CHAPTER VI - THE MALLIAVIN CALCULUS AND APPLICATIONS

JOSEPH G. CONLON

1. INTRODUCTION

We give an introduction here to the Malliavin Calculus and some applications in mathematical finance, in particular to the numerical estimation by Monte-Carlo methods of the so called Greeks-the $\Delta, \rho,$ etc of options. The Greeks are derivatives with respect to some parameter of expectation values of a function of a Gaussian random variable. Therefore, while it is straightforward to write a MC algorithm for the expectation value, it is not so easy to write a reliable algorithm for a derivative.

Our starting point is the Martingale representation theorem. Let $B(t), t \geq 0,$ be Brownian motion and $\mathcal{F}_t$ be the $\sigma-$field generated by $B(s), s \leq t.$ For $0 \leq t \leq T$ let $M_t$ be $\mathcal{F}_t$ measurable and satisfy the conditions

\begin{align}
1.1 \quad E[|M_t|] < \infty, \quad E[M_t | \mathcal{F}_s] = M_s \quad \text{for} \quad 0 \leq s \leq t \leq T.
\end{align}

Then $M_t, 0 \leq t \leq T,$ is said to be a Martingale with respect to the Brownian motion $B(\cdot).$ The Martingale is square integrable if in addition

\begin{align}
1.2 \quad E[M_t^2] < \infty \quad \text{for} \quad 0 \leq t \leq T.
\end{align}

Note that (1.1) implies that the LHS of (1.2) is an increasing function of $t.$ In fact we have

\begin{align}
1.3 \quad E[M_t^2] = E[E[M_t^2 | \mathcal{F}_s]] \geq E[E[M_t | \mathcal{F}_s]^2] = E[M_s^2] \quad \text{for} \quad 0 \leq s \leq t,
\end{align}

where we have used Jensen’s inequality.

**Theorem 1.1** (Martingale Representation Theorem). Suppose $M_t, 0 \leq t \leq T,$ is a square integrable Martingale with respect to Brownian motion $B(\cdot).$ Then there exists for $0 \leq s \leq T$ a square integrable predictable process $\sigma_s$ i.e. $\sigma_s$ is $\mathcal{F}_s$ measurable, such that

\begin{align}
1.4 \quad M_t = M_0 + \int_0^t \sigma_s \, dB(s), \quad E[(M_t - M_0)^2] = E \left[ \int_0^t \sigma_s^2 \, ds \right].
\end{align}

**Proof.** We shall assume wlog that $M_0 \equiv 0,$ and consider the Hilbert space $L^2(\Omega, \mathcal{F}_t, P)$ of square integrable functions which are $\mathcal{F}_t$ measurable. Consider now for any $0 = t_0 < t_1 < t_2, \cdots < t_n = t$ and $\lambda_0, \lambda_1, \ldots, \lambda_{n-1} \in \mathbb{R},$ the function

\begin{align}
1.5 \quad \exp \left[ \sum_{j=1}^{n} \lambda_{j-1} \{ B(t_j) - B(t_{j-1}) \} \right] \in L^2(\Omega, \mathcal{F}_t, P).
\end{align}
It is clear that the closure of the linear span of all such functions (1.5) is the space $L^2(\Omega, \mathcal{F}_t, P)$. Next we consider functions of the form

$$\sum_{j=1}^{n} \sigma_j(\cdot) \{B(t_j) - B(t_{j-1})\}, \quad \sigma_k(\cdot) \in L^2(\Omega, \mathcal{F}_t, P), \quad 0 \leq k < n.$$  

These functions are in $L^2(\Omega, \mathcal{F}_t, P)$ with norm

$$\left( \sum_{j=1}^{n} \sigma_j(\cdot) \{B(t_j) - B(t_{j-1})\} \right)^2 = \sum_{j=1}^{n} (\sigma_j(\cdot))^2 \{t_j - t_{j-1}\}.$$  

We shall show that the closure of the linear span of all such functions (1.6) in $L^2(\Omega, \mathcal{F}_t, P)$ is the subspace of $L^2(\Omega, \mathcal{F}_t, P)$ of functions $f(\cdot)$ satisfying $\langle f(\cdot) \rangle = 0$. Since $M_t \in L^2(\Omega, \mathcal{F}_t, P)$ and $\langle M_t \rangle = 0$ the result follows.

