since the stochastic integral is a true martingale because the integrand is bounded, and because the integrand of the Lebesgue integral is greater or equal to α:

$$E[\sum_{i=0}^{\infty} F_{t+v_{i-1}^n}(U_{t+v_{i-1}^n}^n, v_{t,v_{i-1}^n}) - F_{t+v_{i-1}^n}(U_{t+v_{i-1}^n-}^n, v_{t,v_{i-1}^n})] > \alpha E[\sigma] \tag{4.41}$$

We have therefore established, taking the $\liminf$, that $E[Z(\sigma) - Z(0)] > \alpha E[\sigma]$. But $Z(\sigma) = Z(0) = 0$ almost surely because F is zero on continuous first arguments, and $\mathbb{P}(\sigma > 0) > 0$. A contradiction.

□

4.3.3 Derivatives of a locally regular functional

The following crucial observation, which is an immediate consequence of the definition of optionality, allows to define the derivatives of a locally regular functional.

Lemma 4.1. *Let $(x, v), (x', v') \in C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$, $t < T$ and $\tau_1 \leq \tau_2$ two optional times. Then:*

$$(x_t, v_t) = (x'_t, v'_t) \Rightarrow 1_{t \in [\tau_1(x,v), \tau_2(x,v)[} = 1_{t \in [\tau_1(x,v), \tau_2(x',v')[} \tag{4.42}$$

For all $t < T$, we will therefore define the functional $1_{t \in [\tau_1, \tau_2)}$ on $C_0([0, t], \mathbb{R}^d) \times \mathcal{S}_t$ by its unique value on the pre-image of (x, v) in the set $C_0([0, T], \mathbb{R}^d) \times \mathcal{S}_T$ by the operator of restriction to $[0, t]$.

Theorems 4.2 and 4.3 allow to define the vertical derivatives of a locally regular functional:

Definition 4.12 (Derivatives of a locally regular functional)**.** Define the horizontal and vertical derivatives of a locally regular functional on Υ_{c-} as:

$$\begin{aligned} \mathcal{D}_t F &= \sum_{i \geq 0} D_t F^i 1_{t \in [\tau_i, \tau_{i+1}[} \\ \nabla_x F_t &= \sum_{i \geq 0} \nabla_x F_t^i 1_{t \in [\tau_i, \tau_{i+1}[} \\ \nabla_x^2 F_t &= \sum_{i \geq 0} \nabla_x^2 F_t^i 1_{t \in [\tau_i, \tau_{i+1}[} \end{aligned} \tag{4.43}$$

where the sequence $(\tau_i)_{i\geq 0}$ and $(F^i)_{i\geq 0}$ are as in definition 4.10. These definitions do not depend on the choice of the sequences $(\tau_i)_{i\geq 0}$ and $(F^i)_{i\geq 0}$.

Remark 4.2. Being able to state this definition is the reason why we defined the space of continuous functionals on optional interval $\mathbb{F}^\infty([\tau_1, \tau_2[)$ rather than a space of left-continuous functionals $\mathbb{F}^\infty_l([\tau_1, \tau_2[)$. Had we done that, theorems 4.2 and 4.3 would give coincidence of the vertical derivatives on $]\tau_1, \tau_2[$ instead of $[\tau_1, \tau_2[$, hence the vertical derivatives of a locally regular functional would not have been defined at the countable times $(\tau_i(x, v))_{i\geq 0}$ which depend on the argument (x, v). Martingale representation theorem would however still be true.
Note also that, if the vertical derivatives were defined to belong to a space of right-continuous functionals on optional intervals $\mathbb{F}^\infty_r([\tau_1, \tau_2[)$, we would not have been able to state a local version of the functional Itô formula 4.1 because the subdivision used in the proof is random (see the discussion in the proof of proposition 2.4).

Since vertical derivatives are defined by cadlag perturbations of the path, defining vertical derivatives for functionals on Υ_c requires extending them locally to cadlag paths, and invoking theorems 4.2 and 4.3 to ensure that the definition does not depend on the chosen extension, as is done in definition 4.12. However, since the horizontal extensions of a continuous path are themselves continuous, horizontal derivatives could have been defined directly by equation 2.22. The following lemma ensures that it would be the same notion of horizontal derivative:

Lemma 4.2. *Let $F \in \mathcal{R}_{loc}$, $t < T$ and $(x, v) \in C_0([0, T], \mathbb{R}^d) \times S_T$. Then:*

$$\lim_{h\to 0+} \frac{F_{t+h}(x_{t,h}, v_{t,h}) - F_t(x_t, v_t)}{h} = \mathcal{D}_t F(x_t, v_t) \tag{4.44}$$

Proof. Let i be the index such that $t \in [\tau_i(x, v), \tau_{i+1}(x, v)[$. By lemma 4.1, then also $t \in [\tau_i(x_{t,T-t}, v_{t,T-t}), \tau_{i+1}(x_{t,T-t}, v_{t,T-t})[$, so that for $h > 0$ small enough:

$$t + h < \tau_{i+1}(x_{t,T-t}, v_{t,T-t})$$

. For any such h,

$$F_{t+h}(x_{t,h}, v_{t,h}) = F^i_{t+h}(x_{t,h}, v_{t,h}) \tag{4.45}$$

and hence:

$$\lim_{h\to 0+} \frac{F_{t+h}(x_{t,h}, v_{t,h}) - F_t(x_t, v_t)}{h} = \mathcal{D}_t F^i(x_t, v_t) \tag{4.46}$$

□

4.3.4 Continuity and measurability properties

The two following results on pathwise regularity and measurability for locally continuous functionals are straightforward:

Theorem 4.4 (Pathwise regularity for locally regular functionals)**.** *Let $F \in \mathcal{R}_{loc}$ and $(x, v) \in C_0([0,T], \mathbb{R}^d) \times \mathcal{S}_T$. Then the path $t \mapsto (F_t(x_t, v_t), \nabla_x F_t(x_t, v_t), \nabla_x^2 F_t(x_t, v_t))$ is right-continuous, with a finite number of jump in any interval $[0, t_0]$, $t_0 < T$.*

Proof. Following the proof of theorem 4 for the functionals F^i between times τ_i and τ_{i+1} proves that F and its vertical derivatives are continuous between times τ_i and τ_{i+1}. □

Theorem 4.5 (Measurability properties)**.** *Let $F \in \mathcal{R}_{loc}$. Then: the processes $(F_t(X_t, A_t))$, $(D_t F(X_t, A_t))$, $(\nabla_x F_t(X_t, A_t))$, $(\nabla_x^2 F_t(X_t, A_t))$ are optional.*

Proof. One just has to observe that the indicators $1_{[\tau_i, \tau_{i+1}[}$ are optional, that the processes are adapted by definition, and that their paths are right-continuous. □

4.4 Martingale representation theorem

In this section, we will state the extension of theorem 3.2 to locally regular functionals. The process X is assumed to be a continuous local martingale.

Theorem 4.6. *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 $Y(t)$ admits the functional representation $Y(t) = F_t(X_t, A_t)$ for some $F \in \mathcal{R}_{loc}$, and that F has continuous path on some set $\mathfrak{D}$ such that $\mathbb{P}(X_T \in \mathfrak{D}) = 1$, then:*

$$H = E[H|\mathcal{F}_0] + \int_0^T \nabla_x F_t(X_t, A_t) dX(t) \tag{4.47}$$

Proof. $Y(t) = F_t(X_t, A_t)$ almost surely for some functional $F \in \mathcal{R}_{loc}$. Let the sequences $(\tau_i), (F^i)$ be as in definition 4.10. At a fixed i, let the sequence σ_i^n be as in remark 4.1. Applying the local version of the functional Itô formula 4.1 between times $\tau_i \vee t$ and $\sigma_i^n \wedge t$ yields:

$$\begin{aligned} Y(\sigma_i^n \wedge t) - Y(\tau_i \vee t) = \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} \mathcal{D}_u F(X_u, A_u) du \\ + \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} \mathrm{tr}[{}^t\nabla_x^2 F_u(X_u, A_u)\ d[X](u)] \\ + \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} \nabla_x F_u(X_u, A_u) dX(u) \end{aligned} \tag{4.48}$$

which can be re-written:

$$\begin{aligned} Y(\sigma_i^n \wedge t) - Y(\tau_i \vee t) \quad &= \quad \int_0^t 1_{[\tau_i, \sigma_i^n[}(u) \left[\mathcal{D}_u F(X_u, A_u) du + \mathrm{tr}[{}^t\nabla_x^2 F_u(X_u, A_u)\ d[X](u)] \right] \\ &+ \quad \int_0^t 1_{[\tau_i, \sigma_i^n[}(u) \nabla_x F_u(X_u, A_u) dX(u) \end{aligned} \tag{4.49}$$

By optional sampling theorem, the process $t \mapsto Y(\sigma_i^n \wedge t) - Y(\tau_i \vee t)$ is a martingale, hence by uniqueness of decomposition of a continuous semimartingale in sum of local martingale and continuous variation process, the first line in (4.49) is 0. So we established that:

$$Y(\sigma_i^n \wedge t) - Y(\tau_i \vee t) = \int_0^t 1_{[\tau_i, \sigma_i^n[}(u) \nabla_x F_u(X_u, A_u) dX(u) \tag{4.50}$$

Letting $n \to \infty$, since the process Y is almost surely continuous because $X \in \mathfrak{D}$ with probability 1, we obtain:

$$Y(\tau_{i+1} \wedge t) - Y(\tau_i \wedge t) = \int_0^t 1_{[\tau_i, \tau_{i+1}[}(u) \nabla_x F_u(X_u, A_u) dX(u) \tag{4.51}$$

Summing the above equality on all indices i yields:

$$Y(t) - Y(0) = \int_0^t \nabla_x F_u(X_u, A_u) dX(u) \tag{4.52}$$

Taking the limit $t \to T$ finishes to establish the theorem.

□

4.5 Functional equation for conditional expectations

Consider now as in section 3.5 a continuous semimartingale X satisfying a stochastic differential equation with functional coefficients:

$$dX(t) = b_t(X_t, A_t)dt + \sigma_t(X_t, A_t)dW(t) \tag{4.53}$$

where b, σ are non-anticipative functionals on Υ with values in $\mathbb{R}^d$-valued (resp. $\mathbb{R}^{d\times n}$), and the coordinates of σ are in $\mathbb{F}_r^\infty$. We can state the locally regular version of theorem 3.7. A major difference arises here as this theorem only provides with a *sufficient* condition on a locally regular functional to define a local martingale, rather than a *necessary and sufficient* condition in the case of $\mathbb{C}_b^{1,2}([0,T[) \cup \mathbb{F}_l^\infty([0,T])$ functionals. The reason is that locally regular functionals do not necessarily have regularity on Υ_c endowed with natural topologies, hence a natural support can not be identified for the functional differential equation. For the same reason, uniqueness theorem 3.8 does not have a version for locally regular functionals.

Theorem 4.7 (Functional differential equation for locally regular functionals)**.** *If $F \in \mathcal{R}_{loc}$ satisfies for $t \in [0,T[$:*

$$D_tF(x_t, v_t) + b_t(x_t, v_t)\nabla_x F_t(x_t, v_t) + \frac{1}{2}tr[\nabla_x^2 F(x_t, v_t)\sigma_t{}^t\sigma_t(x_t, v_t)] = 0, \tag{4.54}$$

on the set:

$$\{(x, v) \in C_0([0,T], \mathbb{R}^d) \times D([0,T], S_d^+), v(t) = \sigma_t{}^t\sigma_t(x_t, v_t)\} \tag{4.55}$$

and if F has continuous path on some set $\mathfrak{D}$ such that $\mathbb{P}(X_T \in \mathfrak{D}) = 1$ then the process $Y(t) = F_t(X_t, A_t)$ is a local martingale.

Proof. Define the process $Y(t) = F_t(X_t, A_t)$. Let the sequences $(\tau_i), (F^i)$ be as in definition 4.10. At a fixed i, let the sequence σ_i^n be as in remark 4.1. Applying the local version of the functional Itô formula 4.1 between times $\tau_i \vee t$ and $\sigma_i^n \wedge t$ yields:

$$Y(\sigma_i^n \wedge t) - Y(\tau_i \vee t) = \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} \mathcal{D}_u F(X_u, A_u)du + \frac{1}{2}tr[\nabla_x^2 F_u(X_u, A_u)\sigma_u{}^t\sigma_u(X_u, A_u)]du$$
$$+ \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} b_u(X_u, A_u)\nabla_x F_u(X_u, A_u)du + \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} \nabla_x F_u(X_u, A_u)\sigma_u(X_u, A_u)dW(u)$$

where $A(u) = {\sigma_u}^t \sigma_u(X_u, A_u)$ almost surely, hence taking into account the functional differential equation leads to:

$$Y(\sigma_i^n \wedge t) - Y(\tau_i \vee t) = \int_{\tau_i \vee t}^{\sigma_i^n \wedge t} \nabla_x F_u(X_u, A_u) \sigma_u(X_u, A_u) dW(u) \tag{4.56}$$

Since the process Y is continuous because $X \in \mathfrak{D}$ with probability 1, one can let $n \to \infty$ and obtain:

$$Y(\tau_{i+1} \wedge t) - Y(\tau_i \vee t) = \int_{\tau_i \vee t}^{\tau_{i+1} \wedge t} \nabla_x F_u(X_u, A_u) \sigma_u(X_u, A_u) dW(u) \tag{4.57}$$

Summing over all $i \geq 0$ (since the sum is actually finite) yields:

$$Y(t) - Y(0) = \int_0^t \nabla_x F_u(X_u, A_u) \sigma_u(X_u, A_u) dW(u) \tag{4.58}$$

which proves that Y is a local martingale.

□

Chapter 5

Sensitivity analysis of path-dependent derivatives

5.1 Motivation

In this chapter we use the formalism of functional Itô calculus to study the pricing and hedging of derivative securities. The tools developed in Chapters 2-4 allow to extend the sensitivity analysis which is traditionally developed for non-path-dependent options in Markovian models to settings where the payoff and/or the volatility process are allowed to be path-dependent.

In financial markets, the value of a derivative security depends on market history and may be viewed as a functional of the path of the underlying financial assets from the contract inception date to the current date. This functional is parameterized by relevant market data at the current date, such as the interest rate curve and the prices of traded derivatives on the underlying assets. The notion of vertical derivative of the functional is therefore a natural expression of the sensitivity of the derivative with respect to the prices of underlying assets, while its horizontal derivative is the sensitivity with respect to the passage of time. Sensitivity to the passage of time and first and second order sensitivities to the underlying prices are known by derivatives traders to satisfy a relationship known as Theta - Gamma $(\Theta - \Gamma)$ tradeoff, which can be formalized through a functional differential equation. Derivatives of the functional with respect to the parameters are the sensitivities of the derivatives with respect to observable market variables such as the implied volatility surfaces or the interest rate curves.

We will start this chapter by an introduction to derivatives securities pricing and hedging. The first part will be a short introduction to the usual no-arbitrage pricing of derivatives securities in Quantitative Finance, which is centered around the concept of *replication portfolio*. The second part will consider pricing and hedging from the point of view of a derivatives trader, which is centered around the concept of *sensitivity*, as documented in the few reference books on options from a *trading* point of views [13, 62, 2]. We will then proceed to state a valuation equation (theorem 5.1), similar to the one appearing in Dupire [23]. This equation reconciles the theoretical point of view with the trader's one, as it shows that the replication portfolio corresponds to the hedging of directional sensitivity. Theorem 5.2 then gives rigorous meaning to the $\Theta - \Gamma$ tradeoff. We then use this expression of $\Theta - \Gamma$ tradeoff to link the second order price sensitivity of the derivative to its implied volatility sensitivity in section 5.4, where we will be able to define the sensitivity of a path-dependent derivative to observable market variables, in particuler its "bucket Vegas" ([62], chapter 9). We are then able to define the Black-Scholes Delta and the Delta at a given skew stickiness ratio, which are the actual Deltas that are used on the markets to trade derivative portfolios; and we conclude the section by proposing an efficient numerical algorithm to compute the sensitivites to the market variables and the Deltas of a derivative.

