Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Functional Itô calculus

Bruno Dupire

To cite this article: Bruno Dupire (2019) Functional Itô calculus, Quantitative Finance, 19:5, 721-729, DOI: 10.1080/14697688.2019.1575974

To link to this article: https://doi.org/10.1080/14697688.2019.1575974

Published online: 17 Apr 2019.

Quantitative Finance, 2019
Vol. 19, No. 5, 721–729, https://doi.org/10.1080/14697688.2019.1575974

Functional Itô calculus†

BRUNO DUPIRE*

Bloomberg L.P., New York, NY, USA

(*Received 25 January 2019; accepted 25 January 2019; published online 29 March 2019*)

We extend some results of the Itô calculus to functionals of the current path of a process to reflect the fact that often the impact of randomness is cumulative and depends on the history of the process, not merely on its current value. We express the differential of the functional in terms of adequately defined partial derivatives to obtain an Itô formula. We develop an extension of the Feynman-Kac formula to the functional case and an explicit expression of the integrand in the Martingale Representation Theorem. We establish that under certain conditions, even path dependent options prices satisfy a partial differential equation in a local sense. We exploit this fact to find an expression of the price difference between two models and compute variational derivatives with respect to the volatility surface.

Keywords: Path dependent functionals and options; Functional Itô and Feynman-Kac formulae; Martingale representation; Functional PDE; Delta hedging; Model impact; Submartingale bounds

1. Introduction

1.1. Motivation

In probability theory the unfolding of uncertainty is modeled by stochastic processes and functions of the current state can be manipulated by the Itô calculus. However, in many cases, uncertainty affects the current situation not only through the current state of the process but also through its whole history. For instance, the quality of a harvest does not only depend on the current temperature, but also on the whole pattern of past temperatures; the price of a path dependent option may depend on the whole history of the underlying price; the intensity of an immune response depends on the full history of exposure to antigens derived from microbial infection. The objective of this paper is to lay the foundations of a theory that extends Itô Calculus to functionals of Itô processes.

It connects several well known results (Clark-Ocone formula, max as local time of drawdown) and establishes new ones (BS PDE for path dependent options Black and Scholes 1973) which open the door to many applications.

1.2. Review of usual approaches

In essence, Itô's formula tells us that a smooth function of a semi-martingale is a semi-martingale. The novelty with respect to deterministic differential calculus is the presence of a second order term in the Taylor expansion, due to the non zero quadratic variation of the process. The simplest version is for one dimensional processes and it has extensions for n-dimensional or even infinite dimensional (Da Prato and Zabczyk 1992) processes, but it always considers functions of the current value of the process, as opposed to functions of the current history of the process, which are named functionals.

Malliavin Calculus can be seen as a differential calculus for functionals that are defined on a fixed time interval. It is a variational calculus on the Wiener space. The Malliavin derivative consists in shocking a path at a given time, which impacts its future behavior and any functional.

The object of this paper is to define a differential calculus on functions of the current path of the process, with no reference to a final maturity. It applies to 'running functionals', at every step of their evolution.

1.3. Related literature

The approach presented here seems to be new. It is possible to establish some links with Malliavin calculus but the functional calculus introduced in this paper deals with functionals of the current path, as opposed to functionals of the whole

*Corresponding author. Email: bdupire@bloomberg.net
†The Editors of Quantitative Finance are delighted that Bruno Dupire has accepted their invitation to publish his seminal paper on functional Itô calculus in honour of his 60th birthday. Well-known to both academics and practitioners, Bruno's 2009 paper on SSRN continues to be highly-cited, with new applications in mathematical finance and the theory of stochastic optimal control appearing regularly ever since.

path in the Malliavin case. Neither the Malliavin calculus nor this functional calculus encompasses the other one. The closest related approach we are aware of (thanks to Protter 2005 for pointing to it) is the work of Ahn (1997). He has an Itô formula (Th 2.1) that applies in the very specific case of functionals which are the intrinsic value of a claim (which assumes the process is frozen until the maturity of the claim). He has a PDE (Th 4.1) that assumes the knowledge of the whole path instead of the knowledge of the current path in our case. Eventually, he has (Th 4.2/4.3) Clark-Ocone type representations, but does not treat the case of path dependent dynamics.

Section 2 introduces notations and definitions, Section 3 contains the main results, Section 4 presents a few applications and Section 5 concludes.

2. Notation and definitions

To set the stage, we need to provide definitions of the path space, functionals, distance, continuity and differentiability. As we are working with functionals that are defined on paths of varied lengths, we need some uncommon definitions.

2.1. Space

The paths we consider are cádlág (right continuous with left limits) as opposed to continuous as we will apply some discontinuous shocks to compute some derivatives. It should be understood that this is a restriction on the class of functionals, not an enlargement of the class of processes we will consider.

We denote by Λ_t the set of cadlag functions $[0,t] \longrightarrow \mathbb{R}$ and define $\Lambda \equiv \bigcup_{t\in[0,\overline{T}]} \Lambda_t$ for a given $\overline{T}$. We note that Λ is not a vector space as the notion of addition is not defined.

Paths are denoted by capital letters and processes by lower case letters. For instance, if x is a process, its value at time t is x_t and its path up to time t is $X_t \in \Lambda_t \subset \Lambda$ and for $u \in [0,t], X_t(u) = x_u$.

2.2. Functionals

Definition: A *functional* is a function $f : \Lambda \longrightarrow \mathbb{R}$. It associates a real number to any càdlàg function over $[0,t]$ for any $t \in [0,\overline{T}]$.

Among many possible examples, let us mention the following ones:

- Exponential smoothing
- Bollinger band
- Discrete quadratic variation
- Prepayment of MBS as a function of rates
- Maximum
- Maximum of a rolling average
- Profit and loss of a technical analysis strategy
- First time the current value has been reached
- Running range, drawdown
- Rolling/moving average/Bollinger band
- Conditional expectation of an Asian option
- (Discrete) stochastic integral
- Lower bound of a claim.

Two very significant financial examples are:

- price of a path dependent claim as a function of the path so far (conditional expectation under a model)
- lower/upper bound of a path dependent claim as a function of the path so far (infimum/supremum of the conditional expectations over a class of models).

2.3. Topology and continuity

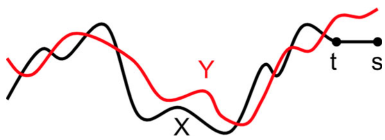
We introduce a distance on Λ together with the induced notion of continuity of a functional.

The set Λ is comprised of paths of various lengths and some care must be taken in defining a distance between two paths of different lengths. We extend the shortest path, freezing its last value, to the length of the longest one, compute the distance in the sense of the max norm and add the difference of the lengths. We thus define a distance on Λ by:

For all X_t, Y_s in Λ (we assume $t \leq s$),

$$d_\Lambda(X_t, Y_s) \equiv \|X_{t,s-t} - Y_s\|_\infty + s - t$$

with $X_{t,s-t}$ the flat extension of X_t, as defined in the next section.



The presence of $s-t$ ensures that d_Λ is a distance, not simply a semi-distance.

