Malliavin Calculus

by David Nualart

1 One-dimensional Gaussian analysis

Consider the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \gamma)$, where

- $\gamma = N(0