A REPRESENTATION THEOREM FOR SMOOTH BROWNIAN MARTINGALES

By Henry Schellhorn, Sixian Jin

Claremont Graduate University

We show that, under certain smoothness conditions, a Brownian martingale, when evaluated 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. 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 can be proved by a passage to the limit of a special case of a backward Taylor expansion of an approximating discrete-time martingale or by a density argument. The latter expansion can also be used for numerical calculations. We use our representation to solve a problem from mathematical finance.

1. Introduction. The problem of representing Brownian martingales has a long and distinguished history. Dambis [2] and Dubins-Schwarz [3] showed that continuous martingales can be represented as time-changed Brownian motions. Doob [4], 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 \(F_T\)-measurable random variable \(F\). When the random variable \(F\) is Malliavin differentiable, the Clark-Ocone formula ([1],[11]) 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 \(F_T\)-measurable random variable \(F\), which is infinitely differentiable in the sense of Malliavin. We show that a Brownian martingale, when evaluated 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 asym-
totic 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 can be 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 [15], 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 [5] 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 this paper we prove the BTE. The reader is referred to Schellhorn [14] for the proof of the exponential formula using the BTE expansion. In this paper we provide a more elegant but probably less general proof.

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 the martingale representation theorem. Three explicit examples are given, which show the usefulness of the Dyson series in analytic calculations. The last example is, as far as we know, a new result. The proofs of the theorems, as well as the notations necessary for the proofs, are relegated to the appendix.

2. Martingale Representation.

2.1. Preliminaries and notations. The uncertainty is described by a complete 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 $\mathbb{R}^d$. Given $\omega \in \Omega$, we denote by $\omega^t$ the path
that "freezes" Brownian motion after time $t$:

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

Now we state some preliminaries on Malliavin calculus and for more details we refer to [10].

Let $\mathbb{H} = L^2([0, T], \mu)$ be the separable Hilbert space of all square integrable real-valued functions on the interval $[0, T]$ with scalar product $\langle \cdot, \cdot \rangle_{\mathbb{H}}$. The norm of a deterministic function $h \in \mathbb{H}$ will be denoted by $\|h\|_{\mathbb{H}}$ and $W(h) = \int_0^T h(t) \, dW(t)$.

Let $\mathcal{S}$ denote the class of smooth random variables such that a random variable $F \in \mathcal{S}$ has the form

$$F = f(W(h_1), \ldots, W(h_n))$$

where $f \in C^\infty_p(\mathbb{R}^n)$ which is the set of all infinitely continuously differentiable function with all their partial derivatives having polynomial growth and $h_i, i = 1, \ldots, n$ belonging to $\mathbb{H}$.

The Malliavin derivative of a smooth random variable $F$ of the form (2.1) is a $\mathbb{H}$-valued random variable given by

$$DF = \sum_{i=1}^n \partial_i f(W(h_1), \ldots, W(h_n)) h_i$$

We denote $\mathbb{D}^{1,2}$ as the closure of the class of smooth random variables $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,2} = (E|F|^2 + E\|DF\|_{\mathbb{H}}^2)^{\frac{1}{2}}$$

The iterated derivative $D^lF$ defined as a random variable with values in $\mathbb{H}^{\otimes l}$. Then for any natural number $l \geq 1$, with the seminorm on $\mathcal{S}$ defined by

$$\|F\|_{l,2} = (E|F|^2 + \sum_{i=1}^l E\|D^iF\|_{\mathbb{H}^{\otimes i}}^2)^{\frac{1}{2}}$$

we denote by $\mathbb{D}^{l,2}$ the completion of the family of smooth random variables $\mathcal{S}$ with respect to the norm $\|\cdot\|_{l,2}$. For any natural number $l \geq 1$ and $F \in \mathbb{D}^{l,2}$, the Malliavin derivative of $F$

$$D^lF = \left\{ D^l_{t_1, \ldots, t_k}F, \; t_i \in [0, T] \right\}$$

\[1\] We could have written it $\omega^t(\omega)$ but for simplicity we just call it $\omega^t$. 

---
is a measurable function on the product space $[0, T] \times \Omega$, which is defined a.e. with respect to the measure $\mu^k \times P$.

We denote $D_{\infty} = \cap_{l=1}^{\infty} D_{l,2}$, to specialize the time scale we will write $D_{\infty}([0, T])$ instead below.

