A Representation Theorem For Smooth Brownian Martingales

Henry Schellhorn
Claremont Graduate University
Henry.Schellhorn@cgu.edu

May 3, 2012

Abstract

We show that, under certain smoothness conditions, a Brownian martingale at a fixed time can be represented as an exponential of its value at a later time. The time-dependent generator of this exponential operator is equal to one half times the Malliavin derivative. This result can also be seen as a generalization of the semi-group theory of parabolic partial differential equations to the parabolic path-dependent partial differential equations introduced by Dupire (2009) and Cont and Founié (2011). The exponential operator can be calculated explicitly in a series expansion, which resembles the Dyson series of quantum mechanics. Our continuous-time martingale representation result is proved by a passage to the limit of a special case of a *backward Taylor expansion* of an approximating discrete-time martingale. The latter expansion can also be used for numerical calculations.

Keywords: Continuous Martingales, Malliavin Calculus, Probabilistic Methods for Partial Differential Equations, Path-Dependent Partial Differential Equations.

Mathematics Subject Classification: 60H07, 60G44, 65M75

1 Introduction

The problem of representing Brownian martingales has a long and distinguished history. Dambis (1965) and Dubins-Schwarz (1965) showed that continuous martingales can be represented as time-changed Brownian motions. Doob (1953), Wiener and Ito developed what is often called Ito's martingale representation theorem: every local Brownian martingale has a version which can be written as an Ito integral plus a constant. In this article, we consider martingales which are conditional expectations of a $\mathcal{F}_T$-measurable random variable F. When the random variable F is Malliavin differentiable, the Clark-Ocone formula (Clark (1970), and Ocone (1984)) states that the integrand in Ito's martingale representation theorem is equal to the conditional expectation of the Malliavin derivative of F. We consider the less general problem of "infinitely smooth" martingales, namely martingales which are conditional expectations of a $\mathcal{F}_T$-measurable random variable F, which is infinitely differentiable in the sense of Malliavin. We show that a Brownian martingale at a fixed time can be represented as an exponential of its value at a later time. The time-dependent generator of this exponential operator is equal to one half times the Malliavin derivative. While smoothness is a severe limitation to our result, our representation formula opens the way to new numerical schemes, and potentially some analytical asymptotic calculations. The exponential operator can be calculated explicitly in a series expansion, which resembles the Dyson series of quantum mechanics. There are two main differences between our martingale representation and the Dyson formula for the initial value problem in quantum mechanics. First, in the case of martingales, time flows backward. Secondly, the time-evolution operator is equal to one half of the second-order Malliavin derivative, while for the initial value problem in quantum mechanics the time-evolution operator is equal to $-2\pi i$ times the time-dependent Hamiltonian divided by the Planck constant.

Our continuous-time martingale representation result is proved by a passage to the limit of a special case of the *backward Taylor expansion* (BTE) of an approximating discrete-time martingale. We introduced (without formal proof) the BTE in Schellhorn and Morris (2009), and applied it to price American options numerically. The idea in that paper was to use the BTE to approximate, over one time-step, the conditional expectation of the option value at the next time-step. While not ideal to price American options because of the lack of differentiability of the payoff, the BTE is better suited to the numerical calculation of the solution of smooth backward stochastic differential equations (BSDE). In a related paper, Hu, Nualart, and Song (2011) introduce a numerical scheme to solve a BSDE with drift using Malliavin calculus. Their scheme can be viewed as a Taylor expansion carried out until the first order. Our BTE can be seen as a generalization to higher order of that idea, where the Malliavin derivative(s) is (are) calculated at the future time-step rather than at the current time-step.

In the second part of the paper, we consider the path-dependent partial differential equation (PPDE) introduced by Dupire (2009). By the functional Feynman-Kac formula introduced by Dupire(2009) and Peng and Wang (2011), the solution of the PPDE is directly related to the martingale solving the corresponding BSDE. We can thus characterize the solution of the PPDE as an exponential of the terminal condition. This result belongs to the classical domain of integration of the evolution equation (see e.g. Yosida (1978), chapter 14). In the smooth case, it generalizes in a natural way the semi-group theory of integration of the classical diffusion equation (see e.g., McOwen (1996), chapter 9.2).

The structure of this paper is the following. We first expose the discrete-time result, namely the Backward Taylor Expansion (BTE) for discrete functionals, and then prove our main result, namely our martingale representation theorem. Two explicit examples are given, which show the usefulness of the Dyson series in analytic calculations. In the second part of the paper we review the definition of the PPDE and provide an explicit exponential formula for its solution. The PPDE notation, which is not necessary to understand the first part, is introduced in the second part. The proofs of the theorems, as well as the notation necessary for the proofs, are relegated to the appendix.

2 Martingale Representation

The uncertainty is described by the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$. Here $\{\mathcal{F}_t\}$ is the filtration generated by Brownian motion W on $\mathbb{R}$. Most results can be easily generalized to Brownian motion on $\mathbb{R}^d$. We denote the Malliavin derivative of order l of F at time t by $D^l_t F$. We call $\mathbb{D}_\infty([0,T])$ the set of random variables which are infinitely Malliavin differentiable and $\mathcal{F}_T$-measurable. Given $\omega \in \Omega$, we denote by $\omega^t(\omega)$ be the path that "freezes" Brownian motion after time t:

$$W(s, \omega^t(\omega)) = \begin{cases} W(s,\omega) & \text{if} \quad s \le t \\ W(t,\omega) & \phantom{\text{if}} \quad t \le s \le T \end{cases}$$

As is conventional for regular derivatives, the notation $D^l_t F(\omega^t(\omega))$ refers to the value of the Malliavin derivative of F along scenario $\omega^t(\omega)$, and not to the value of the composition of $F \circ \omega^t$ evaluated at ω. For instance $\frac{1}{2} D^1_s W^2(\omega^t(\omega)) = 1[s \le T] W(t, \omega)$.

2.1 Backward Taylor Expansion

THEOREM 1 *Let $F \in \mathbb{D}_\infty([0,T])$be a $\sigma\{W(\Delta), .., W(T)\}$-measurable random variable. Suppose that $D^l_{(m+1)\Delta} F = 0$ for $l > L$. Let $\gamma(m, L)$ be given by $\gamma(m, L) = 1$ if $L = 0$ and, otherwise by:*

$$\gamma(m, L) = 1\{L \text{ even}\} \left(\frac{\Delta}{2}\right)^{L/2} \frac{1}{(L/2)!} - \sum_{l=0}^{L-1} \gamma(0, l) \frac{(W((m+1)\Delta) - W(m\Delta))^{L-l}}{(L-l)!} \tag{1}$$

Then

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{l=0}^{L} \gamma(m, l) E[D^l_{(m+1)\Delta} F | \mathcal{F}_{((m+1)\Delta)}] \tag{2}$$

We now consider functionals F which Malliavin derivatives do not vanish. The main application of a Taylor series comes from truncating it. We now proceed to estimate the truncation error. Let $\hat{F}^L_m$ be an approximation of $E[F|\mathcal{F}_{m\Delta}]$ obtained by supposing that F has order L:

$$\hat{F}^L_m = \sum_{l=0}^{L} \gamma(m,l) E[D^l_{(m+1)\Delta} F | \mathcal{F}_{((m+1)\Delta)}]$$

THEOREM 2 *Let $F \in \mathbb{D}_\infty([0,T])$ be a $\sigma\{W(\Delta),..,W(T)\}$-measurable random variable. The mean square truncation error is:*

$$E[(E[F|\mathcal{F}_{m\Delta}] - \hat{F}^L_m)^2 | \mathcal{F}_{m\Delta}] = O(\Delta^{L+1}) \tag{3}$$

Combining theorems 1 and 2, we arrive at the backward Taylor expansion.