5.1.1 A short introduction to no-arbitrage pricing of derivatives securities

In this chapter, we shall introduce the usual setting of mathematical modeling of portfolios and options in quantitative finance literature. A more detailed discussions with less restrictive assumptions can be found in [42]. We consider a world of d tradable assets, modeled by an $\mathbb{R}^d$-valued process S defined on a filtered probability space $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$, satisfying the relation:

$$\frac{dS(t)}{S(t)} = \mu(t)dt + \nu(t)dB(t) \tag{5.1}$$

where $\frac{dS(t)}{S(t)} = (\frac{dS^1(t)}{S^1(t)}, \dots, \frac{dS^d(t)}{S^d(t)})$, for some bounded progressively measurable processes μ and σ, and where B is an n-dimensional Brownian motion on $(\Omega, \mathcal{B}, \mathcal{B}_t, \mathbb{P})$. We will always make the additional assumption that the eigenvalues of ${}^t\nu(t)\nu(t)$ are bounded from below by some $\epsilon > 0$. We assume that there is also an short-term rate of lending-borrowing, which is represented by a bounded progressively measurable process r. A portfolio is represented by an initial value $V_0 \in \mathbb{R}$ and a progressively measurable $\mathbb{R}^d$-valued process δ, where $\delta^i(t)$ represents the number of shares of asset i held at time t, and is such that the following portfolio value process is well-defined:

$$V(t) = \int_0^t (V(s) - \delta(s)S(s))r(s)ds + \int_0^t \delta(s)dS(s) \tag{5.2}$$

This process represents the value of the portfolio at time t. The heuristic meaning of this equation is that the value of the tradable assets in the portfolio at time s is $\delta(s)S(s)$, hence the cash balance is $V(s) - \delta(s)S(s)$. If the cash balance is positive, it can be lent short-term and hence grows at the short-term interest rate $r(s)$, while if it is negative, the portfolio has to be financed at the short-term borrowing rate and hence the negative balance grows at rate $r(s)$, hence the first integral represents the aggregation of interest incomes on positive cash balances and interest expenses on negative cash balances. On other hand, the change in value of the holding in tradable assets between times s and $s+ds$ is $\delta(s)(S(s+ds) - S(s))$, hence the stochastic integral representing the change of value of the portfolio due to the fluctuation of prices of the tradable assets. A rigorous formulation of this heuristic, starting from portfolios with piecewise constant holdings then considering the limit of continuous rebalancing, can be found in [42] or [25].

Our assumptions on μ, σ and r guarantee (see [42]) that there exists a bounded progressively measurable $\mathbb{R}^n$-valued process λ such that $\mu(t) - r(t)\mathbf{1} = \nu(t)\lambda(t)$, where $\mathbf{1}$ is the vector of $\mathbb{R}^d$ whose coordinates are all 1, hence switching to the probability $\mathbb{Q}$, which is defined on the σ-algebra $\mathcal{B}_t$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\int_0^t \lambda(s)dB(s) - \frac{1}{2}\int_0^t |\lambda(s)|^2 ds} \tag{5.3}$$

makes the process $\left(e^{-\int_0^t r(s)ds} S(t)\right)_{t\geq 0}$ a $\mathbb{Q}$-martingale, and consequently, for any portfolio δ with value process V, the process $\left(e^{-\int_0^t r(s)ds} V(t)\right)_{t\geq 0}$ is a local martingale. A portfolio (V_0, δ) will be said admissible if this local martingale is a true martingale. $\mathbb{Q}$ is called a risk-neutral probability, because under this probability risky portfolios have the same instantaneous return than cash, hence there is no reward for bearing risk. The financial securities satisfy the dynamics:

$$dS(t) = r(t)dt + \nu(t)dW(t) \tag{5.4}$$

where $W(t) = B(t) + \int_0^t \lambda(s)ds$ is a Brownian motion under $\mathbb{Q}$. We will denote throughout this chapter $E[.]$ the expectation under the risk-neutral probability $\mathbb{Q}$.

On financial markets, an option with maturity T is a contract between two counterparties, guaranteeing a settlement between the counterparts which depend on publicly available market data. Some options are over the counter, meaning that the definition of the settlement payment between the counterparts is specified in a written contract between both of them and called "term sheet", or exchange traded, meaning that it has standardized terms defined by an exchange, and traders with exchange membership take either side of the contract without explicitly knowing their counterpart, as orders are matched by the exchange. The natural modeling for an option is therefore a functional g defined on $C_0([0,T],\mathbb{R}^d)$, representing the payment from counterpart 1 to counterpart 2 as a function of the observed paths of tradable securities between inception of the option at time 0 and its maturity at time T. The option is said to be replicable, if there exists an admissible portfolio (V_0,δ) such that $V(T) = g(S_T)$ almost surely. In case the option is replicable, the price of the option g at time t is defined to be $V(t)$ because holding the option or holding the portfolio until time T are equivalent. Since $\left(V(t)e^{-\int_0^t r(s)ds}\right)_{t\geq 0}$ is a martingale, then the price at time t of a replicable option g is:

$$V(t) = E\left[g(S_T)e^{-\int_t^T r(s)ds}|\mathcal{B}_t\right] \tag{5.5}$$

A model is said to be complete if every option g with linear growth with respect to the supremum norm, is replicable.

A common procedure used by market practitioners to price options, wether they are replicable or not, is to assume a model such as (5.4), and then define for price $P(t)$ of the option g at time t as:

$$P(t) = E[g(S_T)e^{-\int_t^T r(s)ds}|\mathcal{B}_t] \tag{5.6}$$

This procedure is called risk neutral-pricing. It guarantees that the price of the options is what it should be for options that are replicable on one hand, and on the other hand that the set of option prices it generates is arbitrage-free, meaning that there is no admissible portfolio combining the financial securities S and a set of available options $(g_1, ..., g_N)$ with price processes $(P_1, ..., P_N)$ defined by (5.6), which can generate a positive profit with nonzero probability while having 0 probability of generating a loss (see [42]). Such a property is required for the internal option valuation system of an investment bank, otherwise its traders themselves could arbitrage the system by being credited for taking positions whose final value can only be positive. It also satisfies the important property of being linear, which required within a derivatives house because a derivative must be marked at the same price by an individual trader, a trading desk, the firm as a whole, and independent market makers representing the firm on the exchange floor, although all of them have different aggregate positions.

5.1.2 A short introduction to derivatives pricing and hedging: a sell-side trader's point of view

From a sell-side trading point of view, the price of a derivative is some (deterministic) function of the underlying assets paths from the inception of the derivative contract to the current date, and some observable market variable at the current date, such as interest rate and credit curves, and the prices on other derivative contracts on the same instrument which are liquidly traded on exchanges such as the Chicago Board of Exchange for US markets. Hence the right formalism for the price of a derivative should be a functional of the path of the underlyings up to current time, parameterized by the vector of observable market variables at the current time which are relevant for the pricing of this derivative contract. In practice, this pricing functional is defined as the expectation of the final payoff, given that the underlying will follow from current time to expiry some discrete time, diffusion or jump diffusion model, parameterized by some vector of parameters which are functions of the observable market prices at the current time, typically chosen in order to maximize the fit of the model to the prices of exchange-traded derivatives contract. An example would be to choose as market observable the price of one call or put and choose the volatility parameter of a Black-Scholes model [8], and industry standard in equity markets is to fit a Dupire local volatility model [22] or an implied tree [20] to the traded strikes and expiries for the underlying. The existence of a dynamic replication in the model or not (complete or incomplete market) is of little relevance to the trader since the role of the model is only to give a price to the derivative contract at a given time; at any future revaluation time the model will actually be different since its parameters will have changed. Decision making and hedging is done through sensitivity analysis, where the sensitivities are first and upper derivatives of the price functional with respect to the current underlying price, passage of time and observable market variables. In particular, sensitivity with respect to the underlying price is computed by "bumping" the current price of the underlying by a small value, keeping history up to current date constant, hence corresponds to the notion of vertical derivative introduced in the formalism of functional Itô calculus in definition 2.8. Many notions of sensitivities to the underlying prices can actually be defined, depending on what observable market variables are expected to change when the underlying price is "bumped" (see section 5.4.3), one possibility being a "constant model" assumption, i.e. that they move so that model parameters are unchanged. A derivatives position is summarized by its different sensitivities, and traders choose their exposure by combining instrument in order to achieve the sensitivities they want in different scenarii. Hedging consists of taking positions in order to reduce or annul the sensitivities to the market variables the trader does not want exposure to.

5.2 Functional valuation equation and greeks for exotic derivative

5.2.1 Valuation equation

Within all chapter 5, the open set U is taken to be $U =]0,\infty[^d$.

We assume that the pricing model is a *functional volatility model*, i.e. the assets price process in the model follows a diffusion taking the form:

$$\frac{dS(t)}{S(t)} = \mu_t dt + \sigma_t(S_t, [S]_t) dB(t) \tag{5.7}$$

where μ is a bounded progressively measurable process and σ is a bounded $\mathcal{M}^{d,n}$-valued non-anticipative functional, and that the short-term lending and borrowing rate takes the form $r_t(S_t, [S]_t)$, where r is a bounded functional. Hence, under the risk-neutral measure, S is a weak solution of the stochastic differential equation with functional coefficients:

$$\frac{dS(t)}{S(t)} = r_t(S_t, [S]_t)dt + \sigma_t(S_t, [S]_t)dW(t) \tag{5.8}$$

Denote $a_t = {}^t\sigma_t\sigma_t$. We shall make the further assumption that the eigenvalues of a are bounded away from 0 and that the process $a_t(S_t, [S_t])$ has cadlag trajectories, so that the process $A(t)$ defined by

$$A_{i,j}(t) = a_t^{i,j}(S_t, [S]_t)S_i(t)S_j(t), 1 \leq i, j \leq d$$

is the cadlag representative of $\frac{d[S](t)}{dt}$.

Remark 5.1. In the case where the interest rate r and the functional a do not depend on the second argument (quadratic variation), the assumption that $a_t(S_t, [S_t])$ has cadlag trajectories can be removed, and all results presented in this chapter hold with functionals F not depending on the second argument v.

The following theorem links the price of options to the solution of a functional differential equation. It was stated by Dupire [23] for functionals depending on the first argument only and with regular functionals ($\mathbb{C}_b^{1,2}$ with $\mathbb{F}^\infty$ derivatives) with no dependence in the second argument) and is the generalization of the partial differential equation for pricing in a local volatility model [22], which itself generalizes the original Black-Scholes partial differential equation [8]. The extension to functionals with dependence in A allows to treat the case of functional local volatility depending on the realized quadratic variance-covariance as in section 5.3.6; more importantly, the extension to *local regularity* allows to price almost all real-life derivatives in a local volatility model as solutions of the valuation equation, since most of them fail to satisfy the regularity assumptions in [23].

This theorem is mathematically the expression of theorem 4.7 in the context of option pricing. However, it differs from an economic interpretation since the argument is based on constructing a portfolio that replicates the option rather than computing a conditional expectation; the link between both is that the price of the replication portfolio is its risk-neutral conditional expectation. The important point is that the drift μ of the process under the physical probability $\mathbb{P}$ does not matter for the pricing of options; in the functional differential equation, the instantaneous short-term lending/borrowing rate $r_t(S_t, [S]_t])$ appears as the coefficient of the first derivative of the function. This theorem links the trader's *sensitivity* to the underlying prices to Mathematical Finance's *dynamic replication*, since the replication portfolio consists of $\nabla_x F_t(S_t, A_t)$ shares.

Theorem 5.1. *Let F be a locally regular functional and g be an option. Assume that, for all $t < T$, F satisfies:*

$$\mathcal{D}_t F_t(x_t, v_t) + r(x_t, (\int_0^s v(u)du)_{s\leq t}) \sum_{i=1}^{d} \partial_i F_t(x_t, v_t) x^i(t)$$
$$+\frac{1}{2} \sum_{1\leq i,j\leq d} \partial_{ij} F_t(x_t, v_t) x^i(t) x^j(t) a_t^{ij}(x_t, (\int_0^s v(u)du)_{s\leq t}) = 0 \tag{5.9}$$

with terminal condition:

$$F_T(x_T, v_T) = g(x_T) \tag{5.10}$$

on the following subset of $\mathcal{U}_T^c \times \mathcal{S}_T$:

$$\{(x, v) \in \mathcal{U}_T^c \times \mathcal{S}_T | \forall t \leq T v_{ij}(t) = x_i(t) x_j(t) a_t^{ij}(x_t, (\int_0^t v(s)ds)_{s\leq t})\} \tag{5.11}$$

together with the integrability condition:

$$E[\sup_{t\in[0,T]} |F_t(S_t, A_t)|] < +\infty \tag{5.12}$$

and if F has continuous path on some set $\mathfrak{D}$ such that $\mathbb{P}(S_T \in \mathcal{D}) = 1$, then option g is replicable by the portfolio with initial value $F_0(S_0, a_0(S_0, 0))$ and position $\nabla_x F_t(S_t, a_t(S_t, [S]_t))$, and its price at time $t \leq T$ is $F_t(S_t, a_t(S_t, [S]_t))$.
If the context of this theorem, we will say that the functional F prices the option g.

We will give here the elementary proof of the theorem with the further assumption $F \in \mathbb{C}_b^{1,2}([0, T[) \cap \mathbb{F}_l^{\infty}([0, T])$ in order to express clearly the idea without getting into the technicalities of working with local regularity. A full proof goes applying Itô's formula locally along the lines of the proof of theorem 4.7.

Proof. Applying Itô's formula 3.1 to $F_t(S_t, a_t(S_t, [S]_t))$ and taking the functional differential equation into account yields:

$$dF_t(S_t, a_t(S_t, [S]_t) = [F_t(S_t, a_t(S_t, [S]_t)) - \nabla_x F_t(S_t, a_t(S_t, [S]_t))S(t)]\, r_t(S_t, a_t(S_t, [S]_t)dt \\ +\nabla_x F_t(S_t, a_t(S_t, [S]_t))dS(t)$$

which proves that $F_t(S_t, a_t(S_t, [S]_t))$ is the price process of a portfolio with initial value $F_0(S_0, a_0(S_0, [S]_0))$. Moreover, terminal condition ensures that it coincides with the payoff of the option g at expiry T, and integrability condition ensures that it is an admissible portfolio. Hence it replicates the option and its price process coincides with the price process of the option. □

5.2.2 Delta, gamma and theta

Let g be an option priced by a functional F as in theorem 5.1. The price of the option at time t within the model, $F_t(S_t, A_t)$, is therefore a *deterministic* functional of the path of (S, A) up to time t. Remarking that $F_t(S_t, A_t) = F_t(S_t, A_{t-})$ since F is predictable in the second variable, it can be furthermore seen as a deterministic functional of the current time t and the two observations:

1. The path in the past, which is (S_{t-}, A_{t-}) and models the fixings already determined and the barrier events already triggered, which is fixed and can not be moved.

2. The current prices of assets $S(t)$, of which one can consider *perturbations* in order to perform sensitivity analysis

As functional F can be locally extended to a functional on cadlag paths, one can therefore defines the sensitivities or *greeks* of the option:

Definition 5.1 (Sensitivities of an option)**.** Define the model Delta or Δ of the option g as the $\mathbb{R}^d$-valued process:

$$\Delta_t = \nabla_x F_t(S_t, A_t) \tag{5.13}$$

Δ is the sensitivity of the option to a perturbation of the current prices of the underlyings. Define the Gamma or Γ of the option g as the $\mathcal{M}^{d,d}$-valued process:

$$\Gamma_t = \nabla_x^2 F_t(S_t, A_t) \tag{5.14}$$

Γ is the sensitivity of the Δ to a perturbation of the current prices of the underlyings. Define the Theta or Θ of the option as:

$$\Theta_t = \mathcal{D}_t F(S_t, A_t) \tag{5.15}$$

Θ is the sensitivity of the option to the passage of time.

Remark 5.2. The right notion of Θ as it appears from the partial differential equation does not always correspond to the decay of the price in a "frozen" market. It is a derivative where S is assumed to remain constant but the market is still assumed to realize volatility since the derivative A of the realized quadratic variance-covariance is also assumed to remain constant. A real "frozen market" decay would be defined as:

$$\Theta_{0vol} = \mathcal{D}_t F(S_t, A_t^{-A(t)}) \tag{5.16}$$

where the instantaneous variance-covariance is also bumped to zero. In cases where the interest rate and the functional volatility σ only depend on the first variable, both notions of Θ are the same, but in the case where they have explicit dependence in the realized variance-covariance the good notion of Θ actually differs from the zero-volatility Θ, as shown in section 5.3.6. The real interpretation of Θ for the trader in that case is that at the next revaluation time, the market will trade at the same point where it trades now, but in between it will still have realized the instantaneous variance-covariance $A(t) = a_t(S_t, [S]_t)$ assumed by the model.