As is conventional for regular derivatives, the notation $D_l F(\omega^t)$ refers to the value of the Malliavin derivative of $F$ along scenario $\omega^t$, and not to the value of the composition of $F \circ \omega^t$ evaluated at $\omega$. For instance $\frac{1}{2}D_s W^2(T, \omega^t) = 1[s \leq T]W(s \wedge t)$. In what follows, all equalities between random variables are to be interpreted as almost sure equalities.

2.2. Backward Taylor Expansion.

**Theorem 2.1.** Let $F \in D_{\infty}([0, T])$ be a $\sigma\{W(\Delta), \ldots, W(T)\}$-measurable random variable. Suppose that $D_l^{(m+1)\Delta} F = 0$ for $l > L$. Let $\gamma(m, L)$ be given by

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

Then

\[
E[F|F_{m\Delta}] = \sum_{l=0}^{L} \gamma(m, l)E[D_l^{(m+1)\Delta} F|F_{((m+1)\Delta)}]
\]

We now consider a functional $F$ whose 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 $\hat{F}_m^L$ be an approximation of $E[F|F_{m\Delta}]$ obtained by supposing that $F$ has order $L$:

\[
\hat{F}_m^L = \sum_{l=0}^{L} \gamma(m, l)E[D_l^{(m+1)\Delta} F|F_{((m+1)\Delta)}]
\]

**Theorem 2.2.** Let $F \in D_{\infty}([0, T])$ be a $\sigma\{W(\Delta), \ldots, W(T)\}$-measurable random variable. The mean square truncation error is:

\[
E[(E[F|F_{m\Delta}] - \hat{F}_m^L)^2|F_{m\Delta}] = O(\Delta^{L+1})
\]
Combining theorems 2.1 and 2.2, we arrive at the backward Taylor expansion.

**Theorem 2.3.** Let $F \in \mathbb{D}_\infty([0,T])$ be a $\sigma\{W(\Delta), .., W(T)\}$-measurable random variable. Then

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

Applying (2.5) recursively, one obtains the following corollary.

**Corollary 2.1.**

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

**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 (2.9) below emerges when we choose the paths $\omega \in \Omega$:

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

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

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

For numerical applications, it is crucial to choose a low order of expansion $L$, in order to keep the number of calculations in (2.6) 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 (2.4). We leave all these considerations for future research.
2.3. **Exponential Formula.** For esthetical reasons we introduce a "chronological operator". In this we follow Zeidler [16]. 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 \(t_1', \ldots, t_{n'}\) is a permutation of \(t_1, \ldots, t_n\) such that \(t_1' \geq t_2' \geq \ldots \geq t_{n'}\).

**Example**

It is showed in Zeidler [16] p. 44-45 that

\[
\int_0^{t_1} \int_0^{t_2} H(t_1)H(t_2)dt_1dt_2 = \frac{1}{2} \int_0^t \int_0^t \mathcal{T}(H(t_1)H(t_2))dt_1dt_2
\]

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

**Definition 2.1.** The exponential operator of a time-dependent generator \(H\) is:

\[
(2.8) \quad \mathcal{T} \exp(\int_t^T H(s)ds) = \sum_{k=0}^\infty \int_t^T \ldots \int_t^T \mathcal{T}(H(\tau_1), \ldots, H(\tau_k))d\tau_1 \ldots d\tau_k
\]

In quantum field theory, the series on the right handside of \((2.8)\) is called a Dyson series [16].

**Theorem 2.4.** Suppose \(F \in \mathcal{D}_\infty(\mathbb{R}^n)\), \(t \leq T\). Then in \(L^2(P)\)

\[
(2.9) \quad E[F|\mathcal{F}_t] = \mathcal{T} \exp(\frac{1}{2} \int_t^T D_s^2 ds)F(\omega^t)
\]

The importance of the exponential formula \((2.9)\) stems from the Dyson series representation \((2.8)\), which we rewrite hereafter in a more convenient way:

\[
E[F|\mathcal{F}_t] = F(\omega^t) + \frac{1}{2} \int_t^T D_s^2 F(\omega^t)ds + \frac{1}{4} \int_t^T \int_t^T D_{s_1}^2 D_{s_2}^2 F(\omega^t)ds_2ds_1 + ...
\]

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. A
proof of the exponential formula was obtained in Schellhorn [14] by applying
the ”certainly-equivalent paths” (2.7) and approximating the continuous
functional $F \in \mathbb{D}_\infty([0, T])$ by a discrete functional as in theorem 2.1. When
$\Delta$ tends to zero (2.3) converges to (2.9). We present here a more succinct
but less general proof.