THEOREM 3 *Let $F \in \mathbb{D}_\infty([0,T])$ be a $\sigma\{W(\Delta),..,W(T)\}$-measurable random variable. Then*

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{l=0}^{\infty} \gamma(m,l) E[D^l_{(m+1)\Delta} F | \mathcal{F}_{((m+1)\Delta)}] \tag{4}$$

Applying (4) recursively, one obtains the following corollary.
COROLLARY

$$E[F|\mathcal{F}_{m\Delta}] = \sum_{j_{m+1}=0}^{\infty} .... \sum_{j_M=0}^{\infty} \prod_{k=m+1}^{M} \gamma(k, j_k) D^{j_{m+1}}_{(m+1)\Delta} .. D^{j_M}_{M\Delta} F \tag{5}$$

Observations: A non-intuitive feature of the Backward Taylor expansion is that *any path* can be chosen to approximate conditional expectations backward. Suppose that the paths of Brownian motion are fixed in advance, like in regression-based algorithms to calculate American options. The BTE can be used as a type of "control variate" to speed up the convergence of the backward induction, as we shall make precise in another article. If the paths are not chosen in advance, an interesting equation similar to (7) emerges when we choose the paths:

$$W((m+1)\Delta) = W(m\Delta) + i\sqrt{\Delta}$$

In the next subsection we will choose the "certainty-equivalent" paths:

$$W((m+1)\Delta) = W(m\Delta)$$

to derive our main result. For numerical applications, it is crucial to choose a low order of expansion L, in order to keep the number of calculations in (5) from growing too fast. One could then imagine a scheme where, at each step, the "optimal path" is chosen so as to minimize the global truncation error in (3). We leave all these considerations for future research.

2.2 Exponential Formula

For esthetical reasons we introduce a "chronological operator". In this we follow Zeidler (2006). Let $(H(t))$ be a collection of operators. The chronological operator $\mathcal{T}$ is defined by

$$\mathcal{T}(H(t_1)H(t_2)..H(t_n)) := H(t_{1'})H(t_{2'})...H(t_{n'})$$

where $t_{1'}, .., t_{n'}$ is a permutation of $t_1, .., t_n$ such that $t_{1'} \geq t_{2'} \geq .. \geq t_{n'}$.
Example
It is showed in Zeidler (2006) p. 44-45 that

$$\int_0^t \int_0^{t_2} H(t_1)H(t_2) dt_1 dt_2 = \frac{1}{2} \int_0^t \int_0^t \mathcal{T}(H(t_1)H(t_2)) dt_1 dt_2$$

This will be the only property of the chronological operator we will use in this article.

Definition: the exponential operator of a time-dependent generator H is:

$$\mathcal{T}\exp(\int_t^T H(s)ds) = \sum_{k=0}^{\infty}\int_t^T..\int_t^T \mathcal{T}(H(\tau_1),..H(\tau_k))d\tau_1..d\tau_k \tag{6}$$

In quantum field theory, the series on the right handside of (6) is called a *Dyson series*.

THEOREM 3 *Suppose $F \in D_\infty([0,T])$. Let $M(t) = E[F|\mathcal{F}_t]$ for $t \leq T$. Then, in $L^2(P)$:*

$$M(t,\omega) = \mathcal{T}\exp(\frac{1}{2}\int_t^T D_s^2 ds)F(\omega^t(\omega)) \tag{7}$$

The importance of the exponential formula (7) stems from the Dyson series representation (6), which we rewrite hereafter in a more convenient way:

$$M(t,\omega) = F(\omega^t(\omega)) + \frac{1}{2}\int_t^T D_s^2 F(\omega^t(\omega))ds + \frac{1}{4}\int_t^T\int_{s_1}^T D_{s_1}^2 D_{s_2}^2 F(\omega^t(\omega))ds_2 ds_1 + ...$$

It can be used for either numerical calculations (which we have not tried yet) of for analytical calculations, as we show in the next subsection.

2.2.1 Solution of Some Problems by Dyson Series

We provide two different examples where conditional expectations can be calculated by Dyson series. The first example is a very well-known example, but it illustrates nicely the computation of Dyson series in case the random variable F (seen as a functional of Brownian motion) is not path-dependent. In the second example, the functional F is path-dependent.

Example 1:

Let $\tau > T$ and:

$$F(\omega) = \frac{1}{\sqrt{\tau - T}}\exp(-\frac{W^2(T,\omega)}{2(\tau - T)})$$

We calculate the Dyson series:

$$\begin{aligned}
M(t,\omega) &= M(T,\omega^t(\omega)) + \frac{1}{2}\int_t^T D_{s_1}^2\frac{1}{\sqrt{\tau - T}}\exp(-\frac{W^2(T,\omega^t(\omega))}{2(\tau - T)})ds_1 + .. \\
&= M(T\omega^t(\omega)) + \frac{1}{2}\int_t^T (\frac{W^2(t,\omega)}{(\tau - T)^{5/2}} - \frac{1}{(\tau - T)^{3/2}})\exp(-\frac{W^2(t,\omega)}{2(\tau - T)}) + .. \\
&= \frac{1}{\sqrt{\tau - T}}\exp(-\frac{W^2(t,\omega)}{2(\tau - T)}) + \frac{1}{2}(T - t)(\frac{W^2(t,\omega)}{(\tau - T)^{3/2}} - \frac{1}{(\tau - T)^{1/2}})\exp(-\frac{W^2(t,\omega)}{2(\tau - T)}) + .. \\
&= \frac{1}{\sqrt{\tau - T}}\exp(-\frac{W^2(t,\omega)}{2(\tau - T)}) + (t - T)\frac{\partial}{\partial T}\left(\frac{1}{\sqrt{\tau - T}}\exp(-\frac{W^2(t,\omega)}{2(\tau - T)}).\right) + .. \\
&= \frac{1}{\sqrt{\tau - t}}\exp(-\frac{W^2(t,\omega)}{2(\tau - t)})
\end{aligned}$$

Observation: We deliberately took $\tau > T$ so that the functional F would be infinitely Malliavin differentiable. It remains to be seen whether proper convergence results can be obtained when $\tau \downarrow T$.

Example 2: a path-dependent functional

Let $F(\omega) = \exp(-\int_0^T W(s,\omega)ds)$. We introduce informally the process $\dot{W}$ (singular white noise) in order to ease the calculation of the Malliavin derivatives. The functional G is indirectly defined as:

$$\begin{aligned}
F(\omega) &= G(\dot{W}(\omega)) = \exp(-\int_0^T\int_0^s \dot{W}(u,\omega)du\ ds)) \\
&= \exp(-\int_0^T (T - u)\dot{W}(u,\omega)du))
\end{aligned}$$

Malliavin derivatives can be computed formally as regular derivatives of G. Thus:

$$\begin{aligned} D_s^2 F &= (T-s)^2 F \\ F(\omega^t(\omega)) &= \exp(-\int_0^t W(s,\omega)ds - W(t,\omega)(T-t)) \end{aligned}$$

The Dyson series becomes:

$$\begin{aligned} M(t,\omega) &= F(\omega^t(\omega)) + \frac{1}{2}\int_t^T (T-s_1)^2 F(\omega^t(\omega))ds_1 + .. \\ &= \exp(-\int_0^t W(s,\omega)ds - W(t,\omega)(T-t))(1+\frac{1}{2}\int_t^T (T-s_1)^2 ds_1) + ..) \\ &= \exp(-\int_0^t W(s,\omega)ds)\exp(-W(t,\omega)(T-t) + \frac{1}{6}(T-t)^3) \end{aligned}$$