As pointed out by Dupire in [23], the formalism of functionals of the path together with the notions of vertical and horizontal derivatives allow to give a meaning to the greeks for exotic derivatives, in a way that is coherent with practitioners' understanding. Moreover, theorem 5.1 underlies that the *sensitivity* hedge of the Δ is indeed the portfolio making the *dynamic* hedge of the option within the model (5.7), hence it allows reconciliation between the practitioners' understanding of Δ as a sensitivity and its traditional view in mathematical finance literature as the integrand in a martingale representation theorem. Before [23], this reconciliation have only appeared in literature in the case where the option price is a classic *function* of the underlyings' current price and the current time.

5.2.3 $\Theta - \Gamma$ tradeoff

Having defined the greeks of a derivative contract, one can note that the valuation functional equation can be rewritten (omitting the variables of the functionals):

$$\frac{1}{2} \sum_{1\le i,j\le d} a_{i,j} x_i x_j \Gamma_{i,j} = r(F - \sum_{1\le i\le d} \Delta_i x_i) - \Theta \tag{5.17}$$

This way of writing the valuation functional euqation is actually well-known for vanilla options in Black-Scholes and local volatility model (see for example [2], [13]) as an expression of the $\Theta - \Gamma$ tradeoff. It is also well-known by traders, but not referenced in mathematical finance litterature before [23], that this also should hold for positions in path-dependent derivatives; theorem 5.1 gives precise meaning to this well-known fact. The heuristics is that, as the underlying prices moves from S to $S + \delta S$ and time from t to $t + \delta t$, the price of the derivative contract should move by (this is formally Taylor expansion at order 2 in S and order 1 in t) :

$$\sum_{1\le i\le d} \Delta_i \delta S_i + \frac{1}{2} \sum_{1\le i,j\le d} a_{i,j} \Gamma_{i,j} \delta S_i \delta S_j + \Theta \delta t \tag{5.18}$$

A long derivative delta-neutral's trader position would annul the first-order variation by taking a position $-\Delta$ in the underlyings. His overall position is long a derivative contract and short Δ in the underlyings hence it cost him $r(F - \sum_{1\le i\le d} \Delta_i S_i)\delta t$ to finance it at the short-term rate. His total loss from passage of time is therefore $[r(F - \sum_{1\le i\le d} \Delta_i x_i) - \Theta]\delta t$ while his gain from the underlyings moving is $\frac{1}{2}\sum_{1\le i,j\le d} a_{i,j}\Gamma_{i,j}\delta S_i \delta S_j$. The $\Gamma - \Theta$ tradeoff formula tells that it breaks even when $E[\delta S_i \delta S_j] = a_{i,j} S_i S_j \delta t$, that is if the instantaneous variance-covariance structure assumed in the pricing model corresponds to the realized one. The following theorem is the expression of El Karoui's Black-Scholes robustness formula [26] in the context of a functional volatily model, and is the dynamic expression of $\Theta - \Gamma$ tradeoff as it characterizes the break-even volatility for a trader delta-hedging a derivative at constant model. Its financial meaning has many implications in option pricing and actual volatility trading [31]. It has been stated by Dupire [23] for regular functionals with no dependence in quadratic variation.

Theorem 5.2 (Θ - Γ tradeoff)**.** *Assume that the stock price process is a diffusion with* *bounded coefficients:*

$$\frac{dS(t)}{S(t)} = \mu(t)dt + \nu(t)dB(t) \tag{5.19}$$

and denote $\xi(t) = {}^t\nu(t)\nu(t)$*. Let* g *be an option, and* F *a solution of the valuation functional w' as in theorem 5.1 for a functional volatility model* $\sigma_t(.,.)$*, such that* $\nabla_x F$ *is bounded and*

$$E[|\mathbf{tr}[\nabla_x^2 F_t(S_t, A_t)^t S(t)S(t)]|] \leq h(t) \tag{5.20}$$

for some positive measurable function $h : [0, T[\mapsto \mathbb{R}, \int_0^T h(t)dt < \infty$*. Then the final value of the admissible porfolio* $(F_0(S_0, A_0), \Delta_t(S_t, A_t))$ *is:*

$$g(S_T) + \int_0^T \frac{1}{2} \sum_{1\leq i,j\leq d} (a_t^{i,j}(S_t, [S]_t) - \xi^{i,j}(t))S_i(t)S_j(t)e^{\int_t^T r(S_s,[S]_s)ds}\Gamma_t^{i,j}(S_t, A_t)dt \tag{5.21}$$

This theorem is actually the expression of $\Theta - \Gamma$ tradeoff since it expresses that the infinitesimal gain between time t and $t+dt$ of the strategy consisting of holding the derivative contract and hedging it by being short its replication portfolio according to the functional local volatility model $\sigma_t(.,.)$ is $(a^{i,j}(t) - a_t^{i,j}(S_t, A_t))S_i(t)S_j(t)\Gamma_t^{i,j}(S_t, A_t)dt$, that is the Γ-weighted difference between the actual realized variance-covariance realized by the assets and the one assumed by the pricing model.
We will give the proof here for a functional that is furthermore in $\mathbb{C}_b^{1,2}([0,T[) \cap \mathbb{F}_l^\infty([0,T])$; a proof for the functional being only $\mathcal{R}_{loc}$ would go applying Itô's formula between stopping times and going to the limit along the lines of the proof of theorem 4.7.

Proof. Define the process:

$$V(t) = F_t(S_t, A_t) + \int_0^t \frac{1}{2} \sum_{1\leq i,j\leq d} (a_s^{i,j}(S_s, [S]_s) - a^{i,j}(s))S_i(s)S_j(s)e^{\int_s^t r(S_u,[S]_u)du}\Gamma_s^{i,j}(S_s, A_s)ds \tag{5.22}$$

Applying the functional Itô formula yields:

$$dV(t) = \mathcal{D}_t F(S_t, A_t)dt + r_t(S_t, [S]_t)\left[\int_0^t \frac{1}{2} \sum_{1\leq i,j\leq d} (a_s^{i,j}(S_s, [S]_s) - a^{i,j}(s))S_i(s)S_j(s)e^{\int_s^t r(S_u,[S]_u)du}\Gamma_s^{i,j}(S_s, A_s)ds\right]dt + \Delta_t(S_t, A_t)dS(t) + \frac{1}{2} \sum_{1\leq i,j\leq d} a_t^{i,j}(S_t, [S]_t)S_i(t)S_j(t)\Gamma_t^{i,j}(S_t, A_t)dt$$

Taking into account the functional differential equation satisfied by F yields:

$$dV(t) = r_t(S_t, [S]_t)(V(t) - \Delta_t(S_t, A_t)S(t))dt + \Delta_t(S_t, A_t)dS(t) \tag{5.23}$$

which proves that $V(t)$ is the price process of the portfolio $(V_0 = F_0(S_0, A_0), \Delta_t(S_t, A_t))$. The integrability condition (5.20) and F having continuous path on a set of full measure allow to pass to the limit $t \to T$. The Δ being bounded ensures that the portfolio is admissible. □

5.3 Examples of the valuation equation

We will give here examples of the valuation functional differential equation applied to some commonly encountered options priced in the standard Dupire local volatility model [22], where the volatility is actually a function of time and the current level of the underlying. Sections 5.3.1, 5.3.2, 5.3.3 will show how the valuation standard partial differential equations well known for those products are actually particular cases of the functional valuation partial differential equation. The valuation functional equation can therefore be seen as a unifying general valuation equation in a functional volatility model. One should note that the pricing functional in 5.3.2 is locally regular but fails to be even continuous at fixed times. Section 5.3.4 will treat the case of a theoretical (continuous-time) Variance Swap and show that the functional valuation partial differential equation holds for this product as well, and gives rise to the standard local volatility partial differential equation with an additional source term. Section 5.3.5 shows that the valuation functional equation holds for payoffs which are functionals of a finite number of observation of the stock price, which is actually the case of almost all real-life options (the exception being continuously monitored barrier clauses). In this case as well, the pricing functional is locally regular but fails to be $\mathbb{C}_b^{1,2}([0,T[)$. Finally, section 5.3.6 considers multi-asset functional volatility model, with dependence on the realized quadratic co-variation.

In all the above examples, except in section 5.3.6, the model is a standard local volatility model in dimension 1, i.e. $d = n = 1$ and:

$$\sigma_t(S_t, [S]_t) = \sigma(t, S(t)) \tag{5.24}$$

for some function σ which is assumed to be continuous, bounded by some constant $M > 0$ and bounded from below by some constant $\eta > 0$, and

$$r_t(S_t, [S]_t) = r \tag{5.25}$$

for some constant $r \geq 0$.

5.3.1 Vanilla options

A vanilla option takes the form:

$$g(x) = h(x(T)) \tag{5.26}$$

for some measurable function g with linear growth. A functional of the form:

$$F_t(x_t, v_t) = f(t, x(t)) \tag{5.27}$$

where f is $C^{1,2}$ on $[0, T[\times]0, \infty[$ and continuous on $[0, T]\times]0, \infty[$, satisfies the valuation functional equation if and only if f is a $C^{1,2}$ solution of the classical valuation PDE in Dupire's local volatility model [22], on $[0, T[\times(0, \infty)$:

$$f_t(t, x) + rxf_x(t, x) + \frac{1}{2}x^2\sigma^2(t, x)f_{xx}(t, x) = rf(t, x) \tag{5.28}$$

with terminal condition:

$$f(T, x) = g(x) \tag{5.29}$$

Classical parabolic PDE theory [44] guarantees existence and uniqueness of a solution with at most exponential growth. Hence, the vanilla option valuation PDE in local volatility is a particular case of 5.1.

5.3.2 Continuously monitored barrier options

Continuously monitored barrier options trade in the market. In case of litigation, it would be up to the side benefitting from the barrier to prove that a trade has breached the barrier. He usually places a stop loss order at the barrier with a respectable prime-broker, so that if the barrier is triggered his own trade will be executed at or beyond the barrier. The option is in that case:

$$g(x) = h(x(T))1_{\sup_{0\leq t\leq T} x(t)<U} \tag{5.30}$$

where h is a measurable function with linear growth, for an up-and-out option. In barriers can be written as differences of out options and vanillas, and down barriers would have a similar treatment. We look for a solution to the valuation functional equation of the form:

$$F_t(x_t, v_t) = 1_{\sup_{0\leq t\leq T} x(t)<U} f_0(t, x(t)) + 1_{\sup_{0\leq t\leq T} x(t)\geq U} f_1(t, x(t)) \tag{5.31}$$

where f_0 is $C^{1,2}$ on $[0,T[\times]0,U[$ and continuous on $[0,T]\times]0,U] - \{(T,U)\}$, and f_1 is $C^{1,2}$ on $[0,T[\times]0,\infty[$ and continuous on $[0,T]\times]0,\infty[$. As in example 4.4, such a functional is in $\mathcal{R}_{loc}$, but fails to be continuous at fixed times. It satisfies the valuation functional equation if and only if f_0 satisfies the classical barrier PDE:

$$f_t(t,x) + rxf_x(t,x) + \frac{1}{2}x^2\sigma^2(t,x)f_{xx}(t,x) = rf(t,x) \tag{5.32}$$

with boundary and terminal conditions:

$$f_0(t,U) = 0, f_0(T,x) = h(x) \tag{5.33}$$

while f_1 satisfies the classical vanilla PDE 5.28 with terminal condition h. Hence the classical PDE with boundary condition for pricing continuously monitored barrier options also appears as a particular case of the valuation functional equation.

5.3.3 Continuously monitored Asian options

A continuously monitored Asian option is defined as:

$$g(x) = h\left(\int_0^T x(s)ds\right) \tag{5.34}$$

where the function h is measurable and has linear growth. A functional F of the form:

$$F_t(x_t, v_t) = f(t, \int_0^t x(s)ds, x(t)) \tag{5.35}$$

where f is $C^{1,1,2}$ on $[0,T[\times]0,\infty[\times]0,\infty[$ and continuous at time T is solution of the valuation functional equation if and only if f solves:

$$f_t(t,a,x) + rxf_x(t,a,x) + xf_a(t,a,x) + \frac{1}{2}x^2\sigma^2(t,x)f_{xx}(t,a,x) = rf(t,a,x) \tag{5.36}$$

with terminal condition:

$$f(T,a,x) = h(a) \tag{5.37}$$

Hence the standard Asian PDE is a particular case of the valuation functional equation. Note that in [23] a better parametrization of the functional F is introduced to obtain another PDE which is more suitable for numerical solutions, and that [63] shows how the value of the functional F at time 0 can be recovered by solving a $1+1$-dimensional PDE.

5.3.4 Continuously monitored variance swap

A continuously monitored Variance Swap is the exchange between two counterparties of the realized quadratic variation of the logarithm of the stock price between contract inception and payment date versus an amount determined at inception. It can be priced in the local volatility model as the payoff:

$$g(x) = \int_0^T \sigma^2(t, x(t))dt - K \tag{5.38}$$

A Variance Swap is usually priced using the fact that it has the same value as a static combination of calls and puts [19]. We look here for a solution to the valuation functional equation of the form $F_t(x_t, v_t) = \int_0^t \sigma^2(s, x(s))ds + f(t, x(t))$. Elementary computation shows that, if f is a $C^{1,2}$ solution of the following standard local volatility valuation PDE with an added source term on $[0,T[\times(0,\infty)$, which is continuous on $[0,T]\times(0,\infty)$:

$$f_t(t,x) + rxf_x(t,x) + \frac{1}{2}x^2\sigma^2(t,x)f_{xx}(t,x) - rf(t,x) = -\sigma^2(t,x) \tag{5.39}$$

with terminal condition:

$$f(T,x) = 0 \tag{5.40}$$

then F is a regular solution of the valuation functional equation 5.1.

5.3.5 Path-dependent options with discrete monitoring

With the exception of continuously monitored barriers, all real-life exotic derivatives are actually payoff taking the form:

$$g(x) = h(x(t_0), x(t_1), ..., x(t_m)) \tag{5.41}$$

where $0 = t_0 \leq t_1 < \ldots < t_m = T$. Some frequently encountered examples are:

- Variance Swap:

$$h(x(t_0), x(t_1), ..., x(t_m)) = \frac{1}{T} \sum_{i=0}^{m-1} log^2 \left(\frac{x(t_{i+1})}{x(t_i)} \right)$$

  with typically $t_{i+1} - t_i = 1$ trading day

- Discretely monitored barrier:

$$h(x(t_0), x(t_1), ..., x(t_m)) = 1_{\max_{0 \leq i \leq m} x(t_i) \leq U} (x(t_m) - K)^+$$

  with typically daily but sometimes weekly monitoring.

- Asian:

$$h(x(t_0), x(t_1), ..., x(t_m)) = \left(\frac{1}{m-1} \sum_{i=0}^{m-1} x(t_i) - K \right)^+$$

  where observation can be daily, weekly, monthly.

- Cliquet:

$$h(x(t_0), x(t_1), ..., x(t_m)) = \min \left[\sum_{i=0}^{m-1} \left(\frac{x(t_{i+1})}{x(t_i)} - K \right)^+ , U \right]$$

  with typically $t_{i+1} - t_i = 1$ or 3 months.

- Lookback:

$$h(x(t_0), x(t_1), ..., x(t_m)) = \left[\max_{1 \leq i \leq m} x(t_i) - x(T) \right]$$

  or a similar expression with the min. Observation frequence can be daily, weekly, monthly...

Remark 5.3. The prices of Asians and Variance Swaps described above are approximated by the prices of their theoretical versions, however the behavior of the greeks is different. In trading floors, theoretical continuous-time options are often used to compute prices, but the greeks are computed in a way that takes into account the discrete fixings.