This distance induces the associated notion of continuity:

A functional $f : \Lambda \longrightarrow \mathbb{R}$ is Λ - *continuous at* $X_t \in \Lambda$ if:

$$\forall \varepsilon > 0, \exists \alpha > 0 : \forall Y_s \in \Lambda,$$
$$d_\Lambda(X_t, Y_s) < \alpha \Rightarrow |f(X_t) - f(Y_s)| < \varepsilon.$$

$f : \Lambda \longrightarrow \mathbb{R}$ is Λ-*continuous* if it is Λ- continuous at all $X_t \in \Lambda_t$.

2.4. Space and time derivatives

We now introduce the space and time derivatives of a functional. For a given current path $X_t \in \Lambda_t$, the derivatives correspond to changes in the current value of the process and in the current time, as we respect the fact that we cannot alter the past and ignore the future. This is in contrast with the Malliavin derivative, which applies shocks at any point of a full path.

For $X_t \in \Lambda_t$, we define:

(a) X_t^h as X_t with the endpoint shifted by $h \in \mathbb{R}$:

$$X_t^h(s) = X_t(s) \quad \text{for } s < t \quad X_t^h(t) = X_t(t) + h$$

(b) $X_{t,\delta t}$ with $\delta t \geq 0$ is an extension of X_t, obtained by freezing the endpoint over $[t, t+\delta t]$:

$$X_{t,\delta t}(s) = X_t(s) \quad \text{for } s \leq t \quad X_{t,\delta t}(s) = X_t(t) \quad \text{for } s \in [t, t+\delta t].$$

For a functional $f : \Lambda \longrightarrow \mathbb{R}, X_t \in \Lambda_t$ and define when they exist

$$\Delta_x f(X_t) \equiv \lim_{h\to 0} \frac{f(X_t^h) - f(X_t)}{h}$$

$$\Delta_{xx} f(X_t) \equiv \lim_{h\to 0} \frac{\Delta_x f(X_t^h) - \Delta_x f(X_t)}{h}$$

$$\Delta_t f(X_t) \equiv \lim_{\delta t\to 0^+} \frac{f(X_{t,\delta t}) - f(X_t)}{\delta t}.$$

The increment in the space derivative definition can be positive or negative whilst the increment in the time derivative definition is only positive. It is hence a right-derivative.

We give below a few examples of functionals and their derivatives.

	$f(X_t)$	x_t	$\int_0^t x_u du$	QV_t
	$\Delta_x f$	1	0	$2(x_t - x_{t^-})$
	$\Delta_t f$	0	x_t	0

If $f(X_t) = h(x_t, t)$, then $\begin{cases} \Delta_x f = \frac{\partial h}{\partial x} \\ \Delta_t f = \frac{\partial h}{\partial t} \end{cases}$.

These derivatives satisfy the classical properties: linearity, product, chain rule.

We introduce a last definition: $f : \Lambda \longrightarrow \mathbb{R}$ is a *smooth functional* if it is Λ - continuous, C^2 in x and C^1 in t, with these derivatives themselves Λ - continuous.

So far no probability concepts have been invoked. We now need to introduce them. The stochastic basis is a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t\in[0,\overline{T}]}$ satisfying the usual conditions.

2.5. *Stochastic integral*

For g a functional from $\Lambda \longrightarrow \mathbb{R}$, if $\lim_{t_i - t_{i-1}\to 0} \sum_i g(X_{t_{i-1}})(x_{t_i} - x_{t_{i-1}})$ exists, we denote it by $\int g(X_t)\,\mathrm{d}x_t$, to be understood as evaluated at X_{t^-}.

3. Functional Itô calculus

These definitions being established, we can proceed with the main results.

3.1. *Functional Itô formula*

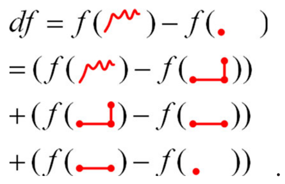
This is the extension of the Itô formula to functionals:

THEOREM 3.1 *If x is a continuous semi-martingale and X_t denotes its path over $[0, t]$ and $f : \Lambda \longrightarrow \mathbb{R}$ is a smooth functional (Λ - continuous, C^2 in x and C^1 in t, with these derivatives themselves Λ - continuous), then, for all $T \geq 0$,*

$$f(X_T) = f(X_0) + \int_0^T \Delta_x f(X_t)\,\mathrm{d}x_t + \int_0^T \Delta_t f(X_t)\,\mathrm{d}t + \frac{1}{2}\int_0^T \Delta_{xx} f(X_t)\,\mathrm{d}\langle x\rangle_t$$

Proof See Appendix. ■

The proof is quite involved but it essentially relies on the following schematic decomposition of the increment of the functional over a small time interval:



The first of the 3 differences goes to 0, the second one gives the first and second space derivative whilst the third one gives the time derivative.

REMARK 1 A more concise notation is

$$\mathrm{d}f(X_T) = \Delta_x f(X_t)\,\mathrm{d}x_t + \Delta_t f(X_t)\,\mathrm{d}t + \tfrac{1}{2}\Delta_{xx} f(X_t)\,\mathrm{d}\langle x\rangle_t$$

or even

$$\mathrm{d}f = \Delta_x f\,\mathrm{d}x_t + \Delta_t f\,\mathrm{d}t + \tfrac{1}{2}\Delta_{xx} f\,\mathrm{d}\langle x\rangle_t.$$

The formula has the same form as the conventional Itô formula, however the derivatives have different meanings.

REMARK 2 If the functional is of the form $f(X_t) = h(x_t, t)$, then $\Delta_x f = \partial h/\partial x$, $\Delta_t f = \partial h/\partial t$ and we are in the classical setting and we regain Itô's formula.

3.2. *Infinitesimal generator*

Assume

$$\mathrm{d}x_t = a(X_t)\,\mathrm{d}t + b(X_t)\,\mathrm{d}w_t$$

with w a standard P- Brownian motion (we recall that in the notation here processes are denoted by lower case letters and their paths by capital letters) and a,b are functionals such that the SDE is well posed (Protter 2005). This expresses the possible dependency of the evolution of the process on its current path. We define the infinitesimal generator A which applies to a smooth functional $f : \Lambda \longrightarrow \mathbb{R}$ by

$$Af(X_t) \equiv \lim_{\delta t\to 0} \frac{E[f(X_{t+\delta t})|X_t] - f(X_t)}{\delta t}.$$

Taking conditional expectation in Theorem 3.1, we have:

$$E[f(X_{t+\delta t})|X_t] - f(X_t) = E\left[\int_t^{t+\delta t} a(X_u)\Delta_x f(X_u)\,\mathrm{d}u + \int_t^{t+\delta t} \Delta_t f(X_u)\,\mathrm{d}u + \frac{1}{2}\int_t^{t+\delta t} b^2(X_u)\Delta_{xx} f(X_u)\,\mathrm{d}u\right].$$

Dividing by δt and passing to the limit:

$$Af(X_t) = \Delta_t f(X_t) + a(X_t)\Delta_x f(X_t) + \frac{b^2(X_t)}{2}\Delta_{xx} f(X_t).$$

This infinitesimal generator generalizes the usual one in two ways: first the dynamics is possibly path dependent; second, f is a functional that may depend on the history of the process, as opposed to a function of merely the current value of the process. We are resolutely in a non Markov setting.