3 Representation of Solutions of Path-dependent Partial Differential Equations

We now introduce some key concepts of the functional Ito calculus introduced by Dupire. For more information, the reader is sent to Cont and Fournié (2011), which we copy herefater almost verbatim. Let X:$[0,T]\times\Omega\to\mathbb{R}$ be a continuous semimartingale. The paths of X lie in $C_0([0,T],\mathbb{R})$. We also consider the space $D([0,t],\mathbb{R})$ of cadlag functions with values in $\mathbb{R}$. For a path $x\in D([0,T],\mathbb{R})$ we denote by $x(t)$ the value of x at t, and by $x_t^p=(x(u),0\le u\le t)$ the restriction of x to $[0,t]$. For a process X, we denote by $X(t)$ the value of X at t, and by $X_t^p=(X(u),0\le u\le t)$ its path on $[0,t]$. Consider a path $x_t^p\in D([0,t],\mathbb{R})$. For $T\ge t$ we define the *horizontal* extension $x_{t,h}^p\in D([0,T],\mathbb{R})$ of x_t^p by:

$$x_{t,h}^p(u) = \begin{cases} x(u) & \text{if } u\in[0,t] \\ x(t) & \text{if } u\in]t,t+h] \end{cases}$$

A non-anticipative functional F is a family of functionals $F=(F(.,t))_{t\in[0,T]}$where:

$$\begin{aligned} F(.,t) &: D([0,t],\mathbb{R})\to\mathbb{R} \\ x &\to F(x,t) \end{aligned} \tag{8}$$

is measurable with respect to $\mathcal{B}_t$, the canonical filtration on $D([0,t],\mathbb{R})$. A non-anticipative functional is said to be continuous at $x\in D([0,T],\mathbb{R})$ if:

$$\forall\varepsilon>0,\exists\eta>0,\forall x' \text{ then } d_\infty(x,x')<\eta\Rightarrow|F(x,t)-F(x',t)|<\varepsilon$$

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

$$x_t^{p,h}(u) = x_t^p(u) + h1[t=u]$$

The horizontal derivative at $x\in D([0,T],\mathbb{R})$ of a non-anticipative functional $F=(F(.,t))$ is defined as[1]:

$$\Delta_t F(x) = \lim_{h\to 0^+}\frac{F(x_{t,h}^p,t+h)-F(x_t^p,t)}{h} \tag{9}$$

whenever it exists. The vertical derivative of $F(.,t)$ at x is defined by:

$$\Delta_x F(x,t) = \lim_{h\to 0^+}\frac{F(x_t^{p,h},t)-F(x_t^p,t)}{h}$$

[1] We depart here from the notation in Cont and Fournié to use the original notation in Dupire. The reason is the risk of typographical confusion betwen the horizontal derivative and the Malliavin derivative.

We define $\mathbb{C}^{1,k}([0,T])$ as the set of functionals which are left-continuous, horizontally differentiable with $\Delta_t F$ continuous at fixed times, and k times vertically differentiable with $\Delta_x^k F$ left-continuous. We define $\mathbb{C}_b^{1,k}([0,T])$ as the set of functionals $F \in \mathbb{C}^{1,k}([0,T])$ such that $\Delta_t F, \Delta_x F, ..., \Delta_x^k F \in \mathbb{B}([0,T])$, the set of non-anticipative functionals which satisfy definition 2.4 in Cont and Fournié (2011).

THEOREM 4 *Suppose that a, b and G are non-anticipative $C_b^{1,\infty}([0,T])$ functionals. Suppose there exists a unique solution $v \in C_b^{1,\infty}([0,T])$ of the semilinear PPDE*

$$\begin{aligned} \Delta_t v + a(x_t^p, t)\Delta_x v + \frac{1}{2} b(x_t^p, t)\Delta_{xx} v &= 0 \\ v(x_T^p, T) &= G(x_T^p, T) \end{aligned}$$

Let X_t be a stochastic process satisfying:

$$dX(t) = a(X_t^p, t)dt + b(X_t^p, t)dW(t)$$

Suppose $G(X_T^p(\omega), T) = F(\omega)$ for some $F \in \mathbb{D}^\infty$. Then, in $L^2(P)$:

$$v(t, X_t^p(\omega)) = \mathcal{T}\exp(\frac{1}{2}\int_t^T D_s^2 ds)F(\omega^t(\omega))$$

Proof
The smoothness conditions on a, b, and G imply the smoothness conditions **(H1)** and **(H2)** in Peng and Wang (2011). By their functional Feynman-Kac formula (see also Dupire (2009)) we have:

$$v(X_t^p, t) = M(t)$$

where

$$M(t) = E[G(X_T^p, T)|\mathcal{F}_t]$$

By proposition 5.6 in Oksendal (1997), for any l and $t \in [0,T]$ we have:

$$\Delta_x^l v(X_t^p, t) = D_t^l M(t)$$

4 Conclusion and Future Work

The main result of this paper is theorem 3. Formally it contributes to building yet another (small) bridge between stochastic processes and quantum field theory. For future work, we intend to design and analyze new numerical schemes that implement the Dyson series to solve BSDEs. The main weakness of theorem 3 is that it currently requires the functional F to be infinitely Malliavin differentiable. A lot of work seems to be needed to relax some of these smoothness requirements. In the Markovian case (i.e., $F(\omega)$ being a function of $W(T,\omega)$), it is known that theorem 4 generally applies when F is only twice differentiable, as can be proved from the theory of quasilinear parabolic PDEs (see for instance McOwen (1995), chapter 11). For path-dependent functionals, the picture is much less clear.
Theorem 3 can certainly be extended to a filtration generated by several Brownian motions, and probably to Levy processes. A generalization from representation of martingales to representation of semimartingales would also be interesting. Because of the growing importance of fractional Brownian motion, an extension of this theorem to fractional Brownian motion would be desirable for applications.

5 Appendix

Following Oksendal (1997), we say that a real function $g : [0,T] \to \mathbb{R}^n$ is *symmetric* if:

$$g(x_{\sigma_1}, .., x_{\sigma_n}) = g(x_1, .., x_n)$$

for all permutations σ of $(1, 2, .., n)$. If in addition

$$||g||^2_{L^2([0,T])^n} = \int\limits_{[0,T]^n} g^2(x_1, .., x_n) dx_1..dx_n < \infty$$

Then we say that $g \in \hat{L}^2([0,T]^n)$ the space of symmetric square integrable functions on $[0,T]^n$. The Wiener chaos expansion of F:

$$F = \sum_{m=0}^{\infty} I_m(f_m) \quad \text{in } L^2(P)$$

where $\{f_m\}_{m=0}^{\infty}$ is a unique sequence of deterministic functions in $\hat{L}^2([0,T]^n)$, and

$$\begin{aligned} I_m(f_m) &= \int_0^T \int_0^{t_m} .. \int_0^{t_2} f(t_1, .., t_m) dW(t_1) dW(t_2)..dW(t_m) \text{ if } m > 0 \\ I_0(f_0) &= f_0 \end{aligned}$$

A notation we use for the Skorohod integral of H is $\int\limits_0^T H(s)\delta W(s)$.

5.1 Proof of Theorem 1

Let $t = m\Delta$ and $T = (m+1)\Delta$. We remind the reader once again of proposition 5.6 in Oksendal(1997), namely that, if $F \in \mathbb{D}_{1,2}$ and $E[F|\mathcal{F}_s] \in \mathbb{D}_{1,2}$ (see Oksendal fro a definition $\mathbb{D}_{1,2}$) then, for $t \leq s$

$$D_t(E[F|\mathcal{F}_s]) = E[D_t F|\mathcal{F}_s] \tag{10}$$

Let $t^+ = t + \varepsilon$ for a small ε. Using (10) and the Clark-Ocone formula (see, e.g. Nualart (1995)), we get, for $t^+ \leq T$ and any u:

$$\begin{aligned} E[D^l_u F|\mathcal{F}_{t^+}] &= E[D^l_u F] + \int_0^{t^+} E[D_s E[D^l_u F|\mathcal{F}_{t^+}]|\mathcal{F}_s] dW(s) &(11)\\ &= E[D^l_u F] + \int_0^{t^+} E[E[D_s D^l_u F|\mathcal{F}_{t^+}]|\mathcal{F}_s] dW(s) &(12)\\ &= E[D^l_u F] + \int_0^{t^+} E[D_s D^l_u F|\mathcal{F}_s] dW(s) &(13)\\ &= E[D^l_u F] + \int_0^{T} E[D_s D^l_u F|\mathcal{F}_s] dW(s) &(14)\\ &\quad - \int_{t^+}^{T} E[D_s D^l_u F|\mathcal{F}_s] dW(s) &(15)\\ &= E[D^l_u F|\mathcal{F}_T] - \int_{t^+}^{T} E[D_s D^l_u F|\mathcal{F}_s] dW(s) &(16) \end{aligned}$$