Define

$$f^m(x_0, \dots, x_m, t_m, x) = h(x_0, \dots, x_m) \tag{5.42}$$

and then backward recursively for $0 \leq i \leq m-1$, $f^i(x_0, x_1, \dots, x_i, t, x)$, as the $C^{1,2}$ solutions of:

$$\begin{aligned} f^i_t(x_0, x_1, \dots, x_i, t, x) + rxf^i_x(x_0, x_1, \dots, x_i, t, x) \\ +\frac{1}{2}x^2\sigma^2(t,x)f^i_{xx}(x_0, x_1, \dots, x_i, t, x) = rf^i(x_0, x_1, \dots, x_i x, t) \end{aligned} \tag{5.43}$$

on $[t_i, t_{i+1}[\times]0, \infty[$ with terminal conditions:

$$f^i(x_0, x_1, \dots, x_i, t_{i+1}, x) = f^{i+1}(x_0, x_1, \dots, x_i, x, t_{i+1}, x) \tag{5.44}$$

and at most exponential growth. Then define the functional:

$$F_t(x_t, v_t) = \sum_{i=0}^{m-1} f^i(x(t_0), \dots, x(t_i), t, x(t))1_{t\in[t_i,t_{i+1}[} + f^n(x(t_0), \dots, x(t_n), t_n, x(t_m)1_{t=t_m}$$

This functional is in $\mathcal{R}_{loc}$ because it satisfies for $x \in C_0([0,T],]0, \infty[)$:

$$F_t(x_t, v_t)1_{t\in[t_i,t_{i+1}[} = f^i(x(t_0-), \dots, x(t_i-), t, x(t)) \tag{5.45}$$

and it satisfies the functional valuation partial differential equation. If it has continuous path on some set $\mathfrak{D}$ with full measure, F then prices the option on periodic fixings.

Remark 5.4. What allows us to show that $F \in \mathcal{R}_{loc}$ is the apparent "trick" to apply the function f^i at points x_{t_i-} rather than x_{t_i} (otherwise the functional defined as such would fail to be vertically differentiable at time t_i). This is more than a trick but points out to the correct understanding of the Δ of an option: if t_j is an observation time, a trader computes the Δ at time t_j by moving the spot "right after" having made the observation which is kept constant.

5.3.6 Options on basket in a model with an unobservable factor

In this section, for the sake of keeping notations as simple as possible, it is assumed that the interest rate is 0. An important segment of the exotics business of equity derivatives houses are the options on basket of indices. Typical payoffs include the following:

- Best-of option: $g(s_T) = \max(\max_{1 \geq i \geq d}(\frac{s^i(T)}{s^i(0)}), 0)$
- Call on basket: $g(s_T) = \max(\sum_{i=1}^{d} c_i \frac{s^i(T)}{s^i(0)}, K)$, where c_i is the weight of asset i in the basket
- Outperformance option: $g(s_T) = max(\frac{s^1(T)}{s^1(0)} - \frac{s^2(T)}{s^2(0)})$

These options are usually written on indices such as S&P500, bond indices, technology companies index, utilities index, etc, which are typically sectors for which institutional investors decide target allocations, and very sensitive to the correlation parameter. A model proposed by Cont [10] takes into account the price-impact effect of the trades of institutional investors on the variance-covariance structure of the assets. The idea is that institutional investors typically profit from the dispersion by allocating constant weights to the different sectors in which they trade, hence the value of their portfolio evolves as:

$$\frac{dV(t)}{V(t)} = \sum x_i \frac{dS_t^i}{S_t^i} \tag{5.46}$$

where x_i represent the weight allocated to sector index i by the institutional investors. It is straightforward that $V(t) = h(S_t, A_t)$ where:

$$h(s_t, v_t) =$$

$$h_0 \exp\left(\sum_i x_i \log \frac{s^i(t)}{s^i(0)} + \frac{1}{2} \sum_i \int_0^t \frac{x_i}{(s^i(s))^2} v^{ii}(s) ds - \frac{1}{2} \sum_{i,j} \int_0^t \frac{x_i x_j}{s^i(s) s^j(s)} v^{i,j}(s) ds \right)$$

defines an element of $\mathbb{F}^\infty$ which is predictable in the second variable. The model for the evolution of the indices incorporates the modification of the local drift and variance-covariance structure due to the trading of the institutional investors; both are actually bounded functions of the spot prices and the size of the institutional investors, so that the model is actually a functional volatility model:

$$\frac{dS_t^i}{S_t^i} = b_i(S(t), h(S_t, A_t))dt + \sigma_i(S(t), h(S_t, A_t)).\Sigma dB_t \tag{5.47}$$

where the matrix Σ is a square-root of fundamental variance-covariance in the market in the absence of price impact. We look for a solution to the functional valuation of the form $F_t(s_t, v_t) = f(t, s(t), h(s_t, v_t))$. Then (omitting the variables for the sake of readability):

$$\begin{aligned}\mathcal{D}_t F(s_t, v_t) = \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial f}{\partial h} h \sum i \frac{x_i}{(s^i(t))^2}\sigma_i \Sigma \sigma_i' \\ -\frac{1}{2}\frac{\partial f}{\partial h} h \sum i, j \frac{x_i x_j}{s^i(t) s^j(t)}\sigma_i \Sigma \sigma_j'\end{aligned} \tag{5.48}$$

$$\partial_j F_t(s_t, v_t) = \frac{\partial f}{\partial s^j} + \frac{\partial f}{\partial h}\frac{x_j}{s^j} h \tag{5.49}$$

$$\begin{aligned}\partial_{ij}^2 F_t(s_t, v_t) = \frac{\partial^2 f}{\partial s^j \partial s^i} + \frac{\partial f}{\partial h}\frac{x_i x_j}{s^i(t) s^j(t)} h \\ + \frac{\partial^2 f}{\partial^2 v}\frac{x_i x_j}{s^i(t) s^j(t)} h^2 + \frac{\partial^2 f}{\partial h \partial s^j}\frac{x_i}{s^i(t)} h, i \neq j\end{aligned} \tag{5.50}$$

$$\begin{aligned}\partial_{ii}^2 F_t(s_t, v_t = \frac{\partial^2 f}{\partial^2 s^i} + \frac{\partial f}{\partial v}\frac{x_i^2 - x_i}{(s^i(t))^2} h \\ + \frac{\partial^2 f}{\partial^2 v}\frac{x_i^2}{(s^i(t))^2} h^2 + \frac{\partial^2 f}{\partial v \partial s^i}\frac{x_i}{s^i(t)} h\end{aligned} \tag{5.51}$$

Therefore, if the function f is a solution of the PDE

$$\begin{aligned}\frac{\partial f}{\partial t} + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial s^i \partial s^j} s^i s^j \sigma_i \Sigma \sigma_j' + \frac{1}{2}\frac{\partial^2 f}{\partial^2 h} h^2 \sum_{i,j} x_i x_j \sigma_i \Sigma \sigma_j' \\ + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial v \partial s^i} x_i h s^j \sigma_i \Sigma \sigma_j' = 0\end{aligned} \tag{5.52}$$

with the terminal condition

$$f(T, s, v) = g(s) \tag{5.53}$$

that is C^1 in t, jointly C^2 in (s, h) on the set $[0, T[\times]0, \infty[^d \times]0, \infty[$, and continuous on $[0, T]\times]0\infty[^d \times]0, \infty[$, then the functional F prices the option g. Hence the Greeks of the option are as follows:

$$\begin{aligned}\Theta_t = \mathcal{D}_t F(S_t, A_t) = \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial f}{\partial h} h \sum_i \frac{x_i}{(S^i(t))^2}\sigma_i \Sigma \sigma_i' \\ -\frac{1}{2}\frac{\partial f}{\partial h} h \sum_{i,j}\frac{x_i x_j}{S^i(t) S^j(t)}\sigma_i \Sigma \sigma_j'\end{aligned} \tag{5.54}$$

$$\Delta_t^i = \frac{\partial f}{\partial s^j} + \frac{\partial f}{\partial h}\frac{x_j}{S^j(t)}h \tag{5.55}$$

$$\Gamma_t^{ii} = \frac{\partial^2 f}{\partial^2 s^i} + \frac{\partial f}{\partial h}\frac{x_i^2 - x_i}{(S^i(t))^2}h$$
$$+\frac{\partial^2 f}{\partial^2 h}\frac{x_i^2}{(S^i(t))^2}h^2 + \frac{\partial^2 f}{\partial h \partial s^i}\frac{x_i}{S^i(t)}h \tag{5.56}$$

$$\Gamma_t^{ij} = \frac{\partial^2 f}{\partial s^j \partial s^i} + \frac{\partial f}{\partial v}\frac{x_i x_j}{S^i(t)S^j(t)}h$$
$$+\frac{\partial^2 f}{\partial^2 v}\frac{x_i x_j}{S^i(t)s^j(t)}h^2 + \frac{\partial^2 f}{\partial v \partial s^j}\frac{x_i}{S^i(t)}h \tag{5.57}$$

Note that in this case $\Theta_{0vol} = \frac{\partial f}{\partial t} \neq \Theta$, because if time passes and the indices realize zero volatility then V does not move either. The correct definition of the Θ is that time has passed, the spot are still at the same level at the next observation time for the trader but they have realized their instantaneous variance-covariance meanwhile, so that V has moved accordingly. Also note that the Δ and the Γ in this case are not the naive derivatives of the function f with respect to the spot.

5.4 Sensitivities to market variables

In this section, we will consider functionals dependent on the first argument only, the short-term interest rate is some deterministic function $r(t)$ and we are working in dimensions $d = n = 1$. Generalization to higher dimensions is straightforward with heavier notations. We consider that options g are priced in a functional volatility model:

$$S(t) = S(t)r(t)dt + S(t)\sigma_t(S_t, V)dW(t) \tag{5.58}$$

where $V \in \mathbb{R}^{m+1}$ is such that $V_0 = x$ is the price of the underlying asset at date 0, and the vector $\tilde{V} = (V_1, \ldots, V_m)$ is a set of observable market variables at date 0, called the calibration data of the model. We assume that, for any $V \in \mathbb{R}^{m+1}$, the coefficients of the stochastic differential equation satisfy the assumptions of theorem B.1. In particular, there exists a unique strong solution to this SDE and it is square-integrable. This situation models the reality of pricing at a trading desk, which is using a pricing model parameterized by calibration data *at a given time* to price the option at this time.

In-house research departments in well-known derivatives houses have been developing various methods to compute the sensitivities to these observable market variables, but most of this work is not publicly available. One common way, when working in a local volatility model and pricing by PDE methods, is to perform a perturbation analysis of the valuation PDE. However, in the general path-dependent case, valuation PDEs are generalized by the valuation functional equation (5.9) and therefore PDE perturbation theory does not apply. Using the functional formalism, we will give in this chapter a precise meaning to the notion of sensitivity to calibration parameters. Building on Dupire's insight [23], we will use our perturbation result for stochastic differential equations with functional coefficients (theorem 5.3) and the $\Theta - \Gamma$ tradeoff formula (theorem 5.2) to actually obtain expressions for these sensitivities which can be used for efficient numerical computation (section 5.4.2). In particular, we will be able to define and compute the Vega buckets and total Vega of a path-dependent derivative. In section 5.4.3, we will define the "Sticky Strike" Delta, which can be viewed as the Black-Scholes delta of any derivative, and Deltas with partial or null realization of the skew, which are the deltas actually used by traders to delta-hedge their derivatives position rather than the model delta from definition 5.1. We will conclude in section 5.4.4 by suggesting an efficient numerical algorithm for computing the sensitivities to market variables and Deltas in a local volatility model.

Remark 5.5. The sensitivities treated in this section do not include the observable points of the interest rate curve, since the short-term rate $t \mapsto r(t)$ of the pricing model does not depend on the market data V.

5.4.1 Directional derivatives with respect to the volatility functional

Let σ be a functional such that $(x_t) \mapsto x(t)\sigma(x_t)$ satisfies the assumptions of theorem B.1, and σ^ϵ be a family of functionals such that $(x_t) \mapsto x(t)\sigma_t^\epsilon(x_t)$ satisfies the assumptions of theorem B.1. Denote S^ϵ the unique strong solution of:

$$\frac{dS^\epsilon(t)}{S^\epsilon(t)} = r(t)dt + \sigma_t^\epsilon(S_t^\epsilon)dW(t) \tag{5.59}$$

We assume that there exists a bounded functional $\tilde{\sigma}$ satisfying:

$$|\sigma_t^\epsilon(x_t) - \sigma_t(x_t) - \epsilon\tilde{\sigma}_t(x_t)| \le \epsilon\phi(\epsilon) \tag{5.60}$$

where ϕ is an increasing function from $(0,\infty)$ to $(0,\infty)$, $\lim_{\epsilon\to 0}\phi(\epsilon) = 0$. Denote S^ϵ the solution of 5.58 for volatility σ^ϵ. Let g be a payoff priced by a functional F in the model σ, satisfying the assumptions of theorem 5.2. We furthermore make the assumption that there exists a positive measurable function $g : [0,T] \mapsto (0,\infty)$, such that:

$$|x^2(t)\nabla_x^2 F_t(x_t) - x'^2(t)\nabla_x^2 F_t(x'_t)| \le g(t) \sup_{s\in[0,t]} |x(t) - x'(t)| \tag{5.61}$$

and that for some constant $C > 0$, for all $t \le T$:

$$|\tilde{\sigma}_t(x_t) - \tilde{\sigma}_t(x'_t)| \le C \sup_{s\in[0,t]} |x(t) - x'(t)| \tag{5.62}$$

The following theorem was first given with a heuristic argument in [23].

Theorem 5.3 (Sensitivity to the functional volatility)**.**

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left| E[g(S_T^\epsilon) - g(S_T)]e^{-\int_0^T r(s)ds} - \epsilon E\left[\int_0^T \tilde{\sigma}_t\sigma_t(S_t)S^2(t)\nabla_x^2 F_t(S_t)e^{-\int_0^t r(s)ds}\right]\right| = 0$$

Proof. Remember that $E[g(S_T)]e^{-\int_0^T r(s)ds} = F_0(S_0)$. Applying theorem 5.2 to F and the process S^ϵ yields:

$$E\left[(S_T^\epsilon)e^{-\int_0^T r(s)ds} + \frac{1}{2}[(\sigma^\epsilon)_t^2(S_t^\epsilon) - \sigma_t(S_t^\epsilon)](S^\epsilon(t))^2\nabla_x^2 F_t(S_t^\epsilon)e^{-\int_0^t r(s)ds}\right] = F_0(S_0) \tag{5.63}$$

Now, $(\sigma^\epsilon)_t^2(S_t^\epsilon) - \sigma_t^2(S_t^\epsilon) = 2\epsilon\tilde{\sigma}_t\sigma_t(S_t^\epsilon) + \epsilon R(S_t^\epsilon,\epsilon)$ where $R(S_t^\epsilon,\epsilon)$ is bounded by $C\phi(\epsilon)$ for some constant C because of assumption (5.60).

Therefore, we just have to prove that:

$$\int_0^T |\tilde{\sigma}_t\sigma_t(S_t^\epsilon)(S^\epsilon(t))^2\nabla_x^2 F_t(S_t^\epsilon) - \tilde{\sigma}_t\sigma_t(S_t)S^2(t)\nabla_x^2 F_t(S_t)|dt \to 0 \tag{5.64}$$

as $\epsilon \to 0$. Let $\eta > 0$, and h an integrable function bounding $E[|S^2(t)\nabla_x^2 F_t(S_t)|]$ as in (5.20), and C a constant bounding $\tilde{\sigma}\sigma$ and for which condition (5.62) is satisfied. We can first fix $R > 0$ such that $\int_{\{t:g(t)\vee h(t)>R\}} h(t) < \frac{\eta}{4C}$, so that the integral (5.64) on the set

$\{t : g(t) \vee h(t) > R\}$ is bounded by $\frac{\eta}{2}$.