3.3. *Functional Feynman-Kac formula*

As an application, we have what can be considered as the Feynman-Kac formula for functionals:

THEOREM 3.2 *Let x follow $dx_t = a(X_t)\,\mathrm{d}t + b(X_t)\,\mathrm{d}w_t$ with a,b such that SDE is well defined. For suitably integrable $g : \Lambda_T \longrightarrow \mathbb{R}, 0 \le T \le \overline{T}$ and $r : \Lambda \longrightarrow \mathbb{R}$, we define $f : \Lambda \longrightarrow \mathbb{R}$ by:*

$$f(Y_t) \equiv E[e^{-\int_t^T r(Z_u)\,\mathrm{d}u} g(Z_T)|Z_t = Y_t]$$

where $\left\{ \begin{smallmatrix} \text{for } u\in[0,t], Z_T(u)=Y_t(u) \\ \text{for } u\in[t,T], dz_u=a(Z_u)\,\mathrm{d}u+b(Z_u)\,\mathrm{d}w_u \end{smallmatrix} \right.$.

Then, if f is a smooth functional, it satisfies

$$\Delta_t f(X_t) + a(X_t)\Delta_x f(X_t) - r(X_t) f(X_t) + \frac{b^2(X_t)}{2}\Delta_{xx} f(X_t) = 0.$$

Proof We apply the Functional Itô formula of Theorem 3.1 to

$$h(Y_t) \equiv E[e^{-\int_0^T r(Z_u)\,\mathrm{d}u} g(Z_T)|Z_t = Y_t] = e^{-\int_0^t r(X_u)\,\mathrm{d}u} f(Y_t).$$

It gives

$$dh = e^{-\int_0^t r\,\mathrm{d}u}\left(\Delta_x f\,\mathrm{d}x + (\Delta_t f - rf)\,\mathrm{d}t + \frac{1}{2}\Delta_{xx} f\,\mathrm{d}\langle x\rangle\right)$$
$$= e^{-\int_0^t r\,\mathrm{d}u}\left(\left(\Delta_t f + a\Delta_x f - rf + \frac{b^2}{2}\Delta_{xx} f\right)\mathrm{d}t + b\Delta_x f\,\mathrm{d}w\right)$$

Expressing that its drift is 0, as h is a martingale, gives the result. ■

REMARK 3 The current realized path Y_t need not be generated by the dynamics of x. For instance, it can have jumps. What matters is that the functional is regular enough to be defined on paths that start with Y_t and follow with the x dynamics.

3.4. *Martingale representation*

Combining Theorems 3.1 and 3.2, we have the following representation of the random variable $g(X_T)$:

THEOREM 3.3 *Under the assumptions of Theorem 3.2, with $r = 0$ and $f(Y_t) \equiv E[g(Z_T)|Z_t = Y_t]$,*

$$g(X_T) = E[g(X_T)|X_0] + \int_0^T b(X_t)\Delta_x f(X_t)\,\mathrm{d}w_t.$$

Proof We apply Theorem 3.2 to f; jointly with Theorem 3.1 it leads to

$$f(X_T) = f(X_0) + \int_0^T b(X_t)\Delta_x f(X_t)\,\mathrm{d}w_t$$

which is the statement of the theorem. ■

This gives an explicit expression for the integrand in the Martingale Representation Theorem and can be compared with the Clark-Ocone formula from Malliavin calculus:

$$g(X_T) = E[g(X_T)|X_0] + \int_0^T b(X_t)E[D_t g(X_t)|X_t]\,\mathrm{d}w_t.$$

When both $E[D_t g(X_T)|X_t]$ and $\Delta_x f(X_t)$ are defined, they are equal. The former is the expectation of a pathwise derivative whilst the latter is the derivative of an expectation, which has more regularity. For instance, if $g(X_T) \equiv h(x_T)$, with $h \in C^0$ but not C^1, the latter is not defined in the classical sense whilst the former is well defined.

One reading of it is that in a complete model, the hedging ratio of a path dependent option is the derivative of its price with respect to the spot price, in the sense of the definition of $\Delta_x f(X_t)$.

3.5. *Functional PDE for exotic options*

We now present a result which is important for pricing and hedging path dependent options with possible path dependent dynamics.

THEOREM 3.4 *Under the assumptions of Theorem 3.2, with instantaneous interest rate $r(X_t)$, assume that the price of the option for the current path X_t is a smooth functional $f(X_t)$, then*

$$\Delta_t f(X_t) + \tfrac{1}{2}b^2\Delta_{xx} f(X_t) + r(X_t)(\Delta_x f(X_t)x_t - f(X_t)) = 0$$

Proof Let us apply the functional Itô formula of Theorem 1 to $f(X_t)$:

$$\mathrm{d}f = \Delta_x f\,\mathrm{d}x + \Delta_t f\,\mathrm{d}t + \tfrac{1}{2}b^2\Delta_{xx} f\,\mathrm{d}t.$$

The portfolio PF of option f with a short position of $\Delta_x f$ stocks gives

$$\mathrm{d}PF = \Delta_t f\,\mathrm{d}t + \tfrac{1}{2}b^2\Delta_{xx} f\,\mathrm{d}t.$$

In the absence of arbitrage, this riskless portfolio has to earn the instantaneous interest rate:

$$\mathrm{d}PF = r(f - \Delta_x f x)\,\mathrm{d}t.$$

This implies

$$\Delta_t f + \tfrac{1}{2}b^2\Delta_{xx} f - r(f - \Delta_x f x) = 0.$$

■

This says that the Black-Scholes PDE (Black and Scholes 1973, Merton 1973) holds for path dependent options (and path dependent dynamics) as well. However, there is a twist: the Greeks $\theta \equiv \Delta_t f$ and $\Gamma \equiv \Delta_{xx} f$ are themselves path dependent.

The Black-Scholes PDE is a hedging PDE that expresses the link between time decay and convexity. In general, the coefficients are path dependent and it cannot be used as a pricing PDE. However, in the case of a small number of state variables, we can deduce a pricing PDE. If the path dependency can be summed up with a finite number of state variables $z_1, \ldots, z_n$ then

$$f(X_t) = g(z_{1,t}, \ldots, z_{n,t})$$

and we can apply the chain rule:

$$\Delta_x f(X_t) = \sum_{i=1}^{n} \frac{\partial g}{\partial z_i} \frac{\partial z_i}{\partial x}.$$

We exemplify this in the case of Asian options and retrieve the usual pricing PDE and a 'better' one.