Since F is assumed discrete, for $s \in (t, T]$ we have:

$$D_s D^l_u F = D_T D^l_u F \tag{17}$$

Thus

$$E[D^l_u F|\mathcal{F}_{t+}] = E[D^l_u F|\mathcal{F}_T] - \int_t^T E[D_T D^l_u F|\mathcal{F}_s]dW(s) \tag{18}$$

Taking $u = T$, we obtain:

$$\begin{aligned} E[F|\mathcal{F}_{t+}] &= E[F|\mathcal{F}_T] - \int_{t+}^T E[D^1_T F|\mathcal{F}_{s_1}]dW(s_1) & (19)\\ &= E[F|\mathcal{F}_T] - [\int_{t+}^T E[D^1_T F|\mathcal{F}_T] - \int_{s_1}^T E[D^2_T F|\mathcal{F}_{s_2}]dW(s_2)\delta W(s_1)] & (20) \end{aligned}$$

where (19) follows from (18) with $l = 0$, and (20) follows from (18) with $l = 1$. We continue the expansion iteratively until we calculate the L^{th} Malliavin derivative, after which all terms are zero. Then, by continuity of martingales generated by Brownian motion, we conclude that:

$$\begin{aligned} E[F|\mathcal{F}_t] &= E[F|\mathcal{F}_T] - \int_t^T E[D^1_T F|\mathcal{F}_T] + \int_{s_1}^T E[D^2_T F|\mathcal{F}_T] - & (21)\\ &\quad .. + (-1)^L \int_{s_{L-1}}^T E[D^L_T F|\mathcal{F}_v]\delta W(s_L)..\delta W(s_1) & (22) \end{aligned}$$

where these integrals are iterated Skorohod integrals. For convenience, we define the iterated time/Skorohod integral $N(b, s_0)$. Let b be a binary vector of dimension r (which will be clear from context). We define $-b$ as the same vector without the first component. It is thus a vector of dimension $r - 1$. For instance, if $r = 5$ and:

$$b = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix} \Longrightarrow -b = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \tag{23}$$

The iterated time/Skorohod integral is:

$$N(b,t) = \begin{cases} 1 & r = 0 \\ \int_t^T N(-b,s)ds & \text{if } r \geq 1 \;\; b_r = 1 \\ \int_t^T N(-b,s)\delta W(s) & r \geq 1 \;\; b_r = 0 \end{cases} \tag{24}$$

For instance, with b as in (23) we have:

$$N(b,t) = \int_t^T \int_{s_1}^T \int_{s_2}^T \int_{s_3}^T \int_{s_4}^T \delta W(s_5)\delta W(s_4)ds_3\delta W(s_2)ds_1$$

We define

$$\begin{aligned} M_{l,0}(t) &= E_T[D^l_T F|\mathcal{F}_T] \\ M_{l,g}(t) &= \int_t^T .. \int_{s_{g-2}}^T \int_{s_{g-1}}^T E[D^l_T F|\mathcal{F}_T]\delta W(s_g)\delta W(s_{g-1})..\delta W(s_1) \qquad g > 0 \end{aligned}$$

Lemma 1.1

$$M_{l,l}(t) = \sum_{h=0}^{\min(l,L-l)} E[D_T^{l+h} F|\mathcal{F}_T] \sum_{b_1+..+b_l=h} (-1)^h N(b,t)$$

Proof of lemma 1.1
From the formula for the Skorohod of a process multiplied by a random variable ((1.49) in Nualart (1995)), we calculate:

$$\begin{aligned} M_{l,1}(s_{l-1}) &= \int_{s_{l-1}}^{T} E[D_T^l F|\mathcal{F}_T] \delta W(s_l) \\ &= \begin{cases} E[D_T^l F|\mathcal{F}_T] N(0,s_{l-1}) - E[D_T^{l+1} F|\mathcal{F}_T] N(1,s_{l-1}) & \text{if } l < L \\ E[D_T^l F|\mathcal{F}_T] N(0,s_{l-1}) & \text{if } l \geq L \end{cases} \end{aligned}$$

We suppose by induction that:

$$M_{l,g}(s_{l-g}) = \sum_{h=0}^{\min(g,n-l)} E[D_T^{l+h} F|\mathcal{F}_T] \sum_{b_1+..+b_g=h} (-1)^h N(b, s_{l-g})$$

Reapplying (1.49) in Nualart (1995), we obtain:

$$\begin{aligned} & M_{l,g+1}(s_{l-g-1}) = \qquad (25) \\ & \sum_{h=0}^{\min(g,n-l)} E[D_T^{l+h} F|\mathcal{F}_T] \int_{s_{l-g-1}}^{T} \left[\sum_{b_1+..+b_g=h} (-1)^h N(b, s_{l-g}) \right] \delta W(s_{l-g}) \\ & -E[D_T^{l+h+1} F|\mathcal{F}_T] \int_{s_{l-g-1}}^{T} \left[\sum_{b_1+..+b_g=h} (-1)^h N(b, s_{l-g}) \right] ds_{l-g} \\ & = \sum_{h=0}^{\min(g+1,n-l)} E[D_T^{l+h} F|\mathcal{F}_T] \sum_{b_1+..+b_{g+1}=h} (-1)^h N(b, s_{l-g-1}) \end{aligned}$$

■
With lemma 1.1, (21) results in:

$$\begin{aligned} E[F|\mathcal{F}_t] &= \sum_{l=0}^{L} (-1)^l M_{l,l}(t) & (26) \\ &= \sum_{l=0}^{L} \sum_{h=0}^{\min(l,L-l)} E_T[D_T^{l+h} F|\mathcal{F}_T] \sum_{b_1+..+b_l=h} (-1)^h N(b,t) & (27) \\ &= \sum_{l=0}^{L} \gamma(m,l) E[D_T^l F|\mathcal{F}_T] & (28) \end{aligned}$$

where $\gamma(m,l)$ does not depend on F or L. One possibility to calculate $\gamma(m,l)$ is to use lemma 1.1. For $l \leq L/2$ we have for instance:

$$\gamma(m,l) = (\left[\frac{l}{2}\right] + 1) \sum_{h=0}^{L} \sum_{b_1+..+b_l=h} (-1)^h N(b,t)$$

This is rather complicated. However, since (28) holds for any differentiable F, we resort to a simpler strategy. Our strategy is to vary F to determine recursively $\gamma(m,l)$. For simplicity, we take $m=0$, i.e., $t=0$. Clearly the first coefficient (take F=constant) is:

$$\gamma(0,0) = 1$$

To determine the second coefficient, $\gamma(0,L)$, with $L=1$, we choose a function F such that $D^{L+1}F = 0$. The only such function is a linear function of $W(T)$. We thus put in (28) $F = W(T)$ and calculate:

$$E_0[F] = \sum_{l=0}^{1} \gamma(0,l) D_T^l F$$

In other terms:

$$0 = \gamma(0,0)W(T) + \gamma(0,1) * 1$$

Thus

$$\gamma(0,L) = -W(T)$$

The general structure of the recursion is then:

$$\gamma(0,L) = \frac{E[F] - \sum_{l=0}^{L-1} \gamma(0,l) D_T^l F}{D_T^L F} \tag{29}$$