On the complementary set $\{t : g(t) \vee h(t) \leq R\}$, we will work with the decomposition:

$$\begin{aligned}\tilde{\sigma}_t\sigma_t(S^\epsilon_t)(S^\epsilon(t))^2\nabla^2_x F_t(S^\epsilon_t) - \tilde{\sigma}_t\sigma_t(S_t)S^2(t)\nabla^2_x F_t(S_t) = \\ [\tilde{\sigma}_t(S^\epsilon_t) - \tilde{\sigma}_t(S_t)]\,\sigma_t(S^\epsilon_t)(S^\epsilon(t))^2\nabla^2_x F_t(S^\epsilon_t) \\ +\tilde{\sigma}_t(S_t)[\sigma_t(S^\epsilon_t) - \sigma_t(S_t)](S^\epsilon(t))^2\nabla^2_x F_t(S^\epsilon_t) \\ +\tilde{\sigma}_t\sigma_t(S_t)[(S^\epsilon(t))^2\nabla^2_x F_t(S^\epsilon_t) - S^2(t)\nabla^2_x F_t(S_t)]\end{aligned} \tag{5.65}$$

The difference under brackets in the three terms in decomposition (5.65) are bounded by $\max(C, R)\sup_{s\leq t}|S^\epsilon_t - S_t|$, because of assumptions (5.62), (5.60), while the rest is bounded by a constant. Hence the stochastic differential equation perturbation theorem 5.3 and Cauchy-Shwarz inequality conclude. □

5.4.2 Sensitivities to market variables

We assume that σ is differentiable with respect to V for the sup norm, uniformly in time, that is for any $V \in \mathbb{R}^{m+1}, t \geq 0$ there exists functionals $\frac{\partial\sigma_t}{\partial V_i}(.,.,V), 0 \leq i \leq m$ such that, for $h \in \mathbb{R}^{m+1}, |h| = 1$:

$$|\sigma_t(x_t, V + \epsilon h) - \sigma_t(x_t, V) - \epsilon\sum_{i=0}^{m} h_i\frac{\partial\sigma_t}{\partial V_i}(x_t, V)| \leq \epsilon\phi(\epsilon, V) \tag{5.66}$$

where $\phi(., V)$ is an increasing function from $(0, \infty)$ to $(0, \infty)$, $\lim_{\epsilon\to 0}\phi(\epsilon, V) = 0$. Let g be a payoff priced in the model $\sigma_t(., V)$ by a functional $F(., V)$ satisfying the assumptions of theorem 5.3. The following proposition is an immediate consequence of theorem 5.3 and allows to compute the sensitivity of the option g to market variables $\tilde{V}$. Let $(e_0, \ldots, e_m)$ denote the canonical basis of $\mathbb{R}^{m+1}$.

The following proposition, which is a direct corollary to theorem 5.3, allows for the explicit computation of the sensitivities of a derivative with respect to observable market variables:

Proposition 5.1.

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{F_0(S_0, V + \epsilon e_i) - F_0(S_0, V)}{\epsilon} = \\ E\left[\int_0^T \frac{\partial\sigma_t}{\partial V_i}(S_t, V)\sigma_t(S_t, V)S^2(t)\nabla^2_x F_t(S_t, V)e^{-\int_0^t r(s)ds}dt\right]\end{aligned} \tag{5.67}$$

This quantity is called the sensitivity of the derivative g in to the market variable V_i.

Implied volatility of call and put options on the underlying S which are traded on exchanges are naturally among the market variables parameterizing the model. From now, we will assume that there are $m_0 \leq m$ traded pairs of strike and expiries $(S_i, T_i), 1 \leq i \leq m_0$ and that the coordinates $V_i, 1 \leq i \leq m_0$ are the implied volatilities at those strike-expiry pairs.

Definition 5.2 (Local Vega)**.** For $i \leq m_0$,

$$Vega_{T_i,K_i} := \lim_{\epsilon \to 0} \frac{F_0(S_0, V + \epsilon e_i) - F_0(S_0, V)}{\epsilon} \tag{5.68}$$

is called the local Vega of the derivative g at the bucket (T_i, K_i).

$$Vega := \sum_{i=1}^{m_0} Vega_{T_i,K_i} = \lim_{\epsilon \to 0} \frac{F_0(S_0, V + \epsilon \sum_{i=1}^{m_0} e_i) - F_0(S_0, V)}{\epsilon} \tag{5.69}$$

is called the Vega of the derivative g.

The Vega of the derivative g in the bucket (T_i, K_i) is its sensitivity to a move of the implied volatility at (T_i, K_i), with the underlying and all other market variables remaining constant. Its Vega is its sensitivity to a parallel shift in the implied volatility surface. These sensitivities are actually the main tools for volatility traders to understand their position, and their decision-making process often consists of deciding a target Vega bucket exposure.

5.4.3 Multiple Deltas of a derivative

The notion of model Δ (definition 5.1) of a derivative is a sensitivity assuming that the model, that is the functional volatility σ, remains constant when the spot in bumped. In terms of observable market quantities, it means that, if the current price $V_0 = x$ is bumped to $x + \epsilon$, the observable market variables $V_i, 1 \leq i \leq m$ are bumped to some new value V'_i so that $\sigma(.,.,V) = \sigma(.,.,(x + \epsilon, \tilde{V}'))$. This joint dynamics imposed by the model to the underlying and the implied volatility smile has no intuitive meaning to practitioners, and has been shown to be unrealistic on the market [3, 4]. On another hand, vanillas are usually traded according to their Black-Scholes Δ, which assumes that implied volatility at a given strike and expiry remains constant as the underlying moves. Consistency when trading a portfolio comprising of exotics and vanillas requires to trade also exotics according to a Δ factoring the same assumption on implied volatilities, which is defined in definition 5.2. This Delta is often called “Sticky Strike” delta by practitioners, but, as it is the natural generalization of the Black-Scholes Delta of a European call or put, we will call it here Black-Scholes Delta of a general path-dependent derivative.

Definition 5.3. Define the Black-Scholes Delta of a derivative priced by the functional F as:

$$\Delta_{BS} = \Delta_0(S_0, V) + \lim_{\epsilon \to 0} \frac{F_0(S_0, V + \epsilon e_0) - F_0(S_0, V)}{\epsilon} \tag{5.70}$$

The following proposition allows for explicit computation of the Black-Scholes Delta:

Proposition 5.2.

$$\begin{aligned} \Delta_{BS} &= \nabla_x F_0(S_0, V) \\ +E &\left[\int_0^T \frac{\partial \sigma_t}{\partial V_0}(S_t, V)\sigma_t(S_t, V)S^2(t)\nabla_x^2 F_t(S_t, V)e^{-\int_0^t r(s)ds} dt \right] \end{aligned} \tag{5.71}$$

Many traders would trade their derivatives portfolio (exotic or vanillas) according to their Black-Scholes Δ, but many would rather incorporate a view on the realized skew in their Δ. To the best of our knowledge, this concept has only appeared in literature in [4]. The valuation system of a trading desk often incorporates an interpolation/extrapolation tool which maps the market data V to a full volatility surface
$\sigma_{BS}(.,.,V)$ such that $\sigma_{BS}(T_i, K_i, V) = V_i$. It is assumed that the implied volatility surface $\sigma_{BS}(T, K, V)$ is differentiable in K. The skew realization ratio is the expected variation of the implied volatility at a relative strike (T, k) for a relative bump in the spot price x of the underlying, expressed in units of the relative skew $K\frac{\partial \sigma_{BS}(T,K,V)}{\partial K}$ in the original implied volatility surface. More precisely, if one defines $\tilde{\sigma}_{BS}(T, k) = \sigma_{BS}(T, kx)$ the reparameterization of the implied volatility surface in terms of relative strike, then:

$$\alpha = \frac{\Delta \tilde{\sigma}_{BS}(T, k)}{K\frac{\partial \sigma_{BS}}{\partial K}(T, K, V)\frac{\Delta x}{x}} \tag{5.72}$$

α is called the skew-stickiness ratio [4]. It should of course be dependent of the strike and maturity, but traders use to choose some constant value for this ratio and compute their delta in consequence. It represents the amount by which implied volatility at the constant *relative* bucket (T, k) varies, for a small relative move $\frac{\Delta x}{x}$ in the underlying, expressed in units of the relative implied volatility skew. $\alpha = 1$ represents a so-called "Sticky Strike" dynamics, which corresponds to the Black-Scholes Delta. $\alpha = 0$ is called a "Sticky Delta" dynamics and represents a volatility surface which is invariant in terms of relative strikes; this is the dynamics that the implied volatility has in exponential Lévy models [11]. Bergomi [4] shows that, at the limit of small skews and small times, diffusion models imply a smile dynamics with $\alpha \approx 2$ at-the-money. He also shows that realistic values observed in the market for the skew stickiness ratio at-the-money tend to be strictly between 1 and 2.
The variation of the implied volatility at the constant *absolute* bucket (T, K), can be expressed as:

$$\Delta\sigma(T, K) = \Delta\tilde{\sigma}_{BS}(T, k) - K\frac{\Delta(x)}{x}\frac{\partial\sigma_{BS}}{\partial K}(T, K, V) \tag{5.73}$$

so that:

$$\frac{\Delta\sigma_{BS}(T, K)}{\Delta x} = (\alpha - 1)\frac{K}{x}\frac{\partial\sigma_{BS}}{\partial K}(T, K, V) \tag{5.74}$$

We can there give a rigorous definition to the Delta at skew stickiness α of a derivative as follows:

Definition 5.4 (Delta with skew stickiness ratio)**.** Define the Delta at skew stickiness α of a derivative priced by the functional F as:

$$\begin{aligned} \Delta_\alpha &= \Delta_0(S_0, V) \\ + \lim_{\epsilon \to 0} &\frac{F_0\left(S_0, V + \epsilon e_0 + \epsilon \sum_{i=1}^{m_0}(\alpha - 1)\frac{K_i}{x}\frac{\partial\sigma_{BS}}{\partial K}(T_i, K_i, V)\right) - F_0(S_0, V)}{\epsilon} \end{aligned}$$

The following proposition allows for the explicit computation of the Delta at skew stickiness ratio α:

Proposition 5.3.

$$\Delta_\alpha = \nabla_x F_0(S_0, V)$$
$$+E[\int_0^T \left(\frac{\partial \sigma}{\partial V_0}(t, S(t), V) + \sum_{i=1}^{m_0} (\alpha - 1) \frac{K_i}{x} \frac{\partial \sigma_{BS}}{\partial K}(T_i, K_i, V) \frac{\partial \sigma_t}{\partial V_i}(S_t, V) \right)$$
$$\sigma_t(S_t, V) S^2(t) \nabla_x^2 F_t(S_t, V) e^{-\int_0^t r(s)ds} dt]$$

Remark 5.6. Deltas at any skew stickiness ratio can be recovered from the Sticky Strike (Black Scholes) Delta and the Sticky Delta Delta:

$$\Delta_\alpha = \alpha \Delta_{BS} + (1 - \alpha)\Delta_0 \tag{5.75}$$

Hence only those two Deltas are returned by valuation systems, and the traders use the linear combination corresponding to their view on the skew stickiness ratio.

5.4.4 Efficient numerical algorithm for the simultaneous computation of Vega buckets and Deltas

Computing the bucket exposure of a derivative and is usually done by bumping the implied volatility in the concerned bucket, re-constructing the local volatility corresponding to the new implied volatility surface, and re-pricing the derivative. Similarly, its Black-Scholes Delta is usually computed by bumping the spot, keeping implied volatilities constant, re-constructing a local volatility surface and then repricing the derivative. Since for well-traded indices buckets are usually quite numerous, such a method is often too time-consuming to be performed intraday. In the case where the functional F (and not only the initial price of the derivative) is computable (which is true for example for barriers, Asians, variance swaps, since the valuation functional equation reduces to low-dimensional ordinary PDEs), propositions 5.1, 5.2 already give a better algorithm to compute those sensitivities by a unique Monte-Carlo under the original local volatility, and where the integral is computed numerically for each path. This will require a numerical integration per sensitivity and per path.

Assume now that the pricing model is a Dupire local volatility model [22]:

$$\sigma_t(S_t, [S]_t) = \sigma(t, S(t)) \tag{5.76}$$

Assume also that the in-house valuation system allows for almost instantaneous precise computation of the prices of vanilla option payoffs, that is quantities taking the form $E[S(t_i)]$. We are then able suggest a more efficient algorithm that only requires a unique Monte-Carlo followed by a single numerical integration per sensitivity. The idea is that:

$$E\left[\int_0^T \frac{\partial\sigma}{\partial V_i}(t,S(t),V)\sigma(t,S(t),V)S^2(t)\nabla_x^2 F_t(S_t,V)e^{-\int_0^t r(s)ds}\right] =$$
$$E\left[\int_0^T \frac{\partial\sigma}{\partial V_i}(t,S(t),V)\sigma(t,S(t),V)S^2(t)E[\nabla_x^2 F_t(S_t,V)e^{-\int_0^t r(s)ds}|S(t)]\right] \tag{5.77}$$

The projection $E[\nabla_x^2 F_t(S_t,V)e^{-\int_0^t r(s)ds}|S(t)]$ can be approximated by the projection on a finite-dimensional subspace of the Hilbert space of the space of functions of $S(t)$. Hence the method is:

- Simulate N paths $S(\omega_1),\ldots,S(\omega_N)$ of the underlying S under the original local volatility:
$$\frac{dS(t)}{S(t)} = r(t)dt + \sigma(t,S(t))dW(t)$$
- choose a finite number of bounded functions $f_j, 1 \le j \le K$
- For each time step t_i in the simulation of the path, perform the linear regression of the Gamma: $\nabla_x^2 F_t(S_{t_i})$ on the explanatory variables $f_j(S(t_i))$ using the draws $S_{t_i}(\omega_1),\ldots,S_{t_i}(\omega_n)$: define $(\alpha_0(t_i),\ldots,\alpha_K(t_i))$ solving the minimization problem:
$$\min_{u\in\mathbb{R}^{K+1}} \sum_{j=1}^{n} |\nabla_x^2 F_t(S_{t_i}(\omega_j)) - u_0 - \sum_{k=1}^{K} u_k f_k(S(t_i)(\omega_j))|^2 \tag{5.78}$$
and define $f(t_i,x) = \alpha_0(t_i) + \sum_{j=1}^K \alpha_j(t_i) f_j(x)$
- For each bucket, compute the sum:
$$\sum_{t_i<T}(t_{i+1}-t_i)E\left[\frac{\partial\sigma}{\partial V_i}(t_i,S(t_i),V)\sigma(t_i,S(t_i),V)S^2(t_i)f(t_i,S(t_i))\right] \tag{5.79}$$

Bibliography

[1] H. Ahn, *Semimartingale integral representation*, Ann. Probab., 25 (1997), pp. 997–1010.

[2] A. Y. Baird, *Option Market Making*, Wiley Finance, 1993.

[3] L. Bergomi, *Dynamic properties of smile models*, Frontiers in Quantitative Finance, (2009), pp. 47–88.

[4] ——, *Smile dynamics 4*, Risk, 22 (2009).

[5] P. Billingsey, *Convergence of Probability Measures*, Wiley, Princeton, N.J., 1999.

[6] J.-M. Bismut, *A generalized formula of Itô and some other properties of stochastic flows*, Z. Wahrsch. Verw. Gebiete, 55 (1981), pp. 331–350.

[7] ——, *Calcul des variations stochastique et processus de sauts*, Z. Wahrsch. Verw. Gebiete, 63 (1983), pp. 147–235.

[8] F. Black and M. S. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy, (1973).

[9] J. M. C. Clark, *The representation of functionals of Brownian motion by stochastic integrals*, Ann. Math. Statist., 41 (1970), pp. 1282–1295.

[10] R. Cont, *Liquidity, feedback effects and correlation risk*, presentation, 2008.

[11] R. Cont and P. Tankov, *Financial modelling with Jump Processes*, Chapman & Hall - CRC Press, 2003.

[12] F. Coquet, J. Mémin, and L. Słomiński, *On non-continuous Dirichlet processes*, J. Theoret. Probab., 16 (2003), pp. 197–216.

[13] J. Cox and M. Rubinstein, *Option markets*, Prentice-Hall, 1985.

[14] M. H. A. Davis, *On stochastic differentiation*, Teor. Veroyatnost. i Primenen, 20 (1975), pp. 887–892.

[15] M. H. A. Davis, *Martingale methods in stochastic control*, Lecture Notes in Control and Information Systems, 16 (1979), pp. 85–117.

[16] M. H. A. Davis, *Functionals of diffusion processes as stochastic integrals*, Math. Proc. Comb. Phil. Soc., 87 (1980), pp. 157–166.

[17] M. H. A. Davis and P. P. Varaiya, *Dynamic programming conditions for partially observable stochastic systems*, SIAM J. Control and Optim., 11 (1973), pp. 226–261.

[18] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential*, vol. 29 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 1978.

[19] K. Demeterfi, E. Derman, M. Kamal, and J. Zou, *More than you ever wanted to know about volatility swaps*, Journal of Derivatives, 6.

[20] E. Derman and I. Kani, *The volatility smile and its implied tree*, Risk, 7.

[21] J. l. Doob, *Stochastic Processes*, Wiley, 1953.

[22] B. Dupire, *Pricing with a smile*, Risk, 7 (1994), pp. 18–20.

[23] ——, *Functional Itô calculus*, Portfolio Research Paper 2009-04, Bloomberg, 2009.

[24] N. El Karoui, *Les Aspects Probabilistes du Contrôle Stochastique*, Springer Verlag, 1981.

[25] ——, *Couverture des risques dans les marchés financiers*, Lecture notes, (2003).