3.6. *Classical PDE for Asian options*

Assume $\mathrm{d}x_t = b(x_t, t)\,\mathrm{d}w_t$. Define $I_t \equiv \int_0^t x_u\,\mathrm{d}u, f(X_t) \equiv g(x_t, I_t, t)$

$$\Delta_t I = x_t \Rightarrow \Delta_t f = \frac{\partial g}{\partial t} + \frac{\partial g}{\partial I} x$$

$$\Delta_x I = 0 \Rightarrow \begin{cases} \Delta_x f = \dfrac{\partial g}{\partial x} \\ \Delta_{xx} f = \dfrac{\partial^2 g}{\partial x^2} \end{cases}$$

$$\Delta_t f + \frac{1}{2} b^2 \Delta_{xx} f = 0 \Rightarrow \frac{\partial g}{\partial t} + x\frac{\partial g}{\partial I} + \frac{1}{2} b^2 \frac{\partial^2 g}{\partial x^2} = 0$$

However, the presence of the convection term (second term) renders the finite difference discretization more difficult (Wilmott *et al.* 1995). This is solved by the following:

Better Asian PDE

Define $J_t \equiv E_t[\int_0^T x_u\,\mathrm{d}u] = \int_0^t x_u\,\mathrm{d}u + (T-t)x_t, f(X_t) \equiv h(x_t, J_t, t)$. Then:

$$\Delta_t J = 0 \Rightarrow \Delta_t f = \frac{\partial h}{\partial t}.$$

$$\Delta_x J = (T-t)x_t \Rightarrow \begin{cases} \Delta_x f = \dfrac{\partial h}{\partial x} + (T-t)\dfrac{\partial h}{\partial J} \\ \Delta_{xx} f = \dfrac{\partial^2 h}{\partial x^2} + 2(T-t)\dfrac{\partial^2 h}{\partial x \partial J} \\ \quad +(T-t)^2 \dfrac{\partial^2 h}{\partial J^2}. \end{cases}$$

$$\Delta_t f + \frac{1}{2} b^2 \Delta_{xx} f = 0 \Rightarrow \frac{\partial h}{\partial t} + \frac{1}{2} b^2 \times \left(\frac{\partial^2 h}{\partial x^2} + 2(T-t)\frac{\partial^2 h}{\partial x \partial J} + (T-t)^2 \frac{\partial^2 h}{\partial J^2} \right) = 0.$$

The choice of J as opposed to I absorbs the drift or carry term.

3.7. *Integration by parts*

In the spirit of the Malliavin calculus, we investigate the adjoint of the space differential operator to obtain an integration by parts formula.

If δ denotes the backward Itô integral

$$\delta(g) \equiv \int g(X_t)\,\delta x_t \equiv \lim_{t_i - t_{i-1} \to 0} \sum_i g(X_{t_i})(x_{t_i} - x_{t_{i-1}}),$$

we have the following relationship for any smooth functional h:

$$\delta(h) = \int_0^T h(X_t)\,\delta x_t = \int_0^T h(X_t)\,\mathrm{d}x_t + \int_0^T \Delta_x h(X_t)\,\mathrm{d}\langle x \rangle_t.$$

In particular, if x is a martingale,

$$E\left[\int_0^T h(X_t)\,\delta x_t\right] = E\left[\int_0^T \Delta_x h(X_t)\,\mathrm{d}\langle x \rangle_t\right].$$

Applying this to $h = fg$:

$$E\left[\int_0^T f(X_t)g(X_t)\,\delta x_t\right] = E\left[\int_0^T \Delta_x f(X_t)g(X_t)\,\mathrm{d}\langle x \rangle_t\right] + E\left[\int_0^T f(X_t)\Delta_x g(X_t)\,\mathrm{d}\langle x \rangle_t\right]$$

or

$$E\left[\int_0^T \Delta_x f(X_t)g(X_t)\,\mathrm{d}\langle x \rangle_t\right] = E\left[\int_0^T f(X_t)g(X_t)\,\delta x_t\right] - E\left[\int_0^T f(X_t)\Delta_x g(X_t)\,\mathrm{d}\langle x \rangle_t\right].$$

We can rewrite it in more concise form, with the notation $\langle f|g \rangle \equiv E[\int_0^T f(X_t)g(X_t)\,\mathrm{d}\langle x \rangle_t]$:

$$\langle \Delta_x f|g \rangle = E\left[\int_0^T f(X_t)g(X_t)\,\delta x_t\right] - \langle f|\Delta_x g \rangle.$$

We eventually obtain the following integration by parts formula:

$$\langle \Delta_x f|g \rangle = E[\delta(fg)] - \langle f|\Delta_x g \rangle.$$

In this formula, both f and g are functionals and $f(X_t), g(X_t), \Delta_x f(X_t)$ and $\Delta_x g(X_t)$ are processes. We can compare it with the integration by parts from the Malliavin calculus:

$$\langle D_t f|g \rangle = (f, \delta(g)),$$

where δ is the Skorokhod integral, f and $\delta(g)$ are random variables, $D_t f$ and g are processes and $(k, l) \equiv E[kl]$.

4. Examples of applications

We expound in this section three applications of the theory.

Example 1: *Profit and Loss from delta hedging and model impact*

Assume a local volatility model (Dupire 1994) $dx_t = \sqrt{v_0(x_t,t)}\,dw_t$. For $g \in \Lambda_T$, define $f \in \Lambda$ by $f(X_t) \equiv E^{Q_{v_0}}[g(X_T)|X_t]$, where $Q_{v_0} \equiv \sqrt{v_0}\,dp$.

By the functional PDE, $\Delta_t f(X_t) + \frac{1}{2}v_0(x_t,t)\Delta_{xx} f(X_t) = 0$.

If y follows $dy_t = \sqrt{v_t}\,dw_t$ with $y_0 = x_0$, by the functional Itô formula

$$
\begin{aligned}
g(Y_T) = f(Y_T) &= f(Y_0) + \int_0^T \Delta_x f(Y_t)\,dy_t + \int_0^T \Delta_t f(Y_t)\,dt \\
&\quad + \frac{1}{2}\int_0^T v_t \Delta_{xx} f(Y_t)\,dt \\
&= f(X_0) + \int_0^T \Delta_x f(Y_t)\,dy_t \\
&\quad + \frac{1}{2}\int_0^T (v_t - v_0(y_t,t))\Delta_{xx} f(Y_t)\,dt.
\end{aligned}
$$

Model impact and volatility expansion

Hence, with $\Pi_g(v) \equiv E^{Q_v}[g(Y_T)|X_0]$ and $\phi^v(x,t)$ the transition density for v,

$$
\begin{aligned}
\Pi_g(v) &= \Pi_g(v_0) + \frac{1}{2}E^{Q_v}\left[\int_0^T (v_t - v_0(x_t,t))\Delta_{xx} f(X_t)\,dt\right] \\
&= \Pi_g(v_0) + \frac{1}{2}\int_0^T \int \phi^v(x,t)E^{Q_v} \\
&\quad \times [(v_t - v_0(x,t))\Delta_{xx} f(X_t)|x_t = x]\,dx\,dt.
\end{aligned}
$$

In the case where v is of the form $v_0 + u : dx_t = \sqrt{v_0(x_t,t) + u(x_t,t)}\,dW_t$

$$
\begin{aligned}
\Pi_g(v_0+u) &= \Pi_g(v_0) + \frac{1}{2}\int_0^T \int \phi^{v_0+u}(x,t)u(x,t)E^{Q_{v_0+u}} \\
&\quad [\Delta_{xx} f(X_t)|x_t = x]\,dx\,dt \\
&= \Pi_g(v_0) + \int_0^T \int m(x,t)u(x,t)\,dx\,dt,
\end{aligned}
$$

where $m(x,t) \equiv \frac{1}{2}\phi^{v_0+u}(x,t)E^{Q_{v_0+u}}[\Delta_{xx} f(X_t)|x_t = x]$.