2.3.1. Solution of Some Problems by Dyson Series. We provide three dif-
erent examples where conditional expectations can be calc ulated 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 var iable
$F$ (seen
as a functional of Brownian motion) is not path-dependent. In the second
example, the functional $F$ is path-dependent. The third example, to the best
of our knowledge, is a new result.

Example 1:
Let $\tau > T$ and:

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

By observing that if

$$ f(x, y) = \frac{1}{\sqrt{\tau - y}} \exp\left(-\frac{x^2}{2(\tau - y)}\right) $$

then $\frac{\partial f}{\partial x^n} = \frac{1}{2^n} \frac{\partial f}{\partial y^n}$ and we can calculate the Dyson series:

$$ E[F|\mathcal{F}_t] = F(\omega^t) + \frac{1}{2} \int_t^T Ds_1 \frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{W^2(T)}{2(\tau - T)}\right)(\omega^t)ds_1 + .. $$

$$ = F(\omega^t) + \frac{1}{2} \int_t^T \left(\frac{W^2(t)}{(\tau - T)^{1/2}} - \frac{1}{(\tau - T)^{3/2}}\right) \exp\left(-\frac{W^2(t)}{2(\tau - T)}\right)ds_1 + .. $$

$$ = \frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{W^2(t)}{2(\tau - T)}\right) + $$

$$ \frac{1}{2} (T - t) \left(\frac{W^2(t)}{(\tau - T)^{5/2}} - \frac{1}{(\tau - T)^{3/2}}\right) \exp\left(-\frac{W^2(t)}{2(\tau - T)}\right) + .. $$

$$ = \frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{W^2(t)}{2(\tau - T)}\right) + $$

$$ (t - T) \frac{\partial}{\partial T} \left(\frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{W^2(t)}{2(\tau - T)}\right)\right) + .. $$

$$ = \frac{1}{\sqrt{\tau - T}} \exp\left(-\frac{W^2(t)}{2(\tau - T)}\right) $$
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 \( F = \exp(-\int_0^T W(u)du) \). By Ito’s formula we have

\[
\int_0^T W(u)du = \int_0^T (T-u)dW_u
\]

Thus

\[
D_s^2 F = (T-s)^2 F \\
F(\omega^t) = \exp(-\int_0^t W(u)du - W(t)(T-t))
\]

The Dyson series becomes:

\[
E[F|\mathcal{F}_t] = F(\omega^t)(1 + \frac{1}{2} \int_t^T (T-s_1)^2 ds_1 + \frac{1}{4} \int_t^T \int_{s_1}^T (T-s_1)^2(T-s_2)^2 ds_2 ds_1 + ...)
\]

\[
= \exp(-\int_0^t W(u)du - W(t)(T-t)) \sum_{n=0}^\infty \frac{1}{n!} \left( \frac{1}{2} \int_t^T (T-s)^2 ds \right)^n
\]

\[
= \exp(-\int_0^t W(u)du) \exp(-W(t)(T-t) + \frac{1}{6}(T-t)^3)
\]

Example 3: a new calculation
Let the interest rate \( r \) follow the Black-Derman-Toy model

\[
\ln r_s = e^{-bs} \ln r_0 + \frac{a(1-e^{-bs})}{b} + \sigma \int_0^s e^{-a(s-u)}dW(u)
\]

and

\[
F = \int_t^T r(s)ds
\]