[26] N. El Karoui and M. Jeanblanc-Picqué, *Sur la robustesse de l'équation de black-scholes*, presentation, 1990.

[27] R. J. Elliott and M. Kohlmann, *A short proof of a martingale representation result*, Statistics & Probability Letters, 6 (1988), pp. 327–329.

[28] W. Fleming and H. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer Verlag, 2005.

[29] 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.

[30] ——, *Dirichlet processes*, in Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), vol. 851 of Lecture Notes in Math., Springer, Berlin, 1981, pp. 476–478.

[31] J. Gatheral, *The Volatility Surface: A Practitioneer's Guide*, Wiley Finance, 2006.

[32] T. Gronwall, *Note on the derivative with respect to a parameter of the solutions of a system of differential equations*, Annals of Mathematics, 20 (1919), pp. 292–296.

[33] U. G. Haussmann, *Functionals of Itô processes as stochastic integrals*, SIAM J. Control Optimization, 16 (1978), pp. 252–269.

[34] ——, *On the integral representation of functionals of Itô processes*, Stochastics, 3 (1979), pp. 17–27.

[35] D. Isaacson, *Stochastic integrals and derivatives*, Ann. Math. Stat., 40 (1969), pp. 1610–1616.

[36] K. Itô, *On a stochastic integral equation*, Proceedings of the Imperial Academy of Tokyo, 20 (1944), pp. 519–524.

[37] ——, *On stochastic differential equations*, Proceedings of the Imperial Academy of Tokyo, 22 (1946), pp. 32–35.

[38] J. Jacod, S. Méléard, and P. Protter, *Explicit form and robustness of martingale representations*, Ann. Probab., 28 (2000), pp. 1747–1780.

[39] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, Springer-Verlag, Berlin, second ed., 2003.

[40] I. Karatzas, D. L. Ocone, and J. Li, *An extension of Clark's formula*, Stochastics Stochastics Rep., 37 (1991), pp. 127–131.

[41] I. Karatzas and S. Shreve, *Brownian motion and Stochastic calculus*, Springer, 1987.

[42] ——, *Methods of Mathematical Finance*, Springer, 1998.

[43] N. Krylov, *Controlled Diffusion Processes*, Springer Verlag, 1980.

[44] 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.

[45] H. Kunita and S. Watanabe, *On square integrable martingales*, Nagoya Math. J., 30 (1967), pp. 209–245.

[46] T. J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana, 14 (1998), pp. 215–310.

[47] T. J. Lyons and T. S. Zhang, *Decomposition of Dirichlet processes and its application*, Ann. Probab., 22 (1994), pp. 494–524.

[48] P. Malliavin, *Stochastic analysis*, Springer, 1997.

[49] P. Meyer, *Un cours sur les integrales stochastiques.* Semin. Probab. X, Univ. Strasbourg 1974/75, Lect. Notes Math. 511, 245-400 (1976)., 1976.

[50] D. Nualart, *Malliavin calculus and its applications*, vol. 110 of CBMS Regional Conference Series in Mathematics, CBMS, Washington, DC, 2009.

[51] D. L. Ocone, *Malliavin's calculus and stochastic integral representations of functionals of diffusion processes*, Stochastics, 12 (1984), pp. 161–185.

[52] E. Pardoux and S. Peng, *Backward stochastic differential equations and quaslinear parabolic partial differential equations*, in Stochastic partial differential equations and their applications, vol. 716 of Lecture Notes in Control and Informatic Science, Springer, 1992.

[53] H. Pham, *On some recent aspects of stochastic control and their applications*, Probability Surveys, 2 (205), pp. 506–549.

[54] P. E. Protter, *Stochastic integration and differential equations*, Springer-Verlag, Berlin, 2005. Second edition.

[55] R. Rishel, *Necessary and suffi cient dynamic programming conditions for continuous time stochastic optimal control*, SIAM J. Control and Optim. 8, 8 (1970), pp. 559–571.

[56] I. Shigekawa, *Derivatives of Wiener functionals and absolute continuity of induced measures*, J. Math. Kyoto Univ., 20 (1980), pp. 263–289.

[57] C. Stricker, *Variation conditionnelle des processus stochastiques*, Ann. Inst. H. Poincaré Probab. Statist., 24 (1988), pp. 295–305.

[58] C. Striebel, *Martingale conditions for the optimal control of continuous time stochastic systems*, Stoch. Proc. and Appl., (1984).

[59] D. W. Stroock, *The Malliavin calculus, a functional analytic approach*, J. Funct. Anal., 44 (1981), pp. 212–257.

[60] 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.

[61] 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.

[62] N. N. Taleb, *Dynamic Hedging - Managing Vanillas and Exotic Options*, Wiley Finance, 1996.

[63] J. Vecer, *A new pde approach for pricing arithmetic average asian options*, Journal of Computational Finance, 4 (2001), pp. 105–113.

[64] 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.

[65] J. Zabczyk, *Remarks on stochastic derivation*, Bull, de L'Acad. Polon. Sci., Ser. Math. Astr. Phys., XXI (1973), pp. 263–269.

Appendix A

Proof of theorems in chapter 2

A.1 Some results on cadlag functions

For a cadlag function $f : [0, T] \mapsto \mathbb{R}^d$ we shall denote $\Delta f(t) = f(t) - f(t-)$ its discontinuity at t.

Lemma A.1. *For any cadlag function $f : [0, T] \mapsto \mathbb{R}^d$*

$$\nabla\epsilon > 0, \quad \exists\eta > 0, \quad |x - y| \leq \eta \Rightarrow |f(x) - f(y)| \leq \epsilon + \sup_{t\in[x,y]} \{|\Delta f(t)|\} \tag{A.1}$$

Proof. Assume the conclusion does not hold. Then there exists a sequence $(x_n, y_n)_{n\geq1}$ such that $x_n \leq y_n$, $y_n - x_n \to 0$ 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 $x_{\psi(n)} \to x$. 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 $(u_n, v_n)_{n\geq1}$ of (x_n, y_n) which converge to x. If $(u_n), (v_n)$ both converge to x from above or from below, $|f(u_n) - f(v_n)| \to 0$ 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 $|f(u_n) - f(v_n)| \to |\Delta f(x)|$, which results in a contradiction as well. Therefore (A.1) must hold. □

The following lemma is a consequence of lemma A.1:

Lemma A.2 (Uniform approximation of cadlag functions by step functions).
Let h be a cadlag function on $[0,T]$. If $(t_k^n)_{n\geq 0, k=0..n}$ is a sequence of subdivisions $0 = t_0^n < t_1 < ... < t_{k_n}^n = t$ of $[0,T]$ such that:

$$\sup_{0\leq i\leq k-1} |t_{i+1}^n - t_i^n| \to_{n\to\infty} 0 \qquad \sup_{u\in[0,T]\setminus\{t_0^n,...,t_{k_n}^n\}} |\Delta f(u)| \to_{n\to\infty} 0$$

then

$$\sup_{u\in[0,T]} |h(u) - \sum_{i=0}^{k_n-1} h(t_i) 1_{[t_i^n, t_{i+1}^n)}(u) + h(t_{k_n}^n) 1_{\{t_{k_n}^n\}}(u)| \to_{n\to\infty} 0 \tag{A.2}$$

A.2 Proof of theorem 2.1

Lemma A.3. *Consider the canonical space $\mathcal{U}_T$ endowed with the natural filtration of the canonical process $X(x,t) = x(t)$. Let $\alpha \in \mathbb{R}$ and σ be an optional time. Then the following functional:*

$$\tau(x) = \inf\{t > \sigma, \quad |x(t) - x(t-)| > \alpha\} \tag{A.3}$$

is a stopping time.

Proof. We can write that:

$$\{\tau(x) \leq t\} = \bigcup_{q\in\mathbb{Q}\bigcap[0,t)} (\{\sigma \leq t - q\} \bigcap \{\sup_{t\in(t-q,t]} |x(u) - x(u-)| > \alpha\} \tag{A.4}$$

and

$$\{\sup_{u\in(t-q,t]} |x(u) - x(u-)| > \alpha\} = \bigcup_{n_0>1} \bigcap_{n>n_0} \{\sup_{1\leq i\leq 2^n} |x(t - q\frac{i-1}{2^n}) - x(t - q\frac{i}{2^n})| > \alpha\} \tag{A.5}$$

thanks to the lemma A.1 in Appendix A.1. □

We can now prove Theorem 2.1 using lemma A.1 from Appendix A.1.

Proof of Theorem 2.1: Let's first prove point 1.; by lemma 4 it implies point 2. for right-continuous functionals and point 3. for left-continuous functionals. Introduce the following random subdivision of $[0,t]$:

$$\tau_0^N(x,v) = 0$$

$$\tau_k^N(x,v) = \inf\{t > \tau_{k-1}^N(x,v) | 2^N t \in \mathbb{N} \text{ or } \quad |v(t) - v(t-)| \vee |x(t) - x(t-)| > \frac{1}{N}\} \wedge t \tag{A.6}$$

From lemma A.3, those functionals are stopping times for the natural filtration of the canonical process. We define the stepwise approximations of x_t and v_t along the subdivision of index N:

$$x^N(s) = \sum_{k=0}^{\infty} x_{\tau_k^N(x,v)} 1_{[\tau_k^N(x,v),\tau_{k+1}^N(x,v)[}(s) + x(t)1_{\{t\}}(s)$$
$$v^N(s) = \sum_{k=0}^{\infty} v_{\tau_k^N(x,v)} 1_{[\tau_k^N(x,v),\tau_{k+1}^N(x,v)[}(t) + v(t)1_{\{t\}}(s) \tag{A.7}$$

as well as their truncations of rank K:

$${}_K x^N(s) = \sum_{k=0}^{K} x_{\tau_k^N} 1_{[\tau_k^N,\tau_{k+1}^N[}(s)$$
$${}_K v^N(t) = \sum_{k=0}^{K} v_{\tau_k^N} 1_{[\tau_k^N,\tau_{k+1}^N[}(t) \tag{A.8}$$

First notice that:

$$F_t(x_t^N, v_t^N) = \lim_{K\to\infty} F_t({}_K x_t^N, {}_K v_t^N) \tag{A.9}$$

because $({}_K x_t^N, {}_K v_t^N)$ coincides with (x_t^N, v_t^N) for K sufficiently large. The truncations

$$F_t^n({}_K x_t^N, {}_K v_t^N)$$

are $\mathcal{F}_t$-measurable as they are continuous functionals of the measurable functions:

$$\{(x(\tau_k^N(x,v)), v(\tau_k^N(x,v))), k \leq K\}$$

so their limit $F_t(x_t^N, v_t^N)$ is also $\mathcal{F}_t$-measurable. Thanks to lemma A.2, x_t^N and v_t^N converge uniformly to x_t and v_t, hence $F_t(x_t^N, v_t^N)$ converges to $F_t(x_t, v_t)$ since F is continuous at fixed times.

Now to show optionality of $Y(t)$ for a left-continuous functional, we will exhibit it as limit of right-continuous adapted processes. For $t \in [0,T]$, define $i^n(t)$ to be the integer such that $t \in [\frac{iT}{n}, \frac{(i+1)T}{n})$. Define the process:

$$Y^n((x,v),t) = F_{\frac{i^n(t)T}{n}}\left(x_{\frac{(i^n(t))T}{n}}, v_{\frac{(i^n(t))T}{n}}\right)$$

, which is piecewise-constant and has right-continuous trajectories, and is also adapted by the first part of the theorem. Now, by d_∞ left-continuity of F, $Y^n(t) \to Y(t)$, which proves that Y is optional.

We similarly prove predictability of $Z(t)$ for a right-continuous functional. We will exhibit it as a limit of left-continuous adapted processes. For $t \in [0,T]$, define $i^n(t)$ to be the integer such that $t \in (\frac{iT}{n}, \frac{(i+1)T}{n}]$. Define the process:

$$Z^n((x,v),t) = F_{\frac{(i^n(t)+1)T}{n}}\left(x_{t-,\frac{(i^n(t)+1)T}{n}-t}, v_{t-,\frac{(i^n(t)+1)T}{n}-t}\right)$$

, which has left-continuous trajectories since as $s \to t-$, $t-s$ sufficiently small, $i^n(s) = i^n(t)$ and $(x_{s-,\frac{(i^n(s)+1)T}{n}-s}, v_{s-,\frac{(i^n(s)+1)T}{n}-s})$ converges to $(x_{t-,\frac{(i^n(t)+1)T}{n}-t}, v_{t-,\frac{(i^n(t)+1)T}{n}-t})$ for d_∞. Moreover, $Z^n(t)$ is $\mathcal{F}_t$-measurable by the first part of the theorem, hence $Z^n(t)$ is predictable. Since $F \in \mathbb{F}_r^\infty$, $Z^n(t) \to Z(t)$, which proves that Y is predictable.

A.3 Measure-theoretic lemmas used in the proof of theorem 2.4 and 2.5

Lemma A.4. *Let f be a bounded left-continuous function defined on $[0,T]$, and let μ_n be a sequence of Radon measures on $[0,T]$ such that μ_n converges vaguely to a Radon measure μ with no atoms. Then for all $0 \le s < t \le T$, with $\mathcal{I}$ being $[s,t]$, $(s,t]$,$[s,t)$ or (s,t):*

$$\lim_n \int_{\mathcal{I}} f(u)d\mu_n(u) = \int_{\mathcal{I}} f(u)d\mu(u) \tag{A.10}$$

Proof. Let M be an upper bound for $|f|$, $F_n(t) = \mu_n([0,t])$ and $F(t) = \mu([0,t])$ the cumulative distribution functions associated to μ_n and μ. For $\epsilon > 0$ and $u \in (s,t]$, define:

$$\eta(u) = \inf\{h > 0 || f(u-h) - f(u)| \ge \epsilon\} \wedge u \tag{A.11}$$

and we have $\eta(u) > 0$ by right-continuity of f. Define similarly $\theta(u)$:

$$\theta(u) = \inf\{h > 0 || f(u-h) - f(u)| \ge \frac{\epsilon}{2}\} \wedge u \tag{A.12}$$

By uniform continuity of F on $[0,T]$ there also exists $\zeta(u)$ such that $\forall v \in [T-\zeta(u),T], F(v+\zeta(u)) - F(v) < \epsilon\eta(u)$. Take a finite covering

$$[s,t] \subset \bigcup_{i=0}^{N}(u_i - \theta(u_i), u_i + \zeta(u_i)) \tag{A.13}$$

where the u_i are in $[s,t]$, and in increasing order, and we can choose that $u_0 = s$ and $u_N = t$. Define the decreasing sequence v_j as follow: $v_0 = t$, and when v_j has been constructed, choose the minimum index $i(j)$ such that $v_j \in (u_{i(j)}, u_{i(j)+1}]$, then either $u_{i(j)} \leq v_j - \eta(v_j)$ and in this case $v_{j+1} = u_{i(j)}$, else $u_{i(j)} > v_j - \eta(v_j)$, and in this case $v_{j+1} = \max(v_j - \eta(v_j), s)$. Stop the procedure when you reach s, and denote M the maximum index of the v_j. Define the following piecewise constant approximation of f on $[s,t]$:

$$g(u) = \sum_{j=0}^{M-1} f(v_j) 1_{(v_{j+1}, v_j]}(u) \tag{A.14}$$

Denote J_1 the set of indices j where v_{j+1} has been constructed as in the first case, and J_2 its complementary. If $j \in J_1$, $|f(u) - g(u)| < \epsilon$ on $[v_j - \eta(v_j), v_j]$, and $v_j - \eta(u_{i(j)})) - v_{j+1} < \zeta(u_{i(j)+1}) = \zeta(v_{j+1})$, because of the remark that $v_j - \eta_{v_j} < u_{i(j)} - \theta(u_{i(j)})$. Hence:

$$\int_{(v_j, v_{j+1}]} |f(u) - g(u)| d\mu(u) \leq \epsilon[F(v_{j+1}) - F(v_j)] + 2M\epsilon\eta(v_{j+1}) \tag{A.15}$$

If $j \in J_2$, $|f(u) - g(u)| < \epsilon$ on $[v_{j+1}, v_j]$. So that summing up all terms we have the following inequality:

$$\int_{[s,t]} |f(u) - g(u)| d\mu(u) \leq \epsilon \left(F(t) - F(s) + 2M(t-s)\right) \tag{A.16}$$

because of the fact that: $\eta(v_j) \leq v_j - v_{j+1}$ for $j < M$. The same argument applied to μ_n yields:

$$\begin{aligned}\int_{[s,t]} |f(u) - g(u)| d\mu_n(u) &\leq \epsilon[F_n(t) - F_n(s-)] \\ &+ 2M \sum_{j=0}^{M-1} F_n(v_{j+1}) - F_n(v_{j+1} - \zeta(v_{j+1}))\end{aligned} \tag{A.17}$$

so that the lim sup satisfies (A.16) since $F_n(u)$ converges to $F(u)$ for every u.