In particular,

$$
\begin{aligned}
\Pi_g(v_0+\varepsilon u) &= \Pi_g(v_0) + \frac{\varepsilon}{2}\int_0^T \int \phi^{v_0+\varepsilon u}(x,t)u(x,t)E^{Q_{v_0+\varepsilon u}} \\
&\quad \times [\Delta_{xx} f(X_t)|x_t = x]\,dx\,dt.
\end{aligned}
$$

The Frêchet derivative satisfies:

$$
\begin{aligned}
\langle \nabla_v g, u\rangle &\equiv \lim_{\varepsilon\to 0}\frac{\Pi_g(v_0+\varepsilon u) - \Pi_g(v_0)}{\varepsilon} \\
&= \frac{1}{2}\int_0^T \int \phi^{v_0}(x,t)u(x,t)E^{Q_{v_0}} \\
&\quad \times [\Delta_{xx} f(X_t)|x_t = x]\,dx\,dt \\
&= \int_0^T \int m(x,t)u(x,t)\,dx\,dt,
\end{aligned}
$$

where

$$
m(x,t) \equiv \frac{1}{2}\phi^{v_0}(x,t)E^{Q_{v_0}}[\Delta_{xx} f(X_t) = x]
$$

is the sensitivity of Π_g to the local variance at (x,t) (i.e. the Frêchet derivative).

Example 2: *Running maximum of a path*

The functional $f(X_t) \equiv \max_{0\le u\le t} X_t(u)$ is not smooth: it is not differentiable in x as it exhibits a kink at $x_t = m_t$ so we will work with a regularized version. For $\alpha > 0$ we define the function

$$
g_\alpha(x,m) \equiv \begin{cases} m - \alpha \text{ for } 0 \le x \le m - 2\alpha \\ m - \alpha + \dfrac{(x-(m-2\alpha))^2}{4\alpha} \text{ for } m - 2\alpha \le x \le m \\ x \text{ for } x \ge m \end{cases}
$$

and the functional

$$
f_\alpha(X_t) \equiv g_\alpha(x_t, m_t).
$$

Its functional derivatives are

$$
\Delta_t f_\alpha(X_t) = 0
$$

$$
\Delta_x f_\alpha(X_t) = \begin{cases} 0 \text{ for } 0 \le x_t \le m_t - 2\alpha \\ \dfrac{x_t - (m_t - 2\alpha)}{2\alpha} \text{ for } m_t - 2\alpha \le x_t \le m_t \\ 1 \text{ for } x_t \ge m_t \end{cases}
$$

$$
\Delta_{xx} f_\alpha(X_t) = \begin{cases} 0 \text{ for } 0 \le x_t \le m_t - 2\alpha \\ \dfrac{1}{2\alpha} \text{ for } m_t - 2\alpha \le x_t \le m_t \\ 0 \text{ for } x_t \ge m_t. \end{cases}
$$

In short, $\Delta_{xx} f_\alpha(X_t) = \frac{1}{2\alpha}1_{[0,2\alpha]}(m_t - x_t)$.

We thus get

$$
\begin{aligned}
f_\alpha(X_t) &= f_\alpha(X_0) + \int_0^t \Delta_x f_\alpha(X_u)\,dx_u + \int_0^t \Delta_t f_\alpha(X_u)\,du \\
&\quad + \frac{1}{2}\int_0^t \Delta_{xx} f_\alpha(X_u)\,d\langle x\rangle_u \\
&= x_0 + \int_0^t \frac{1_{[0,2\alpha]}(m_u - x_u)}{2\alpha}(x_u - (m_u - 2\alpha))\,dx_u \\
&\quad + \int_0^t 1_{(m_u,+\infty)}(x_u)\,dx_u \\
&\quad + \frac{1}{2}\int_0^t \frac{1_{[0,2\alpha]}(m_u - x_u)}{2\alpha}\,d\langle x\rangle_u.
\end{aligned}
$$

The first integral has a bounded integrand on a shrinking interval, the second one is nil and the last one is half the quadratic variation accumulated when the drawdown $x_u - m_u$ is less than 2α, divided by the width of this interval 2α.

Passing to the limit as $\alpha \to 0$, only the last one survives and we obtain that the increment of the maximum from the initial value is the local time at level 0 of the drawdown $m - x$:

$$
m_t = x_0 + \tfrac{1}{2}L_t^{m_t - x_t}.
$$

This equality holds almost surely pathwise as opposed to the classical equalities in law from Paul Levy. For the Brownian motion case, it states that the maximum is equal to the time spent beating the maximum.

Example 3: *Submartingales and lower bounds*

If f is a smooth functional and x_t is a martingale, then $y_t \equiv f(X_t)$ is a submartingale if and only if for all t

$$\tfrac{1}{2}\Delta_{xx} f(X_t)\, \mathrm{d}\langle x\rangle_t + \Delta_t f(X_t)\, \mathrm{d}t \geq 0.$$

If we require that this holds for all martingales, we get the following.

If f is a smooth functional, $y_t \equiv f(X_t)$ is a submartingale for all martingales x_t if and only if

$$\Delta_{xx} f(X_t) \geq 0$$
$$\Delta_t f(X_t) \geq 0.$$

The lower bound process (see El Karoui and Quenez 1995, Kramkov 1996) $LB(X_t) \equiv \inf_Q E^Q[g(X_T)|X_t]$ is a submartingale for all martingale measures in the set. If it corresponds to a smooth functional, we can write it as

$$\begin{aligned} LB(X_T) = LB(X_t) &+ \int_t^T \Delta_x LB(X_u)\, \mathrm{d}x_t \\ &+ \frac{1}{2}\int_t^T \Delta_{xx} LB(X_u)\, \mathrm{d}\langle x\rangle_u \\ &+ \int_t^T \Delta_t LB(X_u)\, \mathrm{d}u \\ = LB + a + k_1 + k_2, \end{aligned}$$

which can be interpreted as the Kramkov decomposition with $a \equiv \int_t^T \Delta_x LB(X_u)\, \mathrm{d}x_t$ the martingale component and the non decreasing process $k_1 + k_2$.

5. Conclusion

This work is motivated by a desire to develop a theoretical framework that reflects the modeling needs of some practical situations. It grew from applications, which ensures precedence of relevance over elegance. In a nutshell, it accounts for the dependency of a consequence on the history of its cause.

The core result is an extension of Itô's formula to functionals of paths. It describes the variations of a functional in terms of some partial derivatives which affect the current state, reflecting the fact that we cannot change the past and do not know the future.

It leads to a functional Feynman-Kac formula, which caters for both path dependent dynamics and path dependent pay-off function, an explicit integrand in the Martingale Representation Theorem. An important result for financial applications is the fact that the price of a path dependent claim written as a function of the price path so far satisfies a functional partial differential equation that expresses the trade-off between the time derivative and the price convexity of the claim. This in turn leads to the computation of the impact of a change of model on the option price and to perturbation analysis. Eventually, we develop an integration by parts formula.

5.1. Extensions

This work can be extended in many directions, both in terms of the underlying dynamics and in terms of the dependent variable represented by the functional.

The dynamics can be extended to the discontinuous case and to multi dimensions, even an infinite number of dimensions, in the spirit of da Prato.