We now compute the conditional expectation \( E[F|\mathcal{F}_t] \) by the exponential formula.
Firstly we denote \( r_s \) by
\[
h(s) e^{\sigma \int_0^s e^{-a(s-u)} dW(u)}
\]
where \( h(s) \) is a deterministic function
\[
h(s) = e^{-bs \ln r_0 + \frac{a(1-e^{-bs})}{b}}.
\]
The second order Malliavin derivative is
\[
D^2_{\tau_1} \int_{\tau_1}^T r_s ds = \int_{\tau_1}^T D^2_{\tau_1} r_s ds = \int_{\tau_1}^T h(s) e^{\sigma \int_0^s e^{-a(s-u)} dW(u)} \sigma^2 e^{-2a(s-\tau_1)} ds
\]
and
\[
D^2_{\tau_n} ... D^2_{\tau_1} \int_{\tau_1}^T r_s ds = \int_{\tau_1}^T h(s) e^{\sigma \int_0^s e^{-a(s-u)} dW(u)} \sigma^{2n} \prod_{j=1}^n e^{-2a(s-\tau_j)} ds
\]
where \( \tau_i \in (t, T) \) for \( i = 1, ..., n \).

Now after acting on the "frozen path" \( \omega^t \) and integrating, one obtains
\[
\frac{1}{2^n n!} \int_t^T \int_{\tau_1}^T ... \int_{\tau_{n-1}}^T \int_{\tau_n}^T h(s) e^{\sigma \int_0^s e^{-a(s-u)} dW(u)} \sigma^{2n} \prod_{j=1}^n e^{-2a(s-\tau_j)} ds d\tau_1 ... d\tau_n
\]
\[
= \frac{1}{2^n} \int_t^T \int_{\tau_{n-1}}^T ... \int_{\tau_1}^T \int_{\tau_n}^T h(s) e^{\sigma e^{-as} \int_0^t e^{au} dW(u)} \sigma^{2n} e^{-2a(s-\tau_j)} ds d\tau_1 ... d\tau_n
\]
\[
= \frac{1}{2^n n!} \int_t^T \int_{\tau_{n-1}}^T ... \int_{\tau_1}^T \int_{\tau_n}^T \sigma^{2n} \prod_{j=1}^n e^{-2a(s-\tau_j)} ds d\tau_1 \left( h(s) e^{\sigma e^{-as} \int_0^t e^{au} dW(u)} \right) ds
\]
\[
= \frac{1}{2^n n!} \int_t^T \left( \int_{\tau_{n-1}}^T \int_{\tau_1}^T \int_{\tau_n}^T \sigma^{2n} \prod_{j=1}^n e^{-2a(s-\tau_j)} ds d\tau_1 \left( h(s) e^{\sigma e^{-as} \int_0^t e^{au} dW(u)} \right) ds \right)
\]
So
\[
E[F|\mathcal{F}_t] = \sum_{n=0}^{\infty} \frac{1}{2^n n!} \int_t^T \int_{\tau_{n-1}}^T \int_{\tau_1}^T (D^2_{\tau_n} ... D^2_{\tau_1} F') (\omega^t) d\tau_1 ... d\tau_n
\]
\[
= \sum_{n=0}^{\infty} \frac{1}{2^n n!} \int_t^T \left( \int_{\tau_{n-1}}^T \int_{\tau_1}^T \int_{\tau_n}^T \sigma^{2n} \prod_{j=1}^n e^{-2a(s-\tau_j)} ds d\tau_1 \left( h(s) e^{\sigma e^{-as} \int_0^t e^{au} dW(u)} \right) ds \right)
\]
\[
= \int_t^T e^{\sigma e^{-2a(s-t)} \int_0^s \sigma^2 e^{-2a(s-u)} dW(u)} ds
\]
\[
(2.10) = \int_t^T e^{\sigma e^{-2a(s-t)} \int_0^s \sigma^2 e^{-2a(s-u)} dW(u)} ds
\]
This result has an application in mathematical finance. We show how to calculate the value function in the problem of optimal portfolio selection with logarithmic utility function and interest rates following the Black-Derman-Toy model. As far as we know this problem has not yet been considered in the literature. Korn and Kraft considered the case of a power utility function of terminal wealth $X^\pi$ (see equation (2.2) in [8]) with Vasicek interest rates, while Liu [9] considers quadratic term structure models. We consider the utility function $U(X^\pi) = \ln(X^\pi)$. In order to ease the exposition, we will consider a market with only one stock following geometric Brownian motion, with drift $b$ and volatility $\sigma$. The underlying Brownian $B$ is not necessarily independent from the Brownian motion $W$. The proportion of wealth invested in the stock (or optimal portfolio) is denoted by $\pi(t)$. Terminal wealth is then:

$$X^\pi(T) = X^\pi(0) \exp\left[\int_0^T (\pi(t)(b - r(t)) + r(t))dt + \int_0^T \pi(t)\sigma dB(t)\right]$$

According to Korn [7] p. 72, the optimal portfolio is:

$$\pi(t) = \frac{b - r(t)}{\sigma^2}$$

A straightforward application of (2.10) provides then the formula for the value function:

$$V(t) \equiv E[\ln X^\pi(T)|F_t]$$

Determining the value function is quite important if the model is generalized to include consumption. We leave this for future research.