On other hand, it is immediately observed that

$$\lim_n \int_{\mathcal{I}} g(u) d\mu_n(u) = \int_{\mathcal{I}} g(u) d\mu(u) \tag{A.18}$$

since $F_n(u)$ and $F_n(u-)$ both converge to $F(u)$ since μ has no atoms (g is a linear combination of indicators of intervals). So the lemma is established.

□

Lemma A.5. *Let $(f_n)_{n\geq 1}, f$ be left-continuous functions defined on $[0,T]$, satisfying:*

$$\forall t \in [0,T], \lim_n f_n(t) = f(t) \qquad \forall t \in [0,T], f_n(t) \leq K \tag{A.19}$$

Let also μ_n be a sequence of Radon measures on $[0,T]$ such that $\mu(n)$ converges vaguely to a Radon measure μ with no atoms. Then for all $0 \leq s < t \leq T$, with $\mathcal{I}$ being $[s,t]$, $(s,t]$,$[s,t)$ or (s,t):

$$\int_{\mathcal{I}} f_n(u) d\mu_n(u) \to_{n\to\infty} \int_s^t f(u) d\mu(u) \tag{A.20}$$

Proof. Let $\epsilon > 0$ and let n_0 such that $\mu(\{\sup_{m\geq n_0} |f_m - f| > \epsilon\}) < \epsilon$. The set $\{\sup_{m\geq n_0} |f_m - f| > \epsilon\}$ is a countable union of disjoint intervals since the functionals are left-continuous, hence it is a continuity set of μ since μ has no atoms; hence, since μ_n converges vaguely to μ [5]:

$$\lim_n \mu_n(\{\sup_{m\geq n_0} |f_m - f| > \epsilon\}) = \mu(\{\sup_{m\geq n_0} |f_m - f| > \epsilon\}) < \epsilon \tag{A.21}$$

since μ_n converges vaguely to μ which has no atoms.

So we have, for $n \geq n_0$:

$$\int_{\mathcal{I}} |f_n(u) - f(u)| d\mu_n(u) \leq 2K\mu_n \left(\{\sup_{n\geq n_0} |f_n - f| > \epsilon\} \right) + \epsilon\mu_n(\mathcal{I}) \tag{A.22}$$

Hence the lim sup of this quantity is less or equal to:

$$2K\mu(\{\sup_{m\geq n_0} |f_m - f| > \epsilon\} + \epsilon\mu(\mathcal{I}) \leq (2K + \mu(\mathcal{I}))\epsilon \tag{A.23}$$

On other hand:

$$\lim_n \int_{\mathcal{I}} f(u) d\mu_n(u) = \int_{\mathcal{I}} f(u) d\mu(u) \tag{A.24}$$

by application of lemma A.4.

□

Appendix B

Stochastic Differential Equations with functional coefficients

B.1 Stochastic differential equations with path dependent coefficients

B.1.1 Strong solutions

In this section, we will state a theorem providing conditions in which the stochastic differential equations with functional coefficients (B.1) and (B.3) have a unique strong solution. It is a non-markovian counterpart of the standard theorem from the Itô theory [37], (theorem 5.2.9 in [41]). Theorem B.1 can be found in a very similar form in [54], however we include it here with a proof in order to remain self-contained for the reader not familiar with stochastic differential equations with path-dependent coefficients. Let b,σ be functionals on $\bigcup_{t\geq 0} C_0([0,t],\mathbb{R}^d)\times A$, respectively $\mathbb{R}^d$ and $\mathcal{M}^{d,n}$ valued, and ξ a $C_0([0,t_0])$-valued random variable, independent from the Brownian motion W, A a subset of $\mathbb{R}^m$ and α be an A-valued admissible control for the filtration $(\sigma(\xi)\vee\mathcal{B}_t)_{t\geq 0}$, in the sense of definition 3.7. We assume that for any $T\geq 0$, the function: $(t,x,u)\to(b_t(x_t,u),\sigma_t(x_t,u))$ defined on $[0,T]\times C_0([0,T],\mathbb{R}^d)\times A$ is Borel-measurable, which ensures that for any filtration $\mathcal{G}$, for any continous $\mathcal{G}$-adapted process X and any admissible control $\alpha\in\mathcal{A}_{\mathcal{G}}$, the process $(b_t(X_t,\alpha(t)),\sigma_t(X_t,\alpha(t))$ is $\mathcal{G}$-progressively measurable.

Definition B.1. Assume that α is an admissible control for the filtration $\sigma(\xi) \vee \mathcal{B}_t$. A strong solution of the Stochastic differential equation:

$$dX(t_0 + t) = b_{t_0+t}(X_{t_0+t}, \alpha(t))dt + \sigma_{t_0+t}(X_{t_0+t}, \alpha(t))dW(t) \tag{B.1}$$

with initial value

$$X_{t_0} = \xi \tag{B.2}$$

is a continuous process X such that $(X_{t_0+t})_{t \geq t_0}$ is a $\sigma(\xi) \vee \mathcal{B}_t$-adapted continuous semimartingale, and:

1. $X_{t_0} = \xi$ a.s.
2. $\int_0^{t-t_0} [|b_{t_0+s}(X_{t_0+s}, \alpha(s))| + |\sigma_{t_0+s}(X_s, \alpha_s)|^2] ds < +\infty$ a.s. for every $t \geq t_0$
3. $X(t) - X(t_0) = \int_0^{t-t_0} b_{t_0+s}(X_{t_0+s}, \alpha(s)) + \sigma_{t_0+s}(X_{t_0+s}, \alpha(s))dW_s$ a.s. for every $t \geq t_0$

Assume now that you are furthermore given a $C_0([0, t_0], S_d^+)$-random variable χ, such that for all $0 \leq s < t \leq t_0$, the increment $\chi(t) - \chi(s)$ is almost surely in S_d^+; and assume that b,σ be functionals on $\bigcup_{t \geq 0} C_0([0, t], \mathbb{R}^d) \times C_0([0, t], \mathcal{S}_d^+) \times A$, respectively $\mathbb{R}^d$ and $\mathcal{M}^{d,n}$ valued, such that for any $T \geq 0$, the function $(t, x, v, u) \to (b_t(x_t, v_t u), \sigma_t(x_t, v_t, u))$ defined on $[0, T] \times C_0([0, T], \mathbb{R}^d) \times C_0([0, T], S_d^+) \times A$ is Borel-measurable.

Definition B.2. Assume that α is an admissible control for the filtration $\sigma(\xi) \vee \sigma(\chi) \vee \mathcal{B}_t$. A strong solution of the Stochastic differential equation:

$$dX(t_0 + t) = b_{t_0+t}(X_{t_0+t}, [X]_{t_0+t}, \alpha(t))dt + \sigma_{t_0+t}(X_{t_0+t}, [X]_{t_0+t}, \alpha(t))dW(t) \tag{B.3}$$

with inital value

$$X_{t_0} = \xi, [X]_{t_0} = \chi \tag{B.4}$$

is a continuous process X such that $(X_{t_0+t})_{t \geq t_0}$ is a $\sigma(\xi) \vee \sigma(\chi) \vee \mathcal{B}_t$-adapted continuous semimartingale, such that, denoting with a slight abuse of notation:

$$[X](t) = 1_{t \leq t_0} \chi(t) + 1_{t > t_0} \left(\chi(t_0) + [X(t_0 + .) - X(t_0)](t - t_0)\right) \tag{B.5}$$

1. $X_{t_0} = \xi$ a.s.

2. $\int_0^{t-t_0}[|b_{t_0+s}(X_{t_0+s}, [X]_{t_0+s}, \alpha(s))| + |\sigma_{t_0+s}(X_s, [X]_{t_0+s}, \alpha_s)|^2]ds < +\infty$ a.s. for every $t \geq t_0$

3. $X(t) - X(t_0) = \int_0^{t-t_0} b_{t_0+s}(X_{t_0+s}, [X]_{t_0+s}, \alpha(s)) + \sigma_{t_0+s}(X_{t_0+s}, [X]_{t_0+s}, \alpha(s))dW_s$ a.s. for every $t \geq t_0$

Theorem B.1. *In the setting of definition B.1, assume that b and σ satisfy the following Lipschitz and linear growth constraints:*

$$|b_t(x, \alpha) - b_t(y, \alpha)| + |\sigma_t(x, \alpha) - \sigma_t(y, \alpha)| \leq K \sup_{s \leq t} |x(s) - y(s)| \tag{B.6}$$

$$|b_t(x, u)| + |\sigma_t(x, u)| \leq K(1 + \sup_{s \leq t} |x(s)| + |u|) \tag{B.7}$$

for all $t \geq t_0$, $x, y \in C_0([0, t], \mathbb{R}^d)$ and u in A. Then the stochastic differential equation B.1 has a unique strong solution. Moreover, if

$$E[\sup_{s \in [0,t_0]} |\xi(s)|^2] < \infty \tag{B.8}$$

then, for every $T < 0$, there exists a constant C depending on T, K and α only such that:

$$\forall t \in [t_0, T], E[\sup_{s \in [0,t]} |X(s)|^2] \leq C(1 + E[\sup_{s \in [0,t_0]} |\xi(s)|^2])e^{C(t-t_0)} \tag{B.9}$$

The proof follows the methodology used to prove theorem 5.2.9 in [41].

Proof. Suppose first that the condition $E[\sup_{s \in [0,t_0]} |\xi(s)|^2] < \infty$ holds. Define the following sequence of processes, starting with $X^0(t) = \xi(t)1_{t \leq t_0} + \xi(t_0)1_{t > t_0}$:

$$\begin{aligned} X_t^{n+1} = \xi(t)1_{t \leq t_0} + [\xi(t_0) + \int_0^{t-t_0} & b_{t_0+s}(X_{t_0+s}^n, \alpha(s))ds \\ + \int_0^{t-t_0} & \sigma_{t_0+s}(X_{t_0+s}^n, \alpha(s))dW(s)]1_{t > t_0} \end{aligned} \tag{B.10}$$

We will first prove by induction that for all $T \geq t_0$, there exists a constant C depending only on K,T, and α such that

$$\forall n \geq 0 \forall t \in [t_0, T], E[\sup_{s \in [0,t]} |X^n(s)|^2] \leq C(1 + E[\sup_{s \in [0,t_0]} |\xi(s)|^2])e^{C(t-t_0)} \tag{B.11}$$

The property is obvious for $n = 0$ for any C, assume it true for n. Define the processes:

$$B(s) = \int_0^s b_{t_0+u}(X^n_{t_0+u}, \alpha(u))ds, M(s) = \int_0^s \sigma_{t_0+u}(X^n_{t_0+u}, \alpha(u))dW(u) \tag{B.12}$$

and define the constant $L = E[\int_0^{T-t_0} |\alpha(s)|^2 ds]$ Then:

$$|B(s)|^2 \leq 2sK^2 \int_0^s [1 + (\sup_{v\in[0,u]} |X^n(t_0+v)|^2 + \sup_{v\in[0,t_0]} |\xi(v)|^2) + |\alpha(u)|^2]du \tag{B.13}$$

(using Cauchy-Schwarz inequality), so that:

$$\begin{aligned} E[\sup_{s\in[0,t-t_0]} |B(s)|^2] \leq 2TK^2[T + L + E[\sup_{v\in[0,t_0]} |\xi(v)|^2]] \\ +2KT^2(1 + E[\sup_{v\in[0,t_0]} |\xi(v)|^2])e^{C(t-t_0)} \end{aligned} \tag{B.14}$$

On other hand:

$$[M](t-t_0) \leq 2K \int_0^{t-t_0} [1 + (\sup_{v\in[0,u]} |X^n(t_0+v)|^2 + \sup_{v\in[0,t_0]} |\xi(v)|^2) + |\alpha(u)|^2]du \tag{B.15}$$

so using Brukholder-Davis-Gundy Inequalities (Theorem 3.3.28 in [41] in dimension 1 and Problem 3.3.29 for multidimensional case), there exists a universal constant Λ such that:

$$\begin{aligned} E[\sup_{s\in[0,t-t_0]} |M(s)|^2] \leq 2K^2\Lambda[T + L + E[\sup_{v\in[0,t_0]} |\xi(v)|^2]] \\ +2K^2\Lambda(1 + E[\sup_{v\in[0,t_0]} |\xi(v)|^2])e^{C(t-t_0)} \end{aligned} \tag{B.16}$$

And hence finally

$$\begin{aligned} E[\sup_{s\in[t_0,t]} |X^{n+1}(s)|^2 ds] &\leq 2[E[\sup_{v\in[0,t_0]} |\xi(v)|^2] + E[\sup_{s\in[0,t-t_0]} |M(s)|^2] + E[\sup_{s\in[0,t-t_0]} |B(s)|^2]] \\ &\leq 2K^2(\Lambda + T)(T + L + E[\sup_{v\in[0,t_0]} |\xi(v)|^2]) \\ &+2K^2(\Lambda + T)(1 + E[\sup_{v\in[0,t_0]} |\xi(v)|^2])e^{C(t-t_0)} \end{aligned}$$

Hence choosing $C = 2K^2(\Lambda+T)(1+T+L+E[\sup_{v\in[0,t_0]} |\xi(v)|^2])$ ensures that the inequality passes by induction. We now define the processes:

$$\begin{aligned} B(s) &= \int_0^s [b_{t_0+u}(X^{n+1}_{t_0+u}, \alpha(u)) - b_{t_0+u}(X^n_{t_0+u}, \alpha(u))]ds \\ M(s) &= \int_0^s [\sigma_{t_0+u}(X^{n+1}_{t_0+u}, \alpha(u)) - \sigma_{t_0+u}(X^n_{t_0+u}, \alpha(u))]dW(u) \end{aligned} \tag{B.17}$$

It is obvious by Brukholder-Davis-Gundy Inequalities that:

$$E[\sup_{s\in[0,t-t_0]}|M(s)|^2] \leq \Lambda K^2 \int_0^{t-t_0} E[\sup_{u\in[0,s]}|X^n(u)-X^{n-1}(u)|^2]ds \tag{B.18}$$

and using Cauchy-Schwarz inequality:

$$E[\sup_{s\in[0,t-t_0]}|B(s)|^2] \leq tK^2 \int_0^{t-t_0} E[\sup_{u\in[0,s]}|X^n(u)-X^{n-1}(u)|^2]ds \tag{B.19}$$

so that:

$$\forall t\leq T, E[\sup_{s\in[0,t-t_0]}|X^{n+1}(s)-X^n(s)|^2] \leq$$
$$2K^2(\Lambda+T)\int_0^{t-t_0} E[\sup_{u\in[0,s]}|X^n(u)-X^{n-1}(u)|^2]ds \tag{B.20}$$

so that reiterating ensures that:

$$\forall t\leq T, \qquad E[\sup_{s\in[0,t-t_0]}|X^{n+1}(s)-X^n(s)|^2] \leq C^*\frac{[2K^2(\Lambda+T)]^n t^n}{n!} \tag{B.21}$$

with $C* = E[\sup_{t\in[t_0,T]}|X^1(t)-X^0(t)|^2] < \infty$ thanks to B.11. Chebychev inequality now ensures that:

$$\mathbb{P}[\sup_{t\in[t_0,T]}|X^{n+1}(t)-X^n(t)| > \frac{1}{2^{n+1}}] \leq 4C^*\frac{8K^2(\Lambda+T)T]^n}{n!} \tag{B.22}$$

Hence by Borel-Cantelly lemma almost surely there exists n_0 such that

$$n\geq n_0 \Rightarrow \sup_{t\in[t_0,T]}|X^{n+1}(t)-X^n(t)|^2 \leq \frac{1}{2^{n+1}} \tag{B.23}$$

. Hence there exists a continuous process X such that almost surely $X^n \to X$ uniformly on compact intervals, and inequality (B.11) passes to the limit by Fatou's lemma. Moreover, inequality (B.11) together with the linear growth condition on b and σ allows to pass to the limit in the Lebesgue and stochastic integrals by dominated convergence, so the limit X satisfies the stochastic differential equation.