The regularity of the functional can be partially relaxed, requiring merely convexity (or a difference of convex functions) in the x dimension. The functional may take values in a larger space than the real line; it can be multi-valued, take the quadratic variation as an argument, be infinite dimensional: The dependent variable could be the shape of a melting structure as a function of the temperature. How is it affected by a sudden change? by a stale temperature?

We hope to have laid the ground for a new field that may attract some interest and lead to new questions.

Acknowledgments

It is my pleasure to thank Nicole El Karoui, Rama Cont, Philip Protter, Raphael Douady, the participants of the Princeton conference on Implied Volatility Models at Huntington Beach in October 2008 where this work was first presented, and my colleagues at Bloomberg, for enriching conversations. Thank you to the Editors for their invitation and copy edit of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Ahn, H., Semimartingale integral representation. *Ann. Probab.*, 1997, **25**(2), 997–1010.
Black, F. and Scholes, M., The pricing of options and corporate liabilities. *J. Political Econ.*, 1973, **81**(3), 637–654.
Da Prato, G. and Zabczyk, J., *Stochastic Equations in Infinite Dimensions*, 1992 (Cambridge University Press: Cambridge).
Dupire, B., Pricing with a smile. *Risk*, 1994, **7**(1), 18–20.
El Karoui, N. and Quenez, M.-C., Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.*, 1995, **33**(1), 29–66.
Kramkov, D.O., Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probab. Theory. Relat. Fields.*, 1996, **105**(4), 459–479.
Le Gall, J.-F., Mouvement brownien et calcul stochastique. *Notes de cours de DEA*. http://www.dma.ens.fr/~legall, 2008.
Merton, R.C., Theory of rational option pricing. *Bell J. Econ. Manage Sci.*, 1973, **4**, 141–183.
Protter, P.E., Stochastic differential equations. In *Stochastic Integration and Differential Equations*, pp. 249–361, 2005 (Springer).
Wilmott, P., Howson, S. and Dewynne, J., *The Mathematics of Financial Derivatives: A Student Introduction*, 1995 (Cambridge University Press: Cambridge).

Appendix

We present here a proof of Theorem 3.1 (restated here for convenience), preceded by a simple technical lemma.

A functional $f : \Lambda \longrightarrow \mathbb{R}$ is Λ-continuous at $X_t \in \Lambda$ if:

$$\forall \varepsilon > 0, \quad \exists \alpha > 0 : \forall Y_s \in \Lambda,$$
$$d_\Lambda(X_t, Y_s) < \alpha \Rightarrow |f(Y_s) - f(X_t)| < \varepsilon.$$

$f : \Lambda \longrightarrow \mathbb{R}$ is Λ-continuous if it is Λ-continuous at all $X_t \in \Lambda$.

This last notion of continuity is weaker than the notion of L^∞-uniform continuity. However, the following Lemma shows that it is equivalent to a seemingly stronger notion:

LEMMA A.1 Λ-Lemma *If f is Λ-continuous, then*

$$\forall X_t \in \Lambda \text{ continuous}, \quad \forall \varepsilon > 0, \quad \exists \alpha > 0 : \forall Y_s \in \Lambda,$$
$$d_\Lambda(X_t, Y_s) < \alpha \Rightarrow |f(Y_s) - f(X_t)| < \varepsilon.$$

Proof The point is to show that α does not depend on t. By contradiction: Assume there is a positive ε and sequences t_p, u_p in $[0,T]$ and $Y^p_{u_p} \in \Lambda_{u_p}$ such that $d_\Lambda(Y^p_{u_p}, X_{t_p}) < 1/p$ and $|f(Y^p_{u_p}) - f(X_{t_p})| > \varepsilon$. Then there is a subsequence $t_{\phi(p)}$ of t_p that converges to some $t^* \in [0,T]$ and we can write

$$|f(Y^{\phi(p)}_{u_{\phi(p)}}) - f(X_{t_{\phi(p)}})| \le |f(Y^{\phi(p)}_{u_{\phi(p)}}) - f(X_{t^*})| + |f(X_{t^*}) - f(X_{t_{\phi(p)}})|.$$

Noticing that

$$\begin{aligned} d_\Lambda(Y^{\phi(p)}_{u_{\phi(p)}}, X_{t^*}) &\le d_\Lambda(Y^{\phi(p)}_{u_{\phi(p)}}, X_{t_{\phi(p)}}) + d_\Lambda(X_{t_{\phi(p)}}, X_{t^*}) \\ &\le \frac{1}{\phi(p)} + d_\Lambda(X_{t_{\phi(p)}}, X_{t^*}) \xrightarrow{p} 0, \end{aligned}$$

we see that the right-hand-side goes to 0 by Λ-continuity of f. This contradicts $|f(Y^{\phi(p)}_{u_{\phi(p)}}) - f(X_{t_{\phi(p)}})| > \varepsilon$ and concludes the Lemma. ∎

THEOREM A.1 *If x is a continuous semi-martingale and X_t denotes its path over $[0,t]$ and is $f : \Lambda \longrightarrow \mathbb{R}$ is Λ-continuous, C^2 in x and C^1 in t, with these derivatives themselves Λ-continuous, then, for all $T \ge 0$,*

$$\begin{aligned} f(X_T) = f(X_0) &+ \int_0^T \Delta_x f(X_t)\, \mathrm{d}x_t + \int_0^T \Delta_t f(X_t)\, \mathrm{d}t \\ &+ \frac{1}{2}\int_0^T \Delta_{xx} f(X_t)\, \mathrm{d}\langle x\rangle_t. \end{aligned}$$

Proof In the classical Itô calculus case, a concise proof applies integration by parts to show that the result holds for polynomials, hence by density to all C^2 functions. This argument does not apply in a simple way to functionals and we follow a more classical approach (Le Gall 2008) by way of Taylor expansion.

We take a sequence of nested subdivisions of $[0,T]$, $0 = t^n_0 < \cdots < t^n_{p_n} = T$ with mesh less than $1/n$.

For $X \in \Lambda_T$ continuous we define X^n_i as short-hand for $X_{t^n_i}$ and Y^n_i and Z^n_i in $\Lambda_{t^n_i}$ by

$$\begin{aligned} Y^n_i(s) &= X_T(t^n_{j-1}), \quad \text{if } s \in [t^n_{j-1}, t^n_j),\ j \le i, \\ Y^n_t(t^n_i) &= X_T(t^n_i). \\ Z^n_i(s) &= Y^n_i(s), \quad \text{if } s < t^n_i, \\ Z^n_i(t^n_i) &= X_T(t^n_{i-1}). \end{aligned}$$

Moreover, we define $\delta x^n_i = x_{t^n_i} - x_{t^n_{i-1}}$ and $\delta t^n_i = t^n_i - t^n_{i-1}$. We notice that $Y_0 = Z_0 = X_0$ so

$$f(X_T) - f(X_0) = f(X_T) - f(Y^n_{p_n}) + \sum_{i=1}^{p_n} f(Y^n_i) - f(Y^n_{i-1}).$$

We perform a Taylor expansion of the summand, applying the Mean Value theorem

(a) in space to $h \mapsto f((Z^n_i)^h)$ as f is twice continuously differentiable in x;
(b) in time to $\delta t \mapsto f(Y^n_{i-1,\delta t})$ as $\Delta_t f$ is right-continuous and $f(Y^n_{i-1,\tau})$ is continuous in τ.