3. Conclusion and Future Work. 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 2.3 is that it currently requires the functional $F$ to be infinitely Malliavin differentiable. Theorem 2.2 and 2.3 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.

4. Appendix. The notation we use for the Skorohod integral of $H$ is

$$\int_0^T H(s)\delta W(s).$$
4.1. Proof of Theorem 2.1. Let \( t = m\Delta \) and \( T = (m + 1)\Delta \). We remind the reader of proposition 5.6 in Oksendal [12], namely that, if \( F \in \mathbb{D}^{1,2} \) and \( E[F|\mathcal{F}_s] \in \mathbb{D}^{1,2} \) then for \( t \leq s \)

\[
D_t(E[F|\mathcal{F}_s]) = E[D_tF|\mathcal{F}_s]
\]

Using (4.1) and the Clark-Ocone formula (see, e.g. [10]), we get, for \( t \leq T \):

\[
E[D_T^l F|\mathcal{F}_t] = E[D_T^l F] + \int_0^t E[D_s E[D_T^l F|\mathcal{F}_t]|\mathcal{F}_s]dW(s)
\]

\[
= E[D_T^l F] + \int_0^t E[D_s D_T^l F|\mathcal{F}_s]dW(s)
\]

\[
= E[D_T^l F] + \int_0^T E[D_s D_T^l F|\mathcal{F}_s]dW(s)
\]

\[
- \int_t^T E[D_s D_T^l F|\mathcal{F}_s]dW(s)
\]

\[
= E[D_T^l F|\mathcal{F}_T] - \int_t^T E[D_s D_T^l F|\mathcal{F}_s]dW(s)
\]

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

\[
D_s D_T^l F = D_T D_T^l F
\]

Thus

\[
E[D_T^l F|\mathcal{F}_t] = E[D_T^l F|\mathcal{F}_T] - \int_t^T E[D_T D_T^l F|\mathcal{F}_s]dW(s)
\]

\[
+ \int_t^{t+0^+} E[(D_T - D_s) D_T^l F|\mathcal{F}_s]dW(s)
\]

Now we are going to show

\[
\int_t^{t+0^+} E[(D_T - D_s) D_T^l F|\mathcal{F}_s]dW(s) = 0
\]

For any \( \varepsilon > 0 \) arbitrarily small, denote

\[
X(t, \varepsilon) = \int_t^{t+\varepsilon} E[(D_T - D_s) D_T^l F|\mathcal{F}_s]dW(s),
\]
we show that

\[(4.5) \quad \lim_{\varepsilon \to 0} X(t, \varepsilon) = 0 \quad \text{a.s.}\]

Firstly \( E[X(t, \varepsilon)] = 0 \) follows directly,

\[(4.6) \quad E[X^2(t, \varepsilon)] = \int_t^{t + \frac{1}{n^2}} E[(D_T - D_s)D_TF|\mathcal{F}_s]^2 ds \leq C \frac{1}{n^2}\]

where \( C \) is a constant and (4.6) follows because Malliavin derivatives are bounded. Then following Chebyshev inequality and the Borel-Cantelli lemma we obtain (4.5).

Then we obtain:

\[(4.7) \quad E[F|\mathcal{F}_t] = E[F|\mathcal{F}_T] - \int_t^T E[D_TF|\mathcal{F}_{s_1}]dW(s_1)\]

\[(4.8) \quad = E[F|\mathcal{F}_T] - \int_t^T E[D^1_TF|\mathcal{F}_T]dW(s_1) - \int_t^T \int_{s_1}^T E[D^2_TF|\mathcal{F}_{s_2}]dW(s_2)dW(s_1)\]

where (4.7) follows from (4.1) with \( l = 0 \), and (4.8) follows from (4.1) with \( l = 1 \). We continue the expansion iteratively until we calculate the \( L^{th} \) Malliavin derivative, after which all terms are zero. We conclude that:

\[(4.9) \quad E[F|\mathcal{F}_t] = E[F|\mathcal{F}_T] - \int_t^T E[D^1_TF|\mathcal{F}_T]dW(s_1) + \]

\[\ldots + (-1)^L \int_t^T \int_{s_1}^T \ldots \int_{s_{L-1}}^T E[D^L_TF|\mathcal{F}_{s_L}]dW(s_L)\ldots dW(s_1)\]