Clearly, formula (29) applies for any coefficient L. To calculate $\gamma(0,L)$ for $L=2$ we thus pick $F = W^2(T)$ (so that $D^{L+1}F = 0$) and obtain:

$$\begin{aligned}
\gamma(0,2) &= \frac{E[W^2(T)] - \sum_{l=0}^{1} \gamma(0,l) D_T^l W^2(T)}{D_T^2 W^2(T)} \\
&= \frac{T - W^2(T) + W(T) * 2W(T)}{2} \\
&= \frac{W^2(T) + T}{2}
\end{aligned}$$

We complete the proof by induction. Suppose that, for

$$\gamma(0,L-1) = 1\{L \text{ even}\} \left(\frac{\Delta}{2}\right)^{L/2} \frac{1}{(L/2)!} - \sum_{l=0}^{L-2} \gamma(0,l) \frac{W(T)^{L-1-l}}{(L-1-l)!} \tag{30}$$

We select $F = W^L(T)$, which we insert together with (30) in (29) to arrive at

$$\gamma(0,L) = 1\{L \text{ even}\} \left(\frac{\Delta}{2}\right)^{L/2} \frac{1}{(L/2)!} - \sum_{l=0}^{L-1} \gamma(0,l) \frac{W(T)^{L-l}}{(L-l)!} \tag{31}$$

The formula for general m obtains by replacing $W(T)$ by $W((m+1)\Delta) - W(m\Delta)$, that is choosing for test functions F above the successive powers of $W((m+1)\Delta) - W(m\Delta)$.

5.2 Proof of Theorem 2

Let $t = m\Delta$ and $T = (m+1)\Delta$. By theorem 1:

$$E[F|\mathcal{F}_t] = \hat{F}_m^L + (-1)^L \int_{s_1=t}^{T} .. \int_{s_{L+1}=s_L}^{T} E[D_{s_L}^{L+1} F|\mathcal{F}_{s_{L+1}}]\delta W(s_{L+1})...\delta W(s_1) \tag{32}$$

We want to reverse the order of integration in (32), and thus show an elementary Fubini-type theorem for Skorohod integrals from first principles. Let f_n be a deterministic function of $n+2$ variables. Suppose the function is symmetric in the first n arguments. We define the following symmetrization operators:

$$\begin{aligned}
Sym^{n,pen}(f_n)(t_1,..,t_n,v,u) &= \frac{1}{n+1}[f_n(t_1,...,t_n,v,u) + \\
&\quad f_n(t_1,...,t_{n-1},v,t_n,u) + .. + f_n(v,t_2..,t_n,t_1,u)] \\
Sym^{n,last}(f_n)(t_1,..,t_n,v,u) &= \frac{1}{n+1}[f_n(t_1,...,t_n,v,u) + \\
&\quad f_n(t_1,...,t_{n-1},u,v,t_n) + .. + f_n(u,t_2,..,t_n,v,t_1)] \\
Sym^{n+1,pen}(f_n)(t_1,..,t_n,v,u) &= \frac{1}{n+2}[f_n(t_1,...,t_n,t_{n+1},v) + \\
&\quad f_n(t_1,...,t_n,v,t_{n+1}) + .. + f_n(v,t_2,..,t_n,t_1,t_{n+1})] \\
Sym^{n+1,last}(f_n)(t_1,..,t_n,v,u) &= \frac{1}{n+2}[f_n(t_1,...,t_n,t_{n+1},u) + \\
&\quad f_n(t_1,...,t_n,u,t_{n+1}) + .. + f_n(u,t_1,..,t_n,t_{n+1},t_1)]
\end{aligned}$$

The operator $Sym^{n,pen}(f)$ sums all the permutations of any of first n variables with the penultimate variable ("v"). The operator $Sym^{n,last}(f)$ sums all the permutations of any of first n variables with the last variable("u"). The operator $Sym^{n+1,pen}(f)$ sums all the permutations of the penultimate variable ("v") with any of the other $n+1$ variables. The operator $Sym^{n+1,last}(f)$ sums all the permutations of the last variable ("u") with any of the other $n+1$ variables. The following elementary lemma translates into our notation the fact that the operations of symmetrization with respect to the penultimate variable ("v") and symmetrization with respect to the last variable ("u") are commutative. So, without proof, we state:

Lemma 2.1

$$Sym^{n+1,pen}(Sym^{n,last}(f)) = Sym^{n+1,last}(Sym^{n,pen}(f))$$

Of course, lemma 2.1 applies to

$$g(t_1,...,t_n,v,u) = f(t_1,...,t_n,v,u) * 1[v \geq u]$$

Lemma 2.2

Suppose $H(s)$ is an $\mathcal{F}_s$-adapted process. Then:

$$\int_{s_1=0}^{T} \int_{s_2=s_1}^{T} H(s_2)dW(s_2)\delta W(s_1) = \int_{s_2=0}^{T} \int_{s_1=0}^{s} H(s_2)dW(s_1)dW(s_1) \tag{33}$$

Proof of lemma 2.2

Since $H(s)$ is adapted, the iterated Ito integrals on the right of (33) are identical to Skorohod integrals, and similarly for the inner integral on the left. It is then sufficient to prove:

$$\int_{s_1=0}^{T} \int_{s_2=s_1}^{T} H(s_2)\delta W(s_2)\delta W(s_1) = \int_{s_2=0}^{T} \int_{s_1=0}^{s_2} H(s_2)\delta W(s_1)\delta W(s_2) \tag{34}$$

Let:

$$\int_{s_2=s_1}^{T} H(s_2)\delta W(s_2) = \sum_{n=0}^{\infty} I_{n+1}(h_n^1(., s_2, s_1))$$

where, for some f_n symmetric in the first n variables:

$$h_n^1(t_1, ..., t_n, s_2, s_1) = Sym^{n,pen}(f_n(t_1, ..., t_n, s_2, s_1) * 1[s_2 \geq s_1])$$

and

$$\int_{s_1=0}^{T} \int_{s_2=s_1}^{T} H(s_2)\delta W(s_2)\delta W(s_1) = \sum_{n=0}^{\infty} I_{n+2}(h_n^2(., s_1))$$

where

$$h^2(t_1, ..., t_n, s_2, s_1) = Sym^{n+1,last}(h^1(t_1, ..., t_n, s_2, s_1))$$

We see that:

$$\int_{s_1=0}^{s_2} H(s_2)\delta W(s_1) = \sum_{n=0}^{\infty} I_{n+1}(g_n^1(., s_2, s_1))$$

where

$$g_n^1(t_1, ..., t_n, s_2, s_1) = Sym^{n,last}(f_n(t_1, ..., t_n, s_2, s_1) * 1[s_2 \geq s_1])$$

Then

$$\int_{s_2=0}^{T} \int_{s_1=0}^{s_2} H(s_2)\delta W(s_1)\delta W(s_1) = \sum_{n=0}^{\infty} I_{n+2}(g_n^2(., s_2, .))$$

where

$$g_n^2(t_1, ..., t_n, s_2, s_1) = Sym^{n+1,pen}(g_n^1(t_1, ..., t_n, s_2, s_1))$$

The left handside of (34) is then equal to:

$$\sum_{n=0}^{\infty} I_{n+2}(Sym^{n+1,last}(Sym^{n,pen}(f_n(t_1, ..., t_n, s_2, s_1) * 1[s_2 \geq s_1]))) \tag{35}$$

The right handside of (34) is equal to:

$$\sum_{n=0}^{\infty} I_{n+2}(Sym^{n+1,pen}(Sym^{n,last}(f_n(t_1, ..., t_n, s_2, s_1) * 1[s_2 \geq s_1]))) \tag{36}$$

By lemma 2.1, (35) is thus equal to (36). ■

Suppose $H(s)$ is an $\mathcal{F}_s$-adapted process. Applying lemma 2.3 we obtain:

$$\int_{s_1=t}^{T} .. \int_{s_{n+1}=s_n}^{T} H(s_{n+1})\delta W(s_{n+1})..\delta W(s_1) = \int_{s_{n+1}=t}^{T} .. \int_{s_1=0}^{s_2} H(s_2)\delta W(s_1)..\delta W(s_{n+1}) \tag{37}$$