We thus get the existence of $y^n_i \in (0, \delta x^n_i)$ and $\theta^n_i \in (0, \delta t^n_i)$ such that:

$$\begin{aligned} f(Y^n_i) - f(Y^n_{i-1}) &= (f(Y^n_i) - f(Z^n_i)) + (f(Z^n_i) - f(Y^n_{i-1})) \\ &= \Delta_x f(Z^n_i)\, \delta x^n_i + \frac{1}{2}\Delta_{xx} f((Z^n_i)^{y^n_i})(\delta x^n_i)^2 \\ &\quad + \Delta_t f(Y^n_{i-1,\theta^n_i})\, \delta t^n_i. \end{aligned}$$

Further splitting δx^n_i into the finite variation part δa^n_i and the local martingale part δm^n_i in the decomposition of the semi-martingale $x_t = x_0 + a_t + m_t$, we get

$$f(X_T) - f(X_0) = A^n + B^n_{1,a} + B^n_{1,m} + \frac{1}{2}B^n_2 + C^n,$$

with

$$\begin{aligned} A^n &\equiv f(X_T) - f(Y^n_{p_n}), \\ B^n_{1,a} &\equiv \sum_{i=1}^{p_n} \Delta_x f(Z^n_i)\, \delta a^n_i, \\ B^n_{1,m} &\equiv \sum_{i=1}^{p_n} \Delta_x f(Z^n_i)\, \delta m^n_i, \\ B^n_2 &\equiv \sum_{i=1}^{p_n} \Delta_{xx} f((Z^n_i)^{y^n_i})(\delta x^n_i)^2, \\ C^n &\equiv \sum_{i=1}^{p_n} \Delta_t f(Y^n_{i-1,\theta^n_i})\, \delta t^n_i. \end{aligned}$$

We are going to tackle these five terms one by one.

(1) $A^n = f(X_T) - f(Y^n_{p_n})$:

Pathwise, x_t is continuous on $[0,T]$, hence uniformly continuous, and $y^n_{p_n}$ converges uniformly (and for the Λ-distance) to X_T. By Λ-continuity of f, $A^n \xrightarrow{n} 0$.

(2) $B^n_{1,a} = \sum_{i=1}^{p_n} \Delta_x f(Z^n_i)\, \delta a^n_i$:

For a given X, define the functions $\phi(t) = \Delta_x f(X_t)$ and $\phi_n(t) = \sum_{i=1}^{p_n} \Delta_x f(Z^n_i)\, 1_{[t^n_{i-1}, t_i)}(t)$. If i is the index such that $t^n_{i-1} \le t < t^n_i$, $d_\Lambda(Z^n_i, X_t) \le \|Z^n_i - X^n_i\|_\infty + t^n_i - t^n_{i-1}$ converges to 0 uniformly in t. As $\Delta_x f$ is Λ-continuous, by the Λ-Lemma, $\phi^n(t) = \Delta_x f(Z^n_i)$ converges uniformly to $\phi(t) = \Delta_x f(X_t)$, which is continuous in t. By ordinary dominated convergence,

$$\begin{aligned} B^n_{1,a} = \sum_{i=1}^{p_n} \Delta_x f(Z^n_i)\, \delta a^n_i &= \int_0^T \phi^n(t)\, \mathrm{d}a_t \xrightarrow{n} \int_0^T \phi(t)\, \mathrm{d}a_t \\ &= \int_0^T \Delta_x f(X_t)\, \mathrm{d}a_t. \end{aligned}$$

(3) $B^n_{1,m} = \sum_{i=1}^{p_n} \Delta_x f(Z^n_i)\, \delta m^n_i$:

We take $g = \Delta_x f$ which is Λ-continuous and define the processes $y_t = g(X_t)$ and $y^n_t = \sum_{i=1}^{p_n} g(Z^n_i) 1_{[t^n_{i-1}, t_i)}(t)$. Both processes are adapted to $\mathcal{F}_t$ as Z^n_i is known at $t^n_{i-1} \le t$. For each integer p, we define stopping times τ^1_p, τ^2_p and τ_p by

a) $\tau^1_p = \inf\{t \in [0,T] : |y_t| + \langle m\rangle_t \ge p\}$.

b) We define A_p as the set of stepwise approximations of X_T with a mesh of size at most $1/p$:

$$A_p \equiv \left\{ Y \in \Lambda_T : Y(t) = \sum_{i=1}^{p_n} X(t_{i-1}) 1_{[t^n_{i-1}, t_i)}(t), \right.$$
$$\left. \times \text{ with } t_{i-1} < t_i < t_{i-1} + \frac{1}{p} \right\}.$$

We define our second stopping time as

$$\tau^2_p = \inf\{t \in [0,T] : \exists\, Y \in A_p : |g(Y_t) - g(X_t)| > 1\},$$

where $g = \Delta_x f$.

LEMMA A.2 $\mathbb{P}(\tau_p^2 = T) \xrightarrow{p} 1$.

Proof We take a given continuous $X \in \Lambda_T$. As g is Λ-continuous, by the Λ-Lemma,

$$\exists\, \alpha > 0 : \|Y_t - X_t\|_\infty < \alpha \Rightarrow |g(Y_t) - g(X_t)| < 1.$$

$X(t)$ is continuous, thus uniformly continuous so there is $\beta > 0$ such that $|u - v| < \beta \Rightarrow |X(u) - X(v)| < \alpha$. Thus, for $p > 1/\beta$, for any $Y \in A_p$ and all $t \in [0, T]$,

$$\begin{aligned}|u - v| < \frac{1}{p} &\Rightarrow |u - v| < \beta \Rightarrow |X(u) - X(v)| < \alpha\\ &\Rightarrow \|Y_t - X_t\|_\infty < \alpha \Rightarrow |g(Y_t) - g(X_t)| < 1.\end{aligned}$$

This means that $\tau_p^2 = T$ and proves the Lemma. ■

c) Finally, we define a third stopping time by $\tau_p \equiv \tau_p^1 \wedge \tau_p^2$ and m^{τ_p} by $m_t^{\tau_p} = m_{t\wedge\tau_p}$.

Equipped with this potent stopping time, we proceed as follows:

$$\begin{aligned}\mathbb{E}&\left[\left(\sum_{i=1}^{p_n} g(Z_i^n)\,\delta(m^{\tau_p})_i^n - \int_0^T g(X_t)\,\mathrm{d}m_t^{\tau_p}\right)^2\right]\\ &= \mathbb{E}\left[\left(\int_0^{T\wedge\tau_p} (y_t^n - y_t)\,\mathrm{d}m_t\right)^2\right]\\ &= \mathbb{E}\left[\int_0^{T\wedge\tau_p} (y_t^n - y_t)^2\,\mathrm{d}\langle m\rangle_t\right],\end{aligned}$$

by Itô isometry.

We now show that we are in position to apply the stochastic dominated convergence theorem:

a) As $d_\Lambda(Z_i^n, X_t) \xrightarrow{n} 0$ and g is Λ-continuous, $g(Z_i^n) \xrightarrow{n} g(X_t)$, or $y_t^n \xrightarrow{n} y_t$.

b) For $t < \tau_p$, $|y_t^n| = |g(Z_i^n)| \le |g(Z_i^n) - g(X_t)| + |g(X_t)| \le 1 + p$ because $Z_i^n \in A_p$ and from the definition of τ_p.