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:

\[(4.10) \quad b = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix} \implies -b = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}\]

Let \( b_i \) denote the \( i - th \) component of vector \( b \).
The iterated time/Skorohod integral is:

\[
N(b, t) = \begin{cases} 
1 & r = 0 \\
\int_t^T N(-b, s) ds & r \geq 1 \ b_0 = 1 \\
\int_t^T N(-b, s) \delta W(s) & r \geq 1 \ b_0 = 0 
\end{cases}
\]

For instance, with \( b \) as in (4.10) we have:

\[
N(b, t) = \int_t^T \int_t^T \int_t^T \int_t^T \delta W(s_5) \delta W(s_4) ds_3 \delta W(s_2) ds_1
\]

We define

\[
M_{l,0}(t) = E_T[D_T^l F | \mathcal{F}_T] \\
M_{l,g}(t) = \int_t^T \int_t^T \int_t^T \int_t^T E[D_T^l F | \mathcal{F}_T] \delta W(s_g) \delta W(s_{g-1}) \delta W(s_1) \quad g > 0
\]

**Lemma 4.1.**

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

**Proof.** From the formula for the Skorohod of a process multiplied by a random variable ((1.49 in [10])), we calculate:

\[
M_{l,1}(s_{l-1}) = \int_{s_{l-1}}^T E[D_T^l F | \mathcal{F}_T] \delta W(s_l)
\]

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

We suppose by induction that:

\[
M_{l,g}(s_{l-g}) = \min(g, L-l) \sum_{h=0}^{\min(g, L-l)} E[D_T^{l+h} F | \mathcal{F}_T] \sum_{b_1+\ldots+b_g=h} (-1)^h N(b, s_{l-g})
\]
Reapplying (1.49) in [10], we obtain:

\[
M_{l,g+1}(s_{l-g-1}) = \min(g,n-l) \sum_{h=0}^{L} E[D_{T}^{l+h}F|\mathcal{F}_{T}] \int_{s_{l-g-1}}^{T} \left[ \sum_{b_{1}+\ldots+b_{g}=h} (-1)^{h} N(b, s_{l-g}) \right] \delta W(s_{l-g}) \]

\[
- E[D_{T}^{l+h+1}F|\mathcal{F}_{T}] \int_{s_{l-g-1}}^{T} \left[ \sum_{b_{1}+\ldots+b_{g}=h} (-1)^{h} N(b, s_{l-g}) \right] ds_{l-g}
\]

\[
= \sum_{h=0}^{\min(g+1,n-l)} E[D_{T}^{l+h}F|\mathcal{F}_{T}] \sum_{b_{1}+\ldots+b_{g+1}=h} (-1)^{h} N(b, s_{l-g-1})
\]

With lemma 4.1, (4.9) results in:

\[
E[F|\mathcal{F}_{T}] = \sum_{l=0}^{L} (-1)^{l} M_{l,t}(t)
\]

\[
= \sum_{l=0}^{L} \min(l,L-l) \sum_{h=0}^{L} E[D_{T}^{l+h}F|\mathcal{F}_{T}] \sum_{b_{1}+\ldots+b_{l}=h} (-1)^{h} N(b, t)
\]

\[
= \sum_{l=0}^{L} \gamma(m,l) E[D_{T}^{l}F|\mathcal{F}_{T}]
\]

where \(\gamma(m,l)\) does not depend on \(F\) or \(L\). One possibility to calculate \(\gamma(m,l)\) is to use lemma 4.1. For \(l \leq L/2\) we have for instance:

\[
\gamma(m,l) = \sum_{h=0}^{[\frac{l}{2}]} \sum_{b_{1}+\ldots+b_{l}=l-h} (-1)^{h} N(b, t)
\]