We now forget the assumption $E[\sup_{s\in[0,t_0]}|\xi(s)|^2] < \infty$, and we will show the uniqueness of the solution. Assume that X and Y are two solutions, let $\tau_N = \inf\{t\geq t_0 | |X(s)|\vee|Y(s)| \geq N\}$. The previous methodology immediately proves that:

$$\forall t\leq T E[\sup_{s\in[0,t-t_0]}|X^{\tau_N}(s)-Y^{\tau_N}(s)|^2] \leq C'\int_0^{t-t_0} E[\sup_{u\in[0,s]}|X^{\tau_N}(u)-Y^{\tau_N}(u)|^2]ds \tag{B.24}$$

for some constant C', which can be reiterated to prove that for any n:

$$\forall t \leq T, E[\sup_{s\in[0,t-t_0]} |X^{\tau_N}(s) - Y^{\tau_N}(s)|^2] \leq \frac{[C'(t-t_0)]^n}{n!} \int_0^{t-t_0} E[\sup_{u\in[0,s]} |X^{\tau_N}(u) - Y^{\tau_N}(u)|^2] ds \tag{B.25}$$

Since this is true for any n, then $E[\sup_{s\in[0,T-t_0]} |X^{\tau_N}(s) - Y^{\tau_N}(s)|^2] = 0$ By Fatou's lemma, we can take the limit $N \to +\infty$ to obtain:

$$E[\sup_{s\in[0,T-t_0]} |X(s) - Y(s)|^2] = 0 \tag{B.26}$$

Which proves the uniqueness of the solution.

We will finally show existence in the general case. Note that the event $\{\sup_{s\in[0,t_0]} \xi \leq N\}$ belongs to $\sigma(\xi)$, and denote for $N \geq 1$ X^N the solution with initial value $\xi 1_{\sup_{s\in[0,t_0]} \xi \leq N}$, and X^0 the solution whose initial value is the identically 0 trajectory. Since for $M < N$, $X^N 1_{\sup_{s\in[0,t_0]} \xi \leq M} + X^0 1_{\sup_{s\in[0,t_0]} \xi > M}$ is solution with initial value $\xi 1_{\sup_{s\in[0,t_0]} \xi \leq M}$, and since uniqueness has been established, then $X^N 1_{\sup_{s\in[0,t_0]} \xi \leq M} = X^M 1_{\sup_{s\in[0,t_0]} \xi \leq M}$. Hence almost surely the sequence X^N is constant from a given rank, hence it has a limit X which is solution of the SDE with initial value ξ.

□

Corollary B.1. *In the setting of definition B.2, assume that b and σ satisfy the following Lipschitz and linear growth constraints:*

$$|b_t(x,v,\alpha) - b_t(y,w,\alpha)| + |\sigma_t(x,v,\alpha) - \sigma_t(y,w,\alpha)| \leq K \sup_{s\leq t} |x(s) - y(s)| + |v(s) - w(s)| \tag{B.27}$$

$$|b_t(x,v,u)| \leq K(1 + \sup_{s\leq t} |x(s)| + \sup_{s\leq t} |v(s)| + |u|) \tag{B.28}$$

$$|\sigma_t(x,v,u)| \leq K \tag{B.29}$$

for all $t \geq t_0$, $x, y \in C_0([0,t], \mathbb{R}^d)$, $v, w \in C_0([0,t], \mathcal{S}_d^+)$ and $u \in A$, then there exists a unique strong solution to the equation B.3. with initial value

$$X_{t_0} = \xi, [X]_{t_0} = \chi \tag{B.30}$$

If moreover

$$E[\sup_{s\in[0,t_0]} |\xi(s)|^2 + |\chi(t_0)|^2] < \infty \tag{B.31}$$

, then for every $T < 0$, there exists a constant C depending on T, K and α only such that:

$$\forall t \in [t_0, T], E[\sup_{s\in[0,t]} |X(s)|^2] \le C(1 + E[\sup_{s\in[0,t_0]} |\xi(s)|^2] + E[|\chi(t_0)|^2])e^{C(t-t_0)} \tag{B.32}$$

Proof. Note that the fact that σ is bounded and Lipschitz ensures that ${}^t\sigma\sigma$ is Lipschitz. Theorem B.1 ensures that the following $d(d+1)$-dimensional SDE has a unique strong solution:

$$\begin{aligned} dX(t_0+t) &= b_{t_0+t}(X_{t_0+t}, V_{t_0+t}, \alpha(t))dt + \sigma_{t_0+t}(X_{t_0+t}, V_{t_0+t}, \alpha(t))dW(t) \\ dV(t_0+t) &= {}^t\sigma_{t_0+t}(X_{t_0+t}, V_{t_0+t}, \alpha(t))\sigma_{t_0+t}(X_{t_0+t}, V_{t_0+t}, \alpha(t))dt \end{aligned} \tag{B.33}$$

with initial value

$$(X_{t_0}, V_{t_0}) = (\xi, \chi) \tag{B.34}$$

□

B.1.2 Continuity in the initial value

We can furthermore state continuity of the solutions in the initial value, in the sense of the following theorem and corollary:

Theorem B.2. *Let b, σ be as in theorem B.1, and let ξ and ξ' be two $C_0([0,t_0],\mathbb{R}^d)$ -valued random variable, independent from the Brownian motion W, satisfying the assumption* (B.8)*, and α be an admissible control. Denote X^ξ and $X^{\xi'}$ the solutions of* (B.1) *with respective initial values ξ and ξ'. Then, for every $T \ge t_0$ and every $\epsilon > 0$, there exists a constant C depending only on ϵ, T and K in assumption* (B.6)*, such that:*

$$\forall t_0 \le t \le T, E[\sup_{0\le s\le t} |X^\xi(s) - X^{\xi'}(s)|^2] \le E[\sup_{s\le t_0} |\xi(s) - \xi'(s)|^2 + \epsilon|\xi(t_0) - \xi'(t_0)|^2]e^{C(t-t_0)}$$

Proof. For $s \le T - t_0$, let $f(s) = E[\sup_{0\le u\le t_0+s} |X^\xi(u) - X^{\xi'}(u)|^2]$. Obviously $f(0) = E[\sup_{s\le t_0} |\xi(s) - \xi'(s)|^2]$. Define the processes:

$$\begin{aligned} B(s) &= \int_0^s [b_{t_0+u}(X^\xi_{t_0+u}, \alpha(u)) - b_{t_0+u}(X^{\xi'}_{t_0+u}, \alpha(u))]ds \\ M(s) &= \int_0^s [\sigma_{t_0+u}(X^\xi_{t_0+u}, \alpha(u)) - \sigma_{t_0+u}(X^{\xi'}_{t_0+u}, \alpha(u))]dW(u) \end{aligned} \tag{B.35}$$

Then using the global Lipschitz assumption (B.6) and Cauchy-Schwarz inequality:

$$E[\sup_{u \le s} |B(u)|^2] \le K^2 T \int_0^s f(u) du \tag{B.36}$$

Using the global Lipschitz assumption (B.6):

$$E[[M](s)] \le K^2 \int_0^s f(u) du \tag{B.37}$$

so that using Brukholder-Davis-Gundy Inequalities there exists a universal constant Λ such that:

$$E[\sup_{u \le s} |M(u)|^2] \le \Lambda K^2 \int_0^s f(u) du \tag{B.38}$$

so that finally:

$$E[\sup_{u \le s} (|M(u)| + |B(u)|)^2] \le 2K^2(\Lambda + T) \int_0^s f(u) du \tag{B.39}$$

Note that:

$$|X^{\xi}(t_0 + u) - X^{\xi'}(t^u)|^2 \le (1 + \epsilon)(\xi(t_0) - \xi'(t_0))^2 + (1 + \frac{1}{\epsilon})(M(u) + B(u))^2 \tag{B.40}$$

so that:

$$f(s) \le E[\sup_{s \le t_0} |\xi(s) - \xi'(s)|^2 + \epsilon |\xi_{t_0} - \xi'(t_0)|^2] + 2(1 + \frac{1}{\epsilon})K^2(\Lambda + T) \int_0^s f(u) ds \tag{B.41}$$

Hence applying Gronwall lemma [32] with $C = 2(1 + \frac{1}{\epsilon})K^2(\Lambda + T)$ concludes the proof.

□

Corollary B.2. *Let b, σ are as in corollary B.1, and (ξ, χ) and (ξ', χ') be two initial values as in corollary B.1 satisfying* (B.31)*, and let α be an admissible control. Then, denoting $X^{\xi,\chi}, X^{\xi',\chi'}$ the strong solutions of the SDE* (B.3) *with respective initial values (ξ, χ),(ξ', χ'), then, for every $T \ge t_0$, and every $\epsilon > 0$, there exists a constant C depending only on ϵ, T and K in assumptions* (B.27), (B.29) *such that:*

$$\forall t_0 \le t \le T, \qquad E[\sup_{0 \le s \le t} |X^{\xi,\chi}(s) - X^{\xi',\chi'}(s)|^2] \le$$
$$E[\sup_{s \le t_0} |\xi(s) - \xi'(s)|^2 + |\chi(s) - \chi'(s)|^2 + \epsilon |\xi(t_0) - \xi'(t_0)|^2] e^{C(t - t_0)}$$

B.1.3 Perturbation of coefficients

A probabilistic analysis of perturbations of the coefficient of the SDE is required to allow us to prove theorem 5.3, which is important for the sensitivity analysis of path-dependent derivatives. We state the following theorem and its corollary:

Theorem B.3. *Let b, σ and ξ be as in theorem B.1, with ξ satisfying the assumption* (B.8)*, and α be an admissible control. Let $(b^\epsilon)_{\epsilon>0}$ and $(\sigma^\epsilon)_{\epsilon>0}$ be families of functionals on $\bigcup_{t\geq 0} C_0([0,t],\mathbb{R}^d)\times A$, respectively $\mathbb{R}^d$ and $\mathcal{M}^{d,n}$ valued, satisfying the usual measurability assumption. Assume that there exists $\phi:\mathbb{R}^+\mapsto\mathbb{R}^+$, such that for any $t\leq T$ and any $x\in C_0([0,t],\mathbb{R}^d)$, $|b_t(x_t)-b^\epsilon_t(x_t)|+|\sigma_t(x_t)-\sigma^\epsilon_t(x_t)|<\sup_{s\in[0,t]}|x(s)|\phi(\epsilon)$. Denote X the solution of the SDE* (B.1) *with coefficients b, σ and initial value ξ, and let $(X^\epsilon)_{\epsilon>0}$ be a family of square-integrable processes satisfying the SDE with coefficients b^ϵ, σ^ϵ and initial value ξ, and such that for any $T>t_0$, there exists constants $A_T>0,\epsilon_T>0,\forall\epsilon<\epsilon_T,\forall t_0\leq t\leq T, E[|X^\epsilon(t)|^2]<A_T$. Then for any $T>t_0$, there exists a constant C depending only on K, T and A such that, for any $t\leq T,\epsilon<\epsilon_T$:*

$$E[\sup_{s\in[0,t-t_0]}|X(s)-X^\epsilon(s)|^2]\leq C\phi^2(\epsilon)\left(e^{C(t-t_0)}-1\right) \tag{B.42}$$

Proof. Let $T\geq t_0+s\geq t_0$.

$$\begin{aligned}X(t_0+s)-X^\epsilon(t_0+s)=&\int_0^s[b_{t_0+u}(X_u,\alpha_u)-b^\epsilon_{t_0+u}(X^\epsilon_u,\alpha_u)]du\\&+\int_0^s[\sigma_{t_0+u}(X_u,\alpha_u)-\sigma^\epsilon_{t_0+u}(X^\epsilon_u,\alpha_u)]dW(u)\end{aligned} \tag{B.43}$$

We have therefore:

$$|X(t_0+s)-X^\epsilon(t_0+s)|\leq|B(s)|+|M(s)|+|R(s)| \tag{B.44}$$

where

$$\begin{aligned}B(s)&=\int_0^s[b_{t_0+u}(X^{t_0+u},\alpha(u))-b_{t_0+u}(X^\epsilon_{t_0+u},\alpha(u))]ds\\M(s)&=\int_0^s[\sigma_{t_0+u}(X_{t_0+u},\alpha(u))-\sigma_{t_0+u}(X^\epsilon_{t_0+u},\alpha(u))]dW(u)\\R(s)&=\int_0^s[b_{t_0+u}(X^\epsilon_{t_0+u},\alpha(u))-b^\epsilon_{t_0+u}(X^\epsilon_{t_0+u},\alpha(u))]du\\&+\int_0^s[\sigma_{t_0+u}(X^\epsilon_{t_0+u},\alpha(u))-\sigma^\epsilon_{t_0+u}(X^\epsilon_{t_0+u},\alpha(u))]dWu\end{aligned} \tag{B.45}$$

Then using the global Lipschitz assumption (B.6) and Cauchy-Schwarz inequality:

$$E[\sup_{u\leq s}|B(u)|^2] \leq K^2 T \int_0^s \sup_{v\in[0,u]} |X(t_0+v) - X^\epsilon(t_0+v)|^2 du \tag{B.46}$$

Using the global Lipschitz assumption (B.6):

$$E[[M](s)] \leq K^2 \int_0^s \sup_{v\in[0,u]} |X(t_0+v) - X^\epsilon(t_0+v)|^2 du \tag{B.47}$$

so that using Brukholder-Davis-Gundy Inequalities there exists a universal constant Λ such that:

$$E[\sup_{u\leq s}|M(u)|^2] \leq \Lambda K^2 \int_0^s \sup_{v\in[0,u]} |X(t_0+v) - X^\epsilon(t_0+v)|^2 du \tag{B.48}$$

Using Cauchy-Schwarz inequality:

$$E[\sup_{u\leq s} R^2(u)] \leq 4s\phi^2(\epsilon) A_T^2 \tag{B.49}$$

so that finally:

$$E[\sup_{u\leq s}(|M(u)| + |B(u)|)^2] \leq 2K^2(\Lambda + T) \int_0^s \sup_{v\in[0,u]} |X(t_0+v) - X^\epsilon(t_0+v)|^2 du \tag{B.50}$$

and hence:

$$E[\sup_{u\leq s}|X(t_0+s) - X^\epsilon(t_0+s)|^2] \leq 8A_T^2 s\phi^2(\epsilon)$$
$$+2K^2(\Lambda+T)\int_0^s sup_{v\in[0,u]}|X(t_0+v) - X^\epsilon(t_0+v)|^2 du \tag{B.51}$$

So Gronwall lemma [32] concludes the proof.

□

Corollary B.3. *Let b, σ and (ξ,χ) be as in corollary B.1, with ξ,χ satisfying the assumption* (B.8)*, and α be an admissible control. Let $(b^\epsilon)_{\epsilon>0}$ and $(\sigma^\epsilon)_{\epsilon>0}$ be families of functionals on $\bigcup_{t\geq 0} C_0([0,t],\mathbb{R}^d) \times A$, respectively $\mathbb{R}^d$ and $\mathcal{M}^{d,n}$ valued, satisfying the usual measurability assumption. Assume that there exists $\phi : \mathbb{R}^+ \mapsto \mathbb{R}^+$, such that for any $t \leq T$ and any $x \in C_0([0,t],\mathbb{R}^d), v \in C_0([0,t], S_d^+)$ $|b_t(x_t,v_t) - b_t^\epsilon(x_t,v_t)| + |\sigma_t(x_t,v_t) - \sigma_t^\epsilon(x_t,v_t)| < \phi(\epsilon)$. Denote X the solution of the SDE* (B.1) *with coefficients b, σ and initial value (ξ,χ), and let $(X^\epsilon)_{\epsilon>0}$ be a family of square-integrable processes satisfying the SDE with coefficients*

b^ϵ, σ^ϵ and initial value (ξ, χ), and such that for any $T > t_0$, there exists constants $A_T > 0, \epsilon_T > 0, \forall \epsilon < \epsilon_T$, such that $\forall t_0 \leq t \leq T, E[|X^\epsilon(t)|^2 + [|[X^\epsilon](t)|^2] < A_T$. Then for any $T > t_0$, there exists a constant C depending only on K, T and A_T such that, for any $t \leq T$:

$$E[\sup_{s \in [0, t-t_0]} |X(s) - X^\epsilon(s)|^2 + |[X](s) - [X^\epsilon](s)|^2] \leq c\phi^2(\epsilon)\left(e^{C(t-t_0)} - 1\right) \tag{B.52}$$