Thus, by stochastic dominated convergence, $\mathbb{E}[\int_0^{T\wedge\tau_p}(y_t^n - y_t)^2\,\mathrm{d}\langle m\rangle_t] \xrightarrow{n} 0$, which tells us that

$$\sum_{i=1}^{p_n} g(Z_i^n)\delta(m^{\tau_p})_i^n \xrightarrow{L^2} \int_0^T g(X_t)\,\mathrm{d}m_t^{\tau_p}.$$

And, as $\mathbb{P}(\tau_p = T) \xrightarrow{p} 1$, we have

$$\sum_{i=1}^{p_n} g(Z_i^n)\delta m_i^n \xrightarrow{L^2} \int_0^T g(X_t)\,\mathrm{d}m_t.$$

4) $B_2^n = \sum_{i=1}^{p_n} \Delta_{xx} f((Z_i^n)^{y_i^n})(\delta x_i^n)^2$:

We define the processes $\phi(t) = \Delta_{xx} f(X_t)$ and $\phi^n(t) = \sum_{i=1}^{p_n} \phi_{n,i} 1_{[t_{i-1}^n, t_i)}(t)$ with $\phi_{n,i} \equiv \Delta_{xx} f((Z_i^n)^{y_i^n})$.

For m < n, we write $B_2^n = (B_2^n - a^{m,n}) + (a^{m,n} - b^m) + (b^m - c) + c$ with

$$B_2^n = \sum_{i=1}^{p_n} \phi_{n,i}(\delta x_i^n)^2,$$

$$a^{m,n} \equiv \sum_{j=1}^{p_m} \phi_{m,j} \sum_{i\in I_j^m} (\delta x_i^n)^2, \quad \text{where } I_j^m \equiv \{i : t_{j-1}^m \le t_i^n < t_j^m\},$$

$$b^m \equiv \sum_{j=1}^{p_m} \phi_{m,j}(\langle x\rangle_{t_j^m} - \langle x\rangle_{t_{j-1}^m}),$$

$$c \equiv \int_0^T \phi(t)\,\mathrm{d}\langle x\rangle_t.$$

a) $B_2^n - a^{m,n}$:

We notice that for each $X \in \Lambda$,

$$\begin{aligned}|B_2^n - a^{m,n}| &= \left|\sum_{i=1}^{p_n} \phi_{n,i}(\delta x_i^n)^2 - \sum_{j=1}^{p_m} h_{m,j} \sum_{i\in I_j^m} (\delta x_i^n)^2\right|\\ &\le \|\phi^n - \phi^m\|_\infty \sum_{i=1}^{p_n} (\delta x_i^n)^2.\end{aligned}$$

As in $B_{1,a}$, the Λ-Lemma tells us that for each $X \in \Lambda$, ϕ^n converges uniformly in t to ϕ, thus $\|\phi^n - \phi^m\|_\infty \le \|\phi^n - \phi\|_\infty + \|\phi - \phi^m\|_\infty$ converges to 0 when n and m go to infinity. As $\sum_{i=1}^{p_n}(\delta x_i^n)^2$ converges to $\langle x\rangle_T$ in probability, we deduce that $B_2^n - a^{m,n}$ converges to 0 in probability.

b) $a^{m,n} - b^m$:

For each m and each $j \le p_m$, $\sum_{i\in I_j^m}(\delta x_i^n)^2$ converges in probability to $\langle x\rangle_{t_j^m} - \langle x\rangle_{t_{j-1}^m}$, so $\sum_{j=1}^{p_m} \phi_{m,j} \sum_{i\in I_j^m}(\delta x_i^n)^2$ converges in probability to $\sum_{j=1}^{p_m} \phi_{m,j}(\langle x\rangle_{t_j^m} - \langle x\rangle_{t_{j-1}^m})$, which means that $a^{m,n} - b^m$ converges to 0 in probability.

c) $b^m - c$:

Similarly to $B_{1,a}^n$, the Λ-Lemma tells us that for each X, ϕ^n converges uniformly in t to ϕ continuous and by bounded convergence,

$$b^m = \sum_{j=1}^{p_m} \phi_{m,j}(\langle x\rangle_{t_j^m} - \langle x\rangle_{t_{j-1}^m}) = \int_0^T \phi^n(t)\,\mathrm{d}\langle x\rangle_t \xrightarrow{n} \int_0^T \phi(t)\,\mathrm{d}\langle x\rangle_t,$$

which says that $b^m - c$ converges to 0.

Combining *a*)–*c*), we get that

$$B_2^n = \sum_{i=1}^{p_n} \Delta_{xx} f((Z_i^n)^{y_i^n})(\delta x_i^n)^2 \xrightarrow{\text{probability}} \int_0^T \Delta_{xx} f(X_t)\,\mathrm{d}\langle x\rangle_t.$$

5) $C^n = \sum_{i=1}^{p_n} \Delta_t f(Y_{i-1,\theta_i^n}^n)\,\delta t_i^n$:

We define $g \equiv \Delta_t f$ and take a given $X \in \Lambda$. We further define the functions $\phi^n(t) = \sum_{i=1}^{p_n} g(Y_{i-1,\theta_i^n}^n) 1_{[t_{i-1}^n, t_i)}(t)$ and $\phi(t) = g(X_t)$ and follow the reasoning of $B_{1,a}^n$.

a) simple convergence:

$$d_\Lambda(Y_{i-1,\theta_i^n}^n, X_t) \le d_\Lambda(Y_{i-1,\theta_i^n}^n, Y_{i-1}^n) + d_\Lambda(Y_{i-1}^n, X_t) \le t_i^n - t_{i-1}^n + \varepsilon.$$

This means that $Y_{i-1,\theta_i^n}^n$ converges to X_t and, as g is Λ-continuous, $g(Y_{i-1,\theta_i^n}^n) \xrightarrow{n} g(X_t)$, or

$$h^n(t) \xrightarrow{n} h(t).$$

b) boundedness:

For n big enough, $|\phi^n(t)| = |g(Y_{i-1,\theta_i^n}^n)| \le |g(Y_{i-1,\theta_i^n}^n) - g(X_{i-1,\theta_i^n})| + |g(X_{i-1,\theta_i^n})| \le \varepsilon + M$.

Thus, by ordinary dominated convergence,

$$\begin{aligned}C^n &= \sum_{i=1}^{p_n} \Delta_t f(Y_{i-1,\theta_i^n}^n)\,\delta t_i^n = \int_0^T \phi^n(t)\,\mathrm{d}t \xrightarrow{n} \int_0^T \phi(t)\,\mathrm{d}t\\ &= \int_0^T \Delta_t f(X_t)\,\mathrm{d}t.\end{aligned}$$

Gathering (1)–(5), we obtain

$$\begin{aligned}f(X_T) - f(X_0) &= \int_0^T \Delta_x f(X_t)\,\mathrm{d}x_t + \int_0^T \Delta_t f(X_t)\,\mathrm{d}t\\ &\quad + \frac{1}{2}\int_0^T \Delta_{xx} f(X_t)\,\mathrm{d}\langle x\rangle_t,\end{aligned}$$

which is the statement of the theorem.