This is rather complicated. However, since (4.15) 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 \(F=\text{constant}\)) 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 (4.15) \( F = W(T) \) and calculate:

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

In other terms:

\[
0 = \gamma(0, 0)W(T) + \gamma(0, 1) \cdot 1
\]

Thus

\[
\gamma(0, L) = -W(T)
\]

The general structure of the recursion is then:

(4.16)
\[
\gamma(0, L) = \frac{E[F] - \sum_{l=0}^{L-1} \gamma(0, l) D_T^l F}{D_T^L F}
\]

Clearly, formula (4.16) 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:

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

We complete the proof by induction. Suppose that, for

(4.17)
\[
\gamma(0, L - 1) = 1\{L - 1 \text{ even}\} \left( \frac{\Delta}{2} \right)^{L-1/2} \frac{1}{(L-1/2)!} - \sum_{l=0}^{L-2} \gamma(0, l) \frac{W(T)^{L-1-l}}{(L-1-l)!}
\]
We select \( F = W^L(T) \), which we insert together with (4.17) in (4.16) to arrive at

\[
(4.18) \quad \gamma(0, L) = 1\{L \text{ even}\} \left( \frac{\Delta}{2} \right)^{L/2} \frac{1}{(L/2)!} - \sum_{l=0}^{L-1} \gamma(0, l) \frac{W(T)^{L-l}}{(L-l)!} + \Delta^2 \frac{L}{2} - L - 1 \sum_{l=0}^{L-1} \gamma(0, l) W(T)^{L-l} \]

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 \( W((m + 1)\Delta) - W(m\Delta) \).

4.2. Proof of Theorem 2.2. Let \( t = m\Delta \) and \( T = (m + 1)\Delta \). By theorem 2.1:

\[
(4.19) \quad E[F|\mathcal{F}_t] = \hat{F}_m^L + (-1)^{L+1} \sum_{s_1=t}^{T} \int_{s_{L+1}=s_L}^{T} E[D_{T}^{L+1}F|\mathcal{F}_{s_{L+1}}] \delta W(s_{L+1}) \ldots \delta W(s_1)
\]

Next we need to bound

\[
(4.20) \quad E[(\int_{s_1=t}^{T} \int_{s_{L+1}=s_L}^{T} E[D_{T}^{L+1}F|\mathcal{F}_{s_{L+1}}] \delta W(s_{L+1}) \ldots \delta W(s_1))^2]
\]

Define:

\[
H(L + 1, s_{L+1}) = E[D_{T}^{L+1}F|\mathcal{F}_{s_{L+1}}]
\]

\[
H(n, s_n) = \int_{s_{n+1}=s_n}^{T} H(n + 1, s_{n+1}) \delta W(s_{n+1}) \quad 0 \leq n \leq L
\]

Thus:

\[
E[F|\mathcal{F}_t] - \hat{F}_m^L = H(0, t)
\]
By relation (1.48) in [10] p.39:

\[
E[H(n, s_n)^2] = \int_{s_{n+1} = s_n}^{T} E[H(n + 1, s_{n+1})^2] ds_{n+1} + \\
\int_{s=s_n}^{T} \int_{u=u_n}^{T} \frac{1}{T} E[D_s E[D_{s+1}^F|\mathcal{F}_u] D_u E[D_{s+1}^F|\mathcal{F}_s] dsdu \\
\leq \sup_{s_{n+1} \in [s_n, T]} E[H(n + 1, s_{n+1})^2](T - s_n) + o(T - s_n)
\]

Since \( F \in D_{\infty}([0, T]) \), for all \( s_{L+1} \), there exists a constant \( K \), such that

\[
E[(E[D_{s+1}^F|\mathcal{F}_{s_{L+1}})^2] \leq E[(E[D_{s+1}^F)^2|\mathcal{F}_{s_{L+1}})] \\
= E[(D_{s+1}^F)^2] \\
\leq K
\]

Thus:

\[
E[H(L, s_L)^2] \leq K(T - s_L) + o(T - s_L) = O(\Delta)
\]

and, by induction:

\[
E[H(0, t)^2] = O(\Delta^{L+1})
\]