With the help of (37), with $E[D_s^{L+1}F|\mathcal{F}_s]$ replacing $H(s)$, relationship (32) becomes:

$$E[F|\mathcal{F}_t] = \hat{F}_m^L + (-1)^L \int_{s_1=t}^{T} .. \int_{s_{L+1}=t}^{s_L} E[D_{s_L}^{L+1}F|\mathcal{F}_{s_L}]dW(s_1)...dW(s_{L+1}) \tag{38}$$

By lemma 5.7.2 in Kloeden and Platen (1992), we have:

$$E[(E[F|\mathcal{F}_t] - \hat{F}_m^L)^2|\mathcal{F}_t] \leq \sup_{s\in(t,T]} E[(E[D_s^{L+1}F|\mathcal{F}_s])^2|\mathcal{F}_t] \frac{\Delta^{L+1}}{(L+1)!} \tag{39}$$

For $s \in (t, T]$, the following process is a martingale:

$$X(s) := E[D_s^{L+1}F|\mathcal{F}_s] \tag{40}$$

Thus:

$$\begin{aligned}
\sup_{s\in[t,T]} E[X(s)^2|\mathcal{F}_t] &= \sup_{s\in[t,T]} E[X(t)^2 + \int_t^s d[X,X](u)|\mathcal{F}_t] \\
&= E[X(t)^2 + \int_t^T d[X,X](u)|\mathcal{F}_t] \\
&= E[X(t)^2 + \int_t^T (E_u[D_u^{L+2}F])^2 du|\mathcal{F}_t] \\
&\leq E[X(t)^2|\mathcal{F}_t] + \Delta \sup_{u\in[t,T]} E[D_u^{L+2}F|\mathcal{F}_u]^2
\end{aligned}$$

By induction,

$$\sup_{s\in[t,T]} E[X(s)^2|\mathcal{F}_t] = \sum_{i=0}^{\infty} \Delta^i (E[D_t^{L+i}F|\mathcal{F}_t])^2$$

By assumption, all Malliavin derivatives are bounded. Thus, if $\Delta < 1$, there is a constant M so that, for all l:

$$\sup_{s\in[t,T]} E[X(s)^2|\mathcal{F}_t] \leq \frac{M}{1-\Delta} \tag{41}$$

By putting together (39), (40), and (41), we have:

$$E[(E[F|\mathcal{F}_t] - \hat{F}_m^L)^2|\mathcal{F}_t] \leq \frac{\Delta^{L+1}}{(L+1)!} \frac{M}{1-\Delta} \tag{42}$$

5.3 Proof of Theorem 4

Let $F^{(T,n)}$ be the approximation "by simple functions" of the random variable F. More precisely,we define the discretizing map $\tau^{(T,n)}$

$$\tau^{(T,n)}(t) = \begin{cases} t & \text{if} \quad t \leq T - 1/n \\ T - 1/n & \text{if} \quad T - 1/n \leq t \leq T \end{cases}$$

For each f_m we define the approximation by simple functions $f_m^{(T,n)}$

$$\begin{aligned}
f_m^{(T,n)}(t_1, t_2, .., t_m) &= f_m(\tau^{(T,n)}(t_1), \tau^{(T,n)}(t_2), .., \tau^{(T,n)}(t_m)) \quad m \geq 1 \\
f_0^{(T,n)} &= f_0
\end{aligned}$$

We now define:

$$F^{(T,n)} = \sum_{m=0}^{\infty} I_m(f_m^{(T,n)})$$

We also define $\mathcal{F}_{G^{(T,n)}}$ to be the sigma-algebra generated by Brownian motion under the condition that:

$$W(s) = W(T-1/n) \qquad T-1/n \le s \le T$$

Lemma 4.1

$$E[(F^{(T,n)} - F)^2 | \mathcal{F}_{G^{(T,n)}}] = 0$$

Proof By orthogonality of the Wiener chaos expansion, we need to prove that

$$E[(I_m(f_m^{(T,n)} - f_m))^2 | \mathcal{F}_{_{G^{(T,n)}}}] = 0$$

The case $m = 1$ spells:

$$\begin{aligned} E[\left(\int_0^T f_1^{(T,n)}(t_1) - f_1(t_1) dW(t_1)\right)^2 | \mathcal{F}_{_{G^{(T,n)}}}] &= E[\left(\int_0^{T-1/N} f_m^{(T,n)}(t_1) - f_m(t_1) dW(t_1)\right)^2] \\ &= 0 \end{aligned}$$

Let

$$\begin{aligned} X_{m-1}^{(T,n)}(t) &= \int_0^t \int_0^{t_{m-1}} .. \int_0^{t_2} f_m^{(T,n)}(t_1, .., t_{m-1}, t) dW(t_1) dW(t_2) .. dW(t_{m-1}) \\ X_{m-1}(t) &= \int_0^t \int_0^{t_{m-1}} .. \int_0^{t_2} f_m(t_1, .., t_{m-1}, t) dW(t_1) dW(t_2) .. dW(t_{m-1}) \end{aligned}$$

By induction:

$$X_{m-1}^{(T,n)}(t) = X_{m-1}(t) \ \text{ for } t \le T - 1/n$$

Thus

$$\begin{aligned} E[(I_m(f_m^{(T,n)}) - I_m(f_m))^2 | \mathcal{F}_{_{G^{(T,n)}}}] &= \\ E[(\int_0^T X_{m-1}^{(T,n)}(t_m) - X_{m-1}(t_m) dW(t_m))^2 | \mathcal{F}_{_{G^{(T,n)}}}] &= \\ E[(\int_0^{T-1/n} X_{m-1}^{(T,n)}(t_m) - X_{m-1}(t_m) dW(t_m))^2] &= 0 \end{aligned}$$

■

Lemma 4.2 For any l integer:

$$E[(D_t^l F^{(T,n)} - D_t^l F)^2 | \mathcal{F}_{G^{(T,n)}}] = 0$$

Proof

We only prove the case $l = 1$. The case of higher derivatives is similar. By lemma 4.16 in Oksendal (1997):

$$D_t F = \sum_{m=1}^{\infty} m I_{m-1}(f_m(., t))$$

and also

$$D_t F^{(T,n)} = \sum_{m=1}^{\infty} m I_{m-1}(f_m^{(T,n)}(., t))$$

Then

$$E[(I_{m-1}(f_m^{(T,n)}(.,t) - f_m(.,t)))^2|\mathcal{F}_{G^{(T,n)}}] =$$
$$E[(\int_0^T X_{m-1}^{(T,n)}(t_m) - X_{m-1}(t_m)dW(t_m))^2|\mathcal{F}_{G^{(T,n)}}] =$$
$$E[(\int_0^{T-1/n} X_{m-1}^{(T,n)}(t_m) - X_{m-1}(t_m)dW(t_m))^2] = 0$$

The convergence of $D_tF^{(T,n)}$ to D_tF follows by orthogonality of the Wiener chaos expansion. ■

Lemma 4.3

$$E[F^{(T,n)} - F|\mathcal{F}_{T-1/n}] = 0$$

Proof
By proposition 5.5 in Oksendal (1997)

$$E[I_m(f_m)|\mathcal{F}_{T-1/n}] = I_m(f_m 1_{[T-1/n]}^{\otimes n})$$

where:

$$(f_m 1_{[0,t]}^{\otimes n})(t_1,..,t_m) = f_m(t_1,..,t_m)1_{[0,T-1/n]}(t_1)...1_{[0,T-1/n]}(t_m)$$

By construction:

$$(f_m^{(T,n)} 1_{[0,t]}^{\otimes n} - f_m 1_{[0,t]}^{\otimes n})(t_1,..,t_m) = 0$$

Thus:

$$E[I_m(f_m^{(T,n)} - f_m)|\mathcal{F}_{T-1/n}] = 0$$

■

Lemma 4.4

$$\lim_{n\to\infty} E[(n(E[F^{(T,n)}|\mathcal{F}_{T-1/n}] - F^{(T,n)}) - \frac{1}{2}D_T^2F)^2|\mathcal{F}_{G^{(T,n)}}] = 0$$

Proof
By the backward Taylor expansion (theorem 3) for any n:

$$E[F^{(T,n)}|\mathcal{F}_{T-1/n}] = \sum_{l=0}^{\infty} \gamma(T-1/n,l)D_T^l F^{(T,n)} \tag{43}$$

We remind the reader that along the path $\omega^{T-1/n}(\omega)$ we have:

$$W(s,\omega^{T-1/n}(\omega)) = W(T-1/n,\omega) \qquad T-1/n \le s \le T$$

Thus:

$$\begin{aligned}
\gamma(T-1/n,1,\omega^{T-1/n}(\omega)) &= 0\\
\gamma(T-1/n,2,\omega^{T-1/n}(\omega)) &= \frac{1}{2n}\\
\gamma(T-1/n,3,\omega^{T-1/n}(\omega)) &= 0\\
\gamma(T-1/n,4,\omega^{T-1/n}(\omega)) &= \frac{1}{8n^2}\\
\gamma(T-1/n,l,\omega^{T-1/n}(\omega)) &= O(\frac{1}{n^3}) \text{ for } l > 4
\end{aligned}$$

By the Cauchy-Schwartz inequality, for $l > 4$:

$$\begin{aligned} E[\gamma(T-1/n,l)D_T^l F^{(T,n)}|\mathcal{F}_{G^{(T,n)}}] &\leq |\gamma(T-1/n,l,\omega^{T-1/n}(\omega))|\sqrt{E[(D_T^l F)^2|\mathcal{F}_{G^{(T,n)}}]} \\ &= O(\frac{1}{n^3}) \end{aligned}$$

Thus:

$$E[E[F^{(T,n)}|\mathcal{F}_{T-1/n}]^2] = E[(F^{(T,n)} + \frac{1}{2n}D_T^l F^{(T,n)} + O(\frac{1}{n^2}))^2|\mathcal{F}_{G^{(T,n)}}] \tag{44}$$

and the conclusion follows.■

Lemma 4.5

$$\lim_{n\to\infty} E[(n(E[F|\mathcal{F}_{T-1/n}] - F) - \frac{1}{2}D_T^2 F)^2|\mathcal{F}_{G^{(T,n)}}] = 0$$

Proof
We decompose the integrand into:

$$n(E[F|\mathcal{F}_{T-1/n}] - F) - \frac{1}{2}D_T^2 F = A_n + n(F - F^{(T,n)}) \tag{45}$$

with:

$$A_n = n(E[F|\mathcal{F}_{T-1/n}] - F^{(T,n)}) - \frac{1}{2}D_t^2 F$$

By the Cauchy-Schwartz inequality and lemma 4.1,

$$\begin{aligned} \lim_{n\to\infty} E[nA_n(F - F^{(T,n)})|\mathcal{F}_{G^{(T,n)}}] &\leq \lim_{n\to\infty} \sqrt{E[A_n^2|\mathcal{F}_{G^{(T,n)}}]}\sqrt{E[n^2(F - F^{(T,n)})^2|\mathcal{F}_{G^{(T,n)}}]} \\ &= \lim_{n\to\infty} \sqrt{E[A_n|\mathcal{F}_{G^{(T,n)}}]} \lim_{n\to\infty} \sqrt{E[n^2(F - F^{(T,n)})^2|\mathcal{F}_{G^{(T,n)}}]} \\ &= 0 \end{aligned}$$

Applying lemma 4.1 and developing (45), we have:

$$\lim_{n\to\infty} E[(n(E[F|\mathcal{F}_{T-1/n}] - F) - \frac{1}{2}D_T^2 F)^2|\mathcal{F}_{G^{(T,n)}}] = \lim_{n\to\infty} E[A_n^2|\mathcal{F}_{G^{(T,n)}}] \tag{46}$$

We further decompose A_n into:

$$A_n = B_n + \frac{1}{2}(D_t^2 F^{(T,n)} - D_t^2 F) \tag{47}$$

with

$$B_n = n(E[F|\mathcal{F}_{T-1/n}] - F^{(T,n)}) - \frac{1}{2}D_t^2 F^{(T,n)}$$

Again, by the Cauchy-Schwartz inequality and lemma 4.2:

$$\lim_{n\to\infty} E[B_n(D_t^2 F^{(T,n)} - D_t^2 F)|\mathcal{F}_{G^{(T,n)}}] = 0$$

Applying lemma 4.2 and developing (47), we have:

$$\lim_{n\to\infty} E[A_n^2|\mathcal{F}_{G^{(T,n)}}] = \lim_{n\to\infty} E[B_n^2|\mathcal{F}_{G^{(T,n)}}] \tag{48}$$

We further decompose B_n into:

$$B_n = C_n + nE[F - F^{(T,n)}|\mathcal{F}_{T-1/n}] \tag{49}$$

where:

$$C_n = n(E[F^{(T,n)}|\mathcal{F}_{T-1/n}] - F^{(T,n)}) - \frac{1}{2}D_t^2 F^{(T,n)}$$

Thus, by lemma 4.3:

$$\lim_{n\to\infty} C_n E[F - F^{(T,n)}|\mathcal{F}_{T-1/n}] = 0$$

Thus, by lemma 4.4:

$$\begin{aligned}\lim_{n\to\infty} E[B_n^2|\mathcal{F}_{G^{(T,n)}}] &= \lim_{n\to\infty} E[C_n^2|\mathcal{F}_{G^{(T,n)}}] \\ &= \lim_{n\to\infty} E[(n(E[F^{(T,n)}|\mathcal{F}_{T-1/n}] - F^{(T,n)}) - \frac{1}{2}D_t^2 F^{(T,n)})^2|\mathcal{F}_{G^{(T,n)}}] \\ &= 0\end{aligned} \tag{50}$$

Putting together (46),(48) and (50), the lemma obtains. ■

We now replace T by s and F by $E[F|\mathcal{F}_s]$ in lemma 4.5 to conclude that, for any $s \in [t,T]$:

$$\lim_{n\to\infty} E[(n(E[F|\mathcal{F}_{s-1/n}] - E[F|\mathcal{F}_s]) - \frac{1}{2}D_s^2 E[F|\mathcal{F}_s])^2|\mathcal{F}_{G^{(s,n)}}] = 0 \tag{51}$$

We define $\mathcal{F}_{H^{(s,T)}}$ to be the sigma-algebra generated by $\omega^t(\omega)$ (for all $\omega \in \Omega$). We clearly have:

$$\lim_{n\to\infty} E[(n(E[F|\mathcal{F}_{s-1/n}] - E[F|\mathcal{F}_s]) - \frac{1}{2}D_s^2 E[F|\mathcal{F}_s])^2|\mathcal{F}_{H^{(s-1/n,T)}}] = 0 \tag{52}$$

We write:

$$M(s,\omega) = E[F|\mathcal{F}_s](\omega) \tag{53}$$

$M(s,\omega^t(\omega))$ is $\mathcal{F}_t$-measurable for each $s \in [t,T]$. Thus, we apply Fatou's lemma to (52), so that, for any $s \in [t+\frac{1}{n},T]$:

$$\lim_{n\to\infty} n^2(M(s,\omega^{s-1/n}(\omega)) - M(s,\omega^{s-1/n}(\omega))) - \frac{1}{2}D_s^2 M(s,\omega^{s-1/n}(\omega)) = 0 \text{ in } L^2(P)$$

From now on, all equality statements for random variables are understood to be in $L^2(P)$. A fortiori, the sigma-algebra generated by $\omega^t(\omega)$ (for all $\omega \in \Omega$) is included in $\mathcal{F}_{H^{(s-1/n,T)}}$, for any $s \in [t+\frac{1}{n},T]$. Thus, with $t \le s - \Delta t$:

$$\lim_{\Delta t\to 0} \frac{M(s-\Delta t,\omega^t(\omega)) - M(s,\omega^t(\omega))}{\Delta t} = \frac{1}{2}D_s^2 M(s,\omega^t(\omega)) \tag{54}$$

$$M(T,\omega^t(\omega)) = F(\omega^t(\omega)) \tag{55}$$

The remainder of the proof is standard, but we provide it nevertheless for completeness. Let $\mathcal{H}_t$ be the Banach space of linear operators mapping a bounded $\mathcal{F}_t$ -measurable random variable to a bounded $\mathcal{F}_t$- measurable random variable in $L^2(P)$, with the operator norm. We now define the propagator $P(s,\tau) \in \mathcal{H}_s$ as a linear operator mapping $M(\tau,\omega^t(\omega))$ to $M(s,\omega^t(\omega))$ in $L^2(P)$ with M defined by (53), with $t \le s \le \tau \le T$. We have:

$$||P(t,\tau)||_{\mathcal{H}_t} = \frac{E[M^2(\tau)|\mathcal{F}_{H^{(t,T)}}]}{E[M^2(t)|\mathcal{F}_{H^{(t,T)}}]}$$

Thus $P(t,\tau)$ is bounded. Also, the semi-group property obtains. For $t \le s \le \tau \le T$:

$$P(s,\tau)P(\tau,T) = P(s,T) \tag{56}$$

Equation (54) and (55) become then:

$$\begin{aligned} \frac{\partial P(s,T)}{\partial t} &= \frac{1}{2} D_s^2 P(s,T) \quad t \leq s \leq T & (57) \\ P(T,T) &= I & (58) \end{aligned}$$

The evolution equation (57) and (58) is equivalent to the integral equation:

$$P(t,T) = I + \frac{1}{2}\int_t^T D_s^2 P(s,T) ds$$

We define the linear operator $V(t,\tau) : \mathcal{H}_t \to \mathcal{H}_t$ by:

$$V(t,\tau)P(t,\tau) = I + \frac{1}{2}\int_t^\tau D_s^2 P(s,\tau) ds$$

We now show that the fixed point equation:

$$P(t,T) = V(t,T)P(t,T) \tag{59}$$

has a unique solution. We show that $V(T-\delta,T)$ is a contraction on the space $\mathcal{H}_{T-\delta}$ for sufficiently small δ. We define the operator norm

$$||V(t,\tau)|| = \sup_{\substack{u \neq 0 \\ u \in \mathcal{H}_t}} \frac{||V(t,\tau)u||_{\mathcal{H}_t}}{||u||_{\mathcal{H}_t}}$$

We now estimate:

$$\begin{aligned} ||V(T-\delta,T)u - V(T-\delta,T)v||_\infty \quad : &= \sup_{T-\delta \leq s \leq T} ||V(s,T)u(s,T) - V(s,T)v(s,T)|| \\ &\leq \frac{1}{2} \sup \int_{T-\delta}^T ||D_s^2 u(s,T) - D_s^2 v(s,T)||_{\mathcal{H}_s} ds \end{aligned}$$

By assumption the Malliavin derivatives are bounded, i.e. there is a constant C such that:

$$||D_s^2 u(s,T) - D_s^2 v(s,T)||_{\mathcal{H}_s} \leq C||u(s,T) - v(s,T)||_{\mathcal{H}_s}$$

We conclude that (with obvious definition of the norm on the right handside):

$$||V(T-\delta,T)u - V(T-\delta,T)v||_\infty \leq C\delta ||u-v||_\infty$$

It follows that, if $\delta < 1/C$ then $V(T-\delta,T)$ is a contraction on the space of operators $\mathcal{H}_{T-\delta}$. By the contraction mapping theorem there is a unique solution to the equation $P(T-\delta,T) = V(T-\delta,T)P(T-\delta,T)$ and the solution can be constructed by the Picard iterations:

$$\begin{aligned} P_0(T-\delta,T) &= I & (60) \\ P_{n+1}(T-\delta,T) &= V(T-\delta)P_n(T-\delta,T) \end{aligned}$$

Following verbatim the argument in Zeidler (2006) p.388 and 389:

$$
\begin{aligned}
P(T-\delta,T) &= I+\sum_{k=1}^{\infty}\int_{T-\delta}^{T}..\int_{T-\delta}^{T}(\frac{1}{2})^k\mathcal{T}(D^2_{\tau_1},..D^2_{\tau_k})d\tau_1..d\tau_k \\
&= \mathcal{T}(\exp(\frac{1}{2}\int_{T-\delta}^{T}D^2_s ds)
\end{aligned}
$$

By the semi-group property (56):

$$
\begin{aligned}
P(T-2\delta,T) &= P(T-2\delta,T-\delta)P(T-\delta,T) \\
&= \mathcal{T}(\exp(\frac{1}{2}\int_{T-2\delta}^{T-\delta}D^2_s ds)\mathcal{T}(\exp(\frac{1}{2}\int_{T-\delta}^{T}D^2_s ds) \\
&= \mathcal{T}(\exp(\frac{1}{2}\int_{T-2\delta}^{T}D^2_s ds)
\end{aligned}
$$

By covering the interval $[t,T]$ by overlapping intervals of length less than $1/C$, theorem 4 obtains.

6 References

Clark, J.M.C. (1970). The Representation of Functionals of Brownian Motion by Stochastic Integrals. *Ann. Math. Statist.* 41 1282-1295.

Cont,R., and D.-A. Fournié (2010). Change of Variable Formulas for non-anticipative Functionals on Path Space. *Journal of Functional Analysis* 259, 1043-1072.

Cont, R., and D.-A. Fournié (2011). Functional Ito Calculus and Stochastic Integral Representation of Martingales. Working Paper on arXiv:1002.2446v4 [math.PR].

Dambis, K.E (1965). On the Decomposition of Continuous Martingales. Theor. Prob. Appl.10, 401-410.

Dubins, L., and G. Schwarz (1965). On Continuous Martingales. Proc. Nat. Acad. Sci. USA 53, 913-916.

Dupire, B. (2009). Functional Ito Calculus. Portfolio Research Paper 2009-04, Bloomberg.

Doob, J.L. (1953). Stochastic Processes. Wiley, New York.

Hu, Y., D. Nualart, and X. Song (2011). Malliavin Calculus for Backward Stochastic Differential Equations and Application to Numerical Solutions. *The Annals of Applied Probability* 21, 6, 2379-2423.

Mc Owen, R. (1995). Partial Differential Equations, Methods and Applications. Prentice Hall.

Kloeden, P., and E. Platen (1999). Numerical Solution of Stochastic Differential Equations. Springer.

Nualart, D. (1995).The Malliavin Calculus and Related Topics. Springer-Verlag.

Ocone, D. (1984). Malliavin Calculus and Stochastic Integral Representation of Diffusion Processes. *Stochastics* 12, 161-185.

Oksendal, B.(1997). An Introduction to Malliavin Calculus with Applications to Economics. Norwegian School of Economics and Business Administration.

Peng, S. and Wang, F. (2011). BSDE, Path-dependent PDE and Nonlinear Feynman-Kac Formula.Working Paper on arXiv:1108.4317v1 [math.PR].

Schellhorn, H., and H. Morris (2009). A Differential Tree Approach to Price Path-Dependent American Options using Malliavin Calculus. *IAENG Transactions on Engineering Technologies Volume II.* American Institute of Physics (2009).

Yosida, K. (1978). Functional Analysis. Springer-Verlag.

Zeidler, E. (2006). Quantum Field Theory 1: Basics in Mathematics and Physics. Springer-Verlag.