For a function $\lambda : [0, t] \to \mathbb{R}$ we consider the stochastic integral $X(s), \quad 0 \leq s \leq t,$ defined by

$$X(s) = \int_0^s \lambda(s') \, dB(s') - \frac{1}{2} \int_0^s \lambda(s')^2 \, ds'.$$

Evidently $X(s)$ is the solution to the SDE

$$dX(s) = \lambda(s) dB(s) - \frac{1}{2} \lambda(s)^2 ds, \quad 0 \leq s \leq t, \quad X(0) = 0.$$  

If $g : \mathbb{R} \to \mathbb{R}$ is a $C^2$ function then Ito’s formula gives

$$dg(X(s)) = g'(X(s)) dX(s) + \frac{1}{2} g''(X(s)) \lambda(s)^2 ds.$$  

Taking $g(x) = e^x$, we conclude from (1.9), (1.10) that if $N(s) = e^{X(s)}$, then

$$dN(s) = N(s) \lambda(s) dB(s) \quad \text{implies } N(t) = 1 + \int_0^t N(s) \lambda(s) dB(s).$$  

Evidently $E[N(t)] = 1$, but it is also easy to obtain an expression for $E[N(t)^2]$ by observing similarly to (1.11) that

$$d[N(s)^2] = 2N(s)^2 \lambda(s) dB(s) + N(s)^2 \lambda(s)^2 ds.$$  

Taking the expectation in (1.12), we see that

$$d\langle N(s)^2 \rangle = \lambda(s)^2 \langle N(s)^2 \rangle, \quad \langle N(0)^2 \rangle = 1,$$

and hence $\langle N(t)^2 \rangle$ is given by the formula

$$\langle N(t)^2 \rangle = \exp \left[ \int_0^t \lambda(s)^2 \, ds \right].$$  

Now let us take $\lambda(\cdot)$ to be a simple function

$$\lambda(s) = \sum_{j=1}^{n} \lambda_{j-1} \chi_{[t_{j-1}, t_j]}(s), \quad 0 \leq s \leq t.$$  

Then the function (1.5) is the same as

$$\exp \left[ \frac{1}{2} \int_0^t \lambda(s)^2 \, ds \right] N(t).$$  

We also have from (1.11), (1.14) that $N(t) - 1$ is in the closure of the linear span of functions (1.6) in $L^2(\Omega, \mathcal{F}_t, P)$. We have therefore shown that the closure of
the linear span of functions (1.6) in \( L^2(\Omega, \mathcal{F}_t, P) \) is the subspace of \( L^2(\Omega, \mathcal{F}_t, P) \) of functions \( f(\cdot) \) satisfying \( \langle f(\cdot) \rangle = 0 \).

We consider a function \( \xi(\cdot) \in L^2(\Omega, \mathcal{F}_T, P) \) and define \( M_t, \ 0 \leq t \leq T \), by

\[
M_t = E[\xi(\cdot) \mid \mathcal{F}_t].
\]

It is clear that \( M_t, \ 0 \leq t \leq T \), is a square integrable Martingale and hence by Theorem 1.1 there is a square integrable predictable process \( \sigma_t, \ 0 \leq t \leq T \), such that (1.4) holds. We can write (1.4) as

\[
\xi(\cdot) = \langle \xi(\cdot) \rangle + \int_0^T \sigma_t(\cdot) \, dB(t).
\]

The Ito calculus, which we have already used, gives us a stochastic version of the fundamental theorem of calculus. Equation (1.18) seems like another stochastic version of the fundamental theorem of calculus. We shall see that this is the case by realizing \( \sigma_t(\cdot) \) as

\[
\sigma_t(\cdot) = E[D_t \xi(\cdot) \mid \mathcal{F}_t],
\]

where \( D_t \xi(\cdot) \) is the Malliavin derivative of \( \xi(\cdot) \) at time \( t \).

To see what we mean by the Malliavin derivative we recall some basics of multivariable calculus. For a \( C^1 \) function \( \psi : \mathbb{R}^n \to \mathbb{R} \), its gradient \( D\psi(\cdot) \) is a vector field \( D\psi : \mathbb{R}^n \to \mathbb{R}^n \), which we can write as

\[
D\psi(x) = \begin{bmatrix}
\frac{\partial \psi(x)}{\partial x_1}, & \ldots, & \frac{\partial \psi(x)}{\partial x_n}
\end{bmatrix}, \quad x = (x_1, \ldots, x_n).
\]

The directional derivative of \( \psi(\cdot) \) in direction \( v = (v_1, \ldots, v_n) \in \mathbb{R}^n \) is then given by the formula

\[
D_v \psi(x) = \lim_{\varepsilon \to 0} \frac{\psi(x + \varepsilon v) - \psi(x)}{\varepsilon} = [D\psi(x), v] = \sum_{j=1}^n \frac{\partial \psi(x)}{\partial x_j} v_j,
\]

where \([\cdot, \cdot]\) denotes the Euclidean inner product on \( \mathbb{R}^n \). If \( \psi(x) = [x, w] \) is a linear function of \( x \), then \( D\psi(x) = w \in \mathbb{R}^n \), and the directional derivative of \( \psi(\cdot) \) in direction \( w \) is 1 if the Euclidean norm \( |w| = 1 \), and the directional derivative of \( \psi(\cdot) \) in directions orthogonal to \( w \) is 0. To generalize these considerations to the space of functions \( L^2(\Omega, \mathcal{F}_T, P) \), we replace the Euclidean space \( \mathbb{R}^n \) by \( L^2([0, T]) \) and the linear functions in \( L^2(\Omega, \mathcal{F}_T, P) \) by the functions

\[
\int_0^T h(s) \, dB(s) = \int_0^T h(s)W(s) \, ds, \quad h(\cdot) \in L^2([0, T]),
\]

where \( W(\cdot) \) is the white noise Gaussian process. Recalling that the Euclidean inner product on \( L^2([0, T]) \) is

\[
[h_1(\cdot), h_2(\cdot)] = \int_0^T h_1(s)h_2(s) \, ds, \quad h_1, h_2 \in L^2([0, T]),
\]

we see that the function (1.22) is simply \( [h(\cdot), W(\cdot)] \). It is clear then how we should define the gradient \( D\xi(\cdot) \) of a linear function (1.22),

\[
\xi(\cdot) = \int_0^T h(s) \, dB(s) \implies D\xi(\cdot) = h(\cdot) \in L^2([0, T])
\]

In equation (1.24) we have defined the Malliavin derivative of a linear function of white noise.
Observe that what we are doing here is identifying the Hilbert space $L^2(\Omega, \mathcal{F}_T, P)$ with a Hilbert space of functions $\xi : L^2([0, T]) \rightarrow \mathbb{R}$. The identification is made through the relationship

\begin{equation}
\phi(\cdot) \leftrightarrow W(\cdot) \text{ for } \phi(\cdot) \in L^2([0, T]) .
\end{equation}

Thus the linear function (1.22) corresponds to the function $\xi : L^2([0, T]) \rightarrow \mathbb{R}$ defined by $\xi(\phi(\cdot)) = [\phi(\cdot), h(\cdot)]$. The function (1.5) corresponds to the function $\xi : L^2([0, T]) \rightarrow \mathbb{R}$ defined by

\begin{equation}
\xi(\phi(\cdot)) = \exp \left[ \sum_{j=1}^{n} \lambda_{j-1} \int_{t_{j-1}}^{t_j} \phi(s) \, ds \right] .
\end{equation}

To complete the identification we need to introduce a measure on $L^2([0, T])$ which corresponds to the Wiener measure of $(\Omega, \mathcal{F}_t, P)$. We choose an orthonormal basis $\phi_j(\cdot)$, $j = 1, 2, \ldots$, of $L^2([0, T])$ which gives a coordinate system $x = (x_1, x_2, \ldots)$ defined by

\begin{equation}
\phi(\cdot) = \sum_{j=1}^{\infty} x_j \phi_j(\cdot) , \quad \text{where } \|\phi(\cdot)\|^2 = |x|^2 = \sum_{j=1}^{\infty} x_j^2 < \infty .
\end{equation}

Then $\xi(\cdot)$ is a function of the variables $x = (x_1, x_2, \ldots)$. The measure now on $L^2([0, T])$ is simply the one for which the coordinates $(x_1, x_2, \ldots)$ are i.i.d. standard normal variables. In particular if $\xi(\cdot)$ is just a function of a finite number $x_1, \ldots, x_n$ of the variables, then

\begin{equation}
\langle \xi(\cdot) \rangle = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \xi(x_1, \ldots, x_n) \exp \left[ -\sum_{j=1}^{n} x_j^2/2 \right] dx_1 \cdots dx_n .
\end{equation}

It is easy to verify that the measure (1.28) gives the same value for the expectation of the function (1.26) as the Wiener measure gives for the function (1.5).

For a function $\xi : L^2([0, T]) \rightarrow \mathbb{R}$ we can define the directional derivative similarly to (1.21) by

\begin{equation}
D_h \xi(\phi(\cdot)) = \lim_{\varepsilon \to 0} \frac{\xi(\phi(\cdot) + \varepsilon h(\cdot)) - \xi(\phi(\cdot))}{\varepsilon} = [D \xi(\phi(\cdot)), h] = \int_{0}^{T} D_t \xi(\phi(\cdot)) h(t) \, dt , \quad h(\cdot) \in L^2([0, T]) .
\end{equation}

Equation (1.29) defines the Malliavin derivative $D_t \xi(\cdot)$ as referred to in (1.19).

### 2. Integration by parts formulae

Let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a $C^2$ function. The first derivative $V'(x) \in \mathbb{R}^n$ of $V(\cdot)$ at $x \in \mathbb{R}^n$ is an $n$ dimensional vector, and the second derivative $V''(x)$ is a symmetric $n \times n$ matrix. We shall assume $V(\cdot)$ is a convex function and that there is a $\lambda > 0$ such that

\begin{equation}
V''(x) \geq \lambda I_n \quad \text{in the quadratic form sense for } x \in \mathbb{R}^n .
\end{equation}

We consider the probability measure $P$ on $\mathbb{R}^n$ with density

\begin{equation}
\exp[-V(x)]/\text{normalization} .
\end{equation}
From (2.1) we see that any exponential function \( \psi : \mathbb{R}^n \to \mathbb{R} \) of the form \( \psi(x) = \exp[a \cdot x] \) for some fixed \( a \in \mathbb{R}^n \), is integrable with respect to \( P \). Suppose now that \( \psi : \mathbb{R}^n \to \mathbb{R} \) is a \( C^1 \) function which satisfies

\[
|\psi(x)| + |D\psi(x)| \leq A\exp[B|x|], \quad x \in \mathbb{R}^n,
\]

for some constants \( A, B \).

We also let \( g : \mathbb{R}^n \to \mathbb{R}^n \) be a \( C^1 \) vector field which satisfies a similar inequality,

\[
|g(x)| + |Dg(x)| \leq A\exp[B|x|], \quad x \in \mathbb{R}^n,
\]

for some constants \( A, B \).

Then if \( \langle \cdot, \cdot \rangle \) denotes expectation with respect to the measure (2.2), it is easy to see from integration by parts that

\[
\langle [D\psi(\cdot), g(\cdot)] \rangle = \langle \psi(\cdot)D^*g(\cdot) \rangle,
\]

where the divergence operator \( D^* \) takes a vector field \( g : \mathbb{R}^n \to \mathbb{R}^n, \ g(x) = [g_1(x), \ldots, g_n(x)] \) to a scalar field \( D^*g : \mathbb{R}^n \to \mathbb{R} \),

\[
D^*g(x) = \sum_{j=1}^n \left[ -\frac{\partial g_j(x)}{\partial x_j} \right] + \frac{\partial V(x)}{\partial x_j} g_j(x).
\]

Observe that for any orthonormal basis \( v_1, \ldots, v_n \) of \( \mathbb{R}^n \) one has

\[
\sum_{j=1}^n \frac{\partial V(x)}{\partial x_j} g_j(x) = [V'(x), g(x)],
\]

\[
\sum_{j=1}^n \frac{\partial g_j(x)}{\partial x_j} = \sum_{j=1}^n D_{v_j}[v_j, g(x)],
\]

where \([\cdot, \cdot]\) is the Euclidean inner product on \( \mathbb{R}^n \) and \( D_v \) denotes the directional derivative (1.21). We conclude that

\[
D^*g(x) = -\sum_{j=1}^n D_{v_j}[v_j, g(x)] + [V'(x), g(x)].
\]

In generalizing (2.6), (2.8) to the infinite dimensional case we shall be concerned with vector fields \( g : \mathbb{R}^n \to \mathbb{R}^n \) for which the first summation in (2.6) or (2.8) is zero. This can happen for example in (2.6) when

\[
g_j(x) \text{ is only a function of the variables } x_1, \ldots, x_{j-1}, \ j = 1, \ldots, n.
\]

Evidently (2.9) implies that

\[
\frac{\partial g_j(x)}{\partial x_j} = 0, \ j = 1, \ldots, n, \ \text{whence } D^*g(x) = [V'(x), g(x)].
\]

In the infinite dimensional case we are concerned with vector fields \( G : L^2([0, T]) \to L^2([0, T]) \). As in (1.27) we choose an orthonormal basis \( \phi_j(\cdot), \ j = 1, 2, \ldots, \) for \( L^2([0, T]) \) for which the Wiener measure has the realization (1.28). In that case

\[
[V'(\phi), G(\phi)] = \sum_{j=1}^\infty x_j[\phi_j, G(\phi)] = \sum_{j=1}^\infty [\phi, \phi_j][\phi_j, G(\phi)] = [\phi, G(\phi)],
\]

which we can write as

\[
[V'(\phi), G(\phi)] = \int_0^T G_s(\phi) \phi(s) \, ds \leftrightarrow \int_0^T G_s(W(\cdot)) W(s) \, ds = \int_0^T G_s(W(\cdot)) dB(s),
\]
where we have made the identification (1.25). In (2.12) we are denoting \( G(\phi) \in L^2([0, T]) \) as the function \( s \rightarrow G_s(\phi), \ 0 \leq s \leq T. \)

We restrict now the vector fields \( G : L^2([0, T]) \rightarrow L^2([0, T]) \) to be predictable, which means as in §1 that \( G_s(\phi(\cdot)) \) is a function only of the variables \( \phi(s') \) with \( s' \leq s \). The vector field corresponding to (1.6) is given by

\[
G_s(\phi(\cdot)) = \sigma_{j-1}(\phi(\cdot)), \quad t_{j-1} < s \leq t_j, \quad j = 1, \ldots, n,
\]

where the functions \( \sigma_j(\phi(\cdot)) \) in (2.13) depend only on the variables \( \phi(t) \) for \( t \leq t_j \).

In fact all we do is construct for each \( j = 1, \ldots, n \), an orthonormal basis \( \phi_j, \ j = 1, 2, \ldots, \) of \( L^2([0, T]) \) such that the vector field (2.13) satisfies

\[
D_{\phi_j}[\phi_j, G(\phi)] = 0, \quad j = 1, 2, \ldots
\]

We conclude from (2.16), (2.17) that

\[
\langle [D\xi(\cdot), G(\cdot)] \rangle = \langle \xi(\cdot)D^*G(\cdot) \rangle = \langle \xi(\cdot)\int_0^T G_s(\phi)(s) \, ds \rangle.
\]

From (1.18) it follows that

\[
\langle \int_0^T G_s(\phi)(s) \, ds \rangle = \langle \int_0^T \sigma_s(\phi)(s)ds \int_0^T G_s(\phi)(s)ds \rangle
\]

\[
= \langle \int_0^T \sigma_s(\cdot)dB(s) \int_0^T G_s(\cdot)dB(s) \rangle = \langle \int_0^T \sigma_s(\cdot)G_s(\cdot)ds \rangle.
\]

We conclude from (2.16), (2.17) that

\[
\langle \int_0^T D_s\xi(\cdot)G_s(\cdot) \, ds \rangle = \langle \int_0^T \sigma_s(\cdot)G_s(\cdot)ds \rangle.
\]

It follows that

\[
\langle \int_0^T \{E[D_s\xi(\cdot)|\mathcal{F}_s] - \sigma_s(\cdot)\}G_s(\cdot) \, ds \rangle = 0
\]

for all predictable \( G : L^2([0, T]) \rightarrow L^2([0, T]) \). Since \( E[D_s\xi(\cdot)|\mathcal{F}_s] - \sigma_s(\cdot) \) itself is predictable, we can take \( G(\cdot) \) to be equal to it, whence (2.19) yields the equation

\[
\langle \int_0^T G_s(\cdot)^2ds \rangle = 0.
\]

We conclude that \( E[D_s\xi(\cdot)|\mathcal{F}_s] - \sigma_s(\cdot) = 0 \) with probability 1.

Proof of Clark-Okone formula (1.19). We have from (2.5), (2.15) that for any predictable \( G : L^2([0, T]) \rightarrow L^2([0, T]), \)

\[
\langle [D\xi(\cdot), G(\cdot)] \rangle = \langle \xi(\cdot)D^*G(\cdot) \rangle = \langle \xi(\cdot)\int_0^T G_s(\phi)(s) \, ds \rangle.
\]

From (1.18) it follows that

\[
\langle \int_0^T G_s(\phi)(s) \, ds \rangle = \langle \int_0^T \sigma_s(\phi)(s)ds \int_0^T G_s(\phi)(s)ds \rangle
\]

\[
= \langle \int_0^T \sigma_s(\cdot)dB(s) \int_0^T G_s(\cdot)dB(s) \rangle = \langle \int_0^T \sigma_s(\cdot)G_s(\cdot)ds \rangle.
\]

We conclude from (2.16), (2.17) that

\[
\langle \int_0^T D_s\xi(\cdot)G_s(\cdot) \, ds \rangle = \langle \int_0^T \sigma_s(\cdot)G_s(\cdot)ds \rangle.
\]

It follows that

\[
\langle \int_0^T \{E[D_s\xi(\cdot)|\mathcal{F}_s] - \sigma_s(\cdot)\}G_s(\cdot) \, ds \rangle = 0
\]

for all predictable \( G : L^2([0, T]) \rightarrow L^2([0, T]) \). Since \( E[D_s\xi(\cdot)|\mathcal{F}_s] - \sigma_s(\cdot) \) itself is predictable, we can take \( G(\cdot) \) to be equal to it, whence (2.19) yields the equation

\[
\langle \int_0^T G_s(\cdot)^2ds \rangle = 0.
\]

We conclude that \( E[D_s\xi(\cdot)|\mathcal{F}_s] - \sigma_s(\cdot) = 0 \) with probability 1.

We wish to compare the Ito fundamental theorem of calculus to the Malliavin fundamental theorem. Recall that Ito’s fundamental theorem yields for a \( C^2 \) function \( f : \mathbb{R} \rightarrow \mathbb{R} \) the identity

\[
f(B(T)) = f(B(0)) + \int_0^T f'(B(s)) \, dB(s) + \frac{1}{2} \int_0^T f''(B(s)) \, ds.
\]
To get the corresponding Malliavin formula, we write
\[ f(B(T)) = f\left( \int_0^T W(s) \, ds \right). \]

Hence we consider the function \( \xi : L^2([0, T]) \to \mathbb{R} \) defined by
\[ \xi(\phi) = f\left( \int_0^T \phi(s) \, ds \right). \]

For \( h \in L^2([0, T]) \), the directional derivative \( D_h \xi(\phi) \) is given by
\[ D_h \xi(\phi) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[ f\left( \int_0^T \phi(s) + \varepsilon h(s) \, ds \right) - f\left( \int_0^T \phi(s) \, ds \right) \right] = f'\left( \int_0^T \phi(s) \, ds \right) \int_0^T h(s') \, ds'. \]

Comparing with (1.29), we conclude that
\[ D_t \xi(\phi) = f'\left( \int_0^T \phi(s) \, ds \right) = f'(B(T)) \quad \text{for} \quad 0 \leq t \leq T. \]

The Clark-Okone formula then yields the equation
\[ f(B(T)) = \langle f(B(T)) \rangle + \int_0^T E[f'(B(T))|\mathcal{F}_s] \, dB(s). \]

Note that (2.26) involves only the first derivative of the function \( f(\cdot) \) whereas the Ito formula (2.21) also involves the second derivative.

We can use the Ito calculus to derive (2.26) directly by using (1.11). Thus we have that
\[ e^{-\lambda B(T)} = e^{\lambda^2 T/2} - \int_0^T \lambda \exp[-\lambda B(s) + \lambda^2(T-s)/2] \, dB(s) \]
\[ = \langle e^{-\lambda B(T)} \rangle - \int_0^T \lambda E[e^{-\lambda B(T)} | \mathcal{F}_s] \, dB(s). \]

If we assume now that the function \( f : \mathbb{R} \to \mathbb{R} \) of (2.26) can be reconstructed from its Laplace transform \( \hat{f}(\lambda) \), \( \lambda \geq 0 \), so that
\[ f(x) = \int_0^\infty \hat{f}(\lambda) e^{-\lambda x} \, d\lambda, \quad x \in \mathbb{R}, \]
then (2.26) is obtained from (2.27) on multiplication by \( \hat{f}(\lambda) \) and integration over \( \lambda \geq 0 \).

3. Applications

We consider for simplicity the one dimensional SDE
\[ \frac{dx(t)}{dt} = b(x(t)) + \sigma(x(t))W(t), \]
where \( W(\cdot) \) is white noise. Let \( f : \mathbb{R} \to \mathbb{R} \) and \( u(x, t) \) be the function
\[ u(x, t) = E[ f(x(T)) | x(t) = x], \quad x \in \mathbb{R}, \ t < T. \]
Then \( u(x,t) \) is a solution of the diffusion equation
\[
\frac{\partial u(x,t)}{\partial t} + b(x) \frac{\partial u(x,t)}{\partial x} + \frac{\sigma(x)^2}{2} \frac{\partial^2 u(x,t)}{\partial x^2} = 0, \quad x \in \mathbb{R}, \ t < T,
\]
with terminal condition
\[
\lim_{t \to T} u(x,t) = f(x), \quad x \in \mathbb{R}.
\]
Suppose now we wish to estimate
\[
(3.5) \quad u(x,0) = E[ f(x(T)) \mid x(0) = x ].
\]
We can do this by numerically solving the PDE (3.3) with terminal condition (3.4) backwards in time to \( t = 0 \). This methodology is efficient for low dimensional PDE, but in higher dimension (say larger than 3) it becomes numerically very expensive. An alternative, which is much less sensitive to dimension, is to use Monte-Carlo (MC) methods. Thus we solve the SDE (3.1) numerically by several MC simulations and take the average. We can solve (3.1) by using a stochastic Euler method. Let \( \xi_j, \ j = 0,1,2,\ldots, \) be i.i.d. standard normal variables drawn from a random number generator and let \( M \Delta t = T \). We define \( X_j, \ j = 0,\ldots, M, \) by \( X_0 = x \) and
\[
(3.6) \quad X_j = X_{j-1} + b(X_{j-1}) \Delta t + \sigma(X_{j-1}) \sqrt{\Delta t} \xi_{j-1}, \quad j = 1,\ldots, M.
\]
Then in the MC method we set
\[
(3.7) \quad E[ f(x(T)) \mid x(0) = x ] \simeq \text{average of } N \text{ simulations of } f( X_M ),
\]
with error of order \( 1/\sqrt{N} \). The MC method to estimate \( u(x,0) \) to a given accuracy is efficient and reliable even if the coefficients \( b(\cdot), \sigma(\cdot) \) are not particularly well behaved.

We can extend this method in an obvious way to estimate derivatives, say \( \frac{\partial u(x,0)}{\partial x} \), by setting
\[
(3.8) \quad \frac{\partial u(x,0)}{\partial x} \simeq \text{average of } N \text{ simulations of } \frac{f( X^\varepsilon_M ) - f( X_M )}{\varepsilon},
\]
where \( X^\varepsilon_j, \ j = 0,\ldots, M, \) is the solution to (3.6) with \( X^\varepsilon_0 = x + \varepsilon \) and the same \( \xi_0,\ldots, \xi_{M-1} \). The problem with (3.8) is that we are trying to average the ratio of two small quantities, which can be numerically very sensitive. We shall show how the Malliavin calculus gives us a representation for \( \frac{\partial u(x,0)}{\partial x} \) which avoids this issue, and is therefore much more accurate and numerically stable.

Our first task is to represent \( \frac{\partial u(x,0)}{\partial x} \) as an expectation value. Let \( x_\varepsilon(t) \) be the solution to (3.1) with initial condition \( x_\varepsilon(0) = x + \varepsilon \). Then we have that
\[
(3.9) \quad \frac{d}{dt} [ x_\varepsilon(t) - x_0(t) ] = \left[ \int_0^1 b'(\lambda x_\varepsilon(t) + (1 - \lambda)x_0(t)) \, d\lambda \right] [ x_\varepsilon(t) - x_0(t) ]
\]
\[
+ \left[ \int_0^1 \sigma'(\lambda x_\varepsilon(t) + (1 - \lambda)x_0(t)) \, d\lambda \right] [ x_\varepsilon(t) - x_0(t) ] W(t),
\]
with initial condition \( x_\varepsilon(0) - x_0(0) = \varepsilon \). Note that in (3.9) we are generating \( x_\varepsilon(t) \) and \( x_0(t) \) for \( 0 \leq t \leq T \) with the same realization of the white noise \( W(t) \), \( 0 \leq t \leq T \). Letting \( \varepsilon \to 0 \) we see that \( x_\varepsilon(t) - x_0(t) = \varepsilon Y(t) \) to leading order in \( \varepsilon \), where \( Y(t) \) satisfies the first variation equation
\[
(3.10) \quad \frac{dY(t)}{dt} = Y(t) [ b'(x_0(t)) + \sigma'(x_0(t)) W(t) ],
\]
with initial condition $Y(0) = 1$. Note that in (3.10) we are using the same realization of $W(\cdot)$ to obtain for $0 \leq t \leq T$ both $x_0(t)$ from (3.1), and then $Y(t)$ from (3.10). Now we have that

$$
\frac{\partial u(x,0)}{\partial x} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E[ f(x_\varepsilon(T)) - f(x_0(T))] = E[ f'(x_0(T))Y(T) ] .
$$

If $f(\cdot)$ is a smooth function one can reasonably expect to be able to accurately estimate $\partial u(x,0)/\partial x$ by doing MC directly on the expression (3.11) similarly to the way we estimated $u(x,0)$. In the case when $f(\cdot)$ is not $C^1$ as for the payoff on a call option for example when $f(x) = \max[\exp x - K, 0]$, this approach is not suitable.

The Malliavin calculus enables us to obtain a representation for $\partial u(x,0)/\partial x$ as an expectation value which involves just the function $f(\cdot)$ and not its derivative. MC simulations corresponding to this representation are therefore much more stable relative to the $C^1$ norm of $f(\cdot)$ than simulations based on (3.11) directly. To see this we obtain a formula for the Malliavin derivative $Dx_0(T)$ for the solution $x_0(t)$ to (3.1) with $x_0(0) = x$. Now from (3.1) we can regard $x_0(t), \ 0 \leq t \leq T,$ as a function $x_0(t, \phi)$ on $L^2([0,T])$ which satisfies the integral equation

$$
x_0(t, \phi) = x + \int_0^t b(x_0(s, \phi)) \ ds + \int_0^t \sigma(x_0(s, \phi))b(s) \ ds, \ \ \phi \in L^2([0,T]).
$$

For $h \in L^2([0,T])$ and $\varepsilon > 0$ we have therefore that

$$
nx_0(t, \phi+\varepsilon h) - x_0(t, \phi) = \int_0^t \left[ \int_0^1 b'(\lambda x_0(s, \phi + \varepsilon h) + (1 - \lambda)x_0(s, \phi)) \ d\lambda \right] [x_0(s, \phi+\varepsilon h) - x_0(s, \phi)] \ ds
$$

$$
+ \int_0^t \left[ \int_0^1 \sigma'(\lambda x_0(s, \phi + \varepsilon h) + (1 - \lambda)x_0(s, \phi)) \ d\lambda \right] [x_0(s, \phi+\varepsilon h) - x_0(s, \phi)]b(s) \ ds
$$

$$
+ \varepsilon \int_0^t \sigma(x_0(s, \phi + \varepsilon h))h(s) \ ds.
$$

If we divide (3.13) by $\varepsilon$ and let $\varepsilon \to 0$ then we obtain an integral equation for the directional derivative $D_h x_0(t, \phi),$

$$
D_h x_0(t, \phi) = \int_0^t b'(x_0(s, \phi))D_hx_0(s, \phi) \ ds
$$

$$
+ \int_0^t \sigma'(x_0(s, \phi))D_hx_0(s, \phi)b(s) \ ds + \int_0^t \sigma(x_0(s, \phi))h(s) \ ds .
$$

We can write (3.14) in differential form as

$$
\frac{d}{dt} D_h x_0(t, \cdot) = D_h x_0(t, \cdot) [b'(x_0(t, \cdot)) + \sigma'(x_0(t, \cdot))W(t)] + \sigma(x_0(t, \cdot))h(t), \ \ \ 0 \leq t \leq T,
$$

with initial condition $D_h x_0(0, \cdot) = 0$. Comparing (3.10) to (3.15), we see that they are both the same linear equation except that (3.10) is homogeneous and (3.15) is inhomogeneous. Hence by Duhamel’s principle we can write the solution to the inhomogeneous equation (3.15) in terms of the fundamental solution of the homogeneous equation (3.10). Let $Y(t, s), \ s \leq t \leq T,$ be the solution to (3.10) with
initial condition \( Y(s, s) = 1 \). Then by Duhamel we have the formula

\[
D_h x_0(t, \cdot) = \int_0^t Y(t, s) \sigma(x_0(s, \cdot)) h(s) \, ds ,
\]

from which we conclude that

\[
D_s x_0(t, \cdot) = Y(t, s) \sigma(x_0(s, \cdot)), \quad 0 \leq s \leq t .
\]

Observe next that

\[
Y(t, s) = Y(t) Y(s)^{-1}, \quad 0 \leq s \leq t \leq T.
\]

It follows then that the formula (3.11) can be written as

\[
\frac{\partial u(x, 0)}{\partial x} = E \left[ \frac{f'(x_0(T, \cdot)) D_s x_0(T, \cdot) Y(s)}{\sigma(x_0(s, \cdot))} \right] = E \left[ \frac{D_s \{ f(x_0(T, \cdot)) \} Y(s)}{\sigma(x_0(s, \cdot))} \right].
\]

Let \( \alpha : [0, T] \to \mathbb{R} \) be an arbitrary function such that

\[
\int_0^T \alpha(s) \, ds = 1,
\]

and consider the vector field \( G : L^2([0, T]) \to L^2([0, T]) \) with \( G_s(\cdot) = \alpha(s) Y(s, \cdot)/\sigma(x_0(s, \cdot)) \), \( 0 \leq s \leq T \). Then it follows from (3.19) that

\[
\frac{\partial u(x, 0)}{\partial x} = \langle [Df(x_0(T, \cdot))], G(\cdot) \rangle .
\]

It is clear from (3.1), (3.10) that \( G(\cdot) \) is predictable, and so from (2.15) we have that

\[
D^* G(\cdot) = \int_0^T \frac{Y(s, \cdot)}{\sigma(x_0(s, \cdot))} \alpha(s) dB(s) .
\]

We conclude then from (3.21), (3.22) and the integration by parts formula (2.5) that

\[
\frac{\partial u(x, 0)}{\partial x} = \langle f(x_0(T, \cdot)) \cdot D^* G(\cdot) \rangle = \left\{ f(x_0(T, \cdot)) \int_0^T \frac{Y(s, \cdot)}{\sigma(x_0(s, \cdot))} \alpha(s) dB(s) \right\} .
\]

The formula (3.23) is known as the Bismut-Elworthy-Li formula. It is easy by using the stochastic Euler method described above to do a MC simulation for the RHS of (3.23). The accuracy and reliability of this method is superior to MC simulation of the RHS of (3.11).

4. The Poincaré and Logarithmic Sobolev Inequalities

References