4.3. Proof of Theorem 2.4. In this proof the equivalences are all in \( L^2 \) sense. Following lemma 1.1.2 in [10], we know \( \varepsilon = \{ e^{\int_0^T f(s)dW(s)} | f \in L^2 \} \) is a total subset of \( L^2 [0, T] \). Then we only need to show (2.9) satisfies for any function \( \varepsilon (f) \in \varepsilon \) with the form

\[
\varepsilon (f) = e^{\int_0^T f(s)dW(s)}
\]

Since \( \int_0^T f(s)dW(s) \) is Gaussian

\[
(4.21) \quad E[\varepsilon (f) | \mathcal{F}_t] = e^{\int_0^t f(s)dW(s) + \frac{1}{2} \int_0^t (f(s))^2 ds}
\]

While we have

\[
\varepsilon (f)(\omega^t) = e^{\int_0^t f(s)dW(s)}(\omega^t) = e^{\int_0^t f(s)dW(s)}
\]

and

\[
D_s^2 \varepsilon (f) = \varepsilon (f)(f(s))^2
\]
where \( s \in (t, T) \). Then
\[
\frac{1}{2^i i!} \int_{[t, T]} (D^2_{s_i} \ldots D^2_{s_1} \varepsilon (f)) (\omega^i) \ ds_i \ldots ds_1
\]
\[
= \varepsilon (f) (\omega^i) \frac{1}{2^i i!} \int_{[t, T]} (f (s_1))^2 \ldots (f (s_i))^2 ds_i \ldots ds_1
\]
\[
= e^\int_0^1 f(s) dW(s) \frac{1}{i!} \left( \frac{1}{2} \int_t^T (f (s))^2 ds \right)^i
\]
so
\[
\left( e^{\frac{1}{2} \int_t^T D^2_{s} ds} \right) \varepsilon (f) (\omega^i) = e^\int_0^1 f(s) dW(s) \sum_{i=0}^\infty \frac{1}{i!} \left( \frac{1}{2} \int_t^T (f (s))^2 ds \right)^i
\]
\[
= e^\int_0^1 f(s) dW(s) + \frac{1}{2} \int_t^T (f(s))^2 ds
\]
\[
= E [\varepsilon (f) | \mathcal{F}_t]
\]

Linearity is obvious. Indeed for \( f \) and \( g \in L^2 \) and \( \alpha \) and \( \beta \) be scalars, it is obvious that
\[
\alpha \left( e^{\frac{1}{2} \int_t^T D^2_{s} ds} \right) \varepsilon (f) (\omega^i) + \beta \left( e^{\frac{1}{2} \int_t^T D^2_{s} ds} \right) \varepsilon (g) (\omega^i) =
\]
\[
\alpha E [\varepsilon (f) | \mathcal{F}_t] + \beta E [\varepsilon (f) | \mathcal{F}_t]
\]
\[
\square
\]

Acknowledgements. We thank Qidi Peng for reviewing our manuscript.

References.
[1] CLARK, J.M.C. (1970). The Representation of Functionals of Brownian Motion by Stochastic Integrals. Ann. Math. Statist 41 1282-1295.
[2] Dambis, K.E. (1965). On the Decomposition of Continuous Martingales. Theor. Prob. Appl. 10, 401-410.
[3] Dubins, L. and Schwarz,G. (1965). On Continuous Martingales. Proc. Nat. Acad. Sci. USA 53, 913-916.
[4] Doob, J.L. (1953). Stochastic Processes. Wiley, New York.
[5] Hu, Y., Nualart, D. and Song, X. (2011). Malliavin Calculus for Backward Stochastic Differential Equations and Application to Numerical Solutions. The Annals of Applied Probability 21, 6, 2379-2423.
[6] Kloeden, P. and Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*. Springer.

[7] Korn, R. (1997). *Optimal Portfolios*. World Scientific.

[8] Korn, R. and Kraft, H. (2001). A Stochastic Control Approach to Portfolio Problems with Stochastic Interest rates. *SIAM Journal of Control and Optimization*, vol. 40, no 4, 1250-1269.

[9] Liu, J. (2007). Portfolio Selection in Stochastic Environments. *Review of Financial Studies*. vol 20, no 1, 1-39.

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

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

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

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

[14] Schellhorn, H. (2011). A Representation Theorem for Smooth Brownian Martingales. Working Paper on arXiv:1205.0295 [math.PR].

[15] Schellhorn, H. and Morris, H. (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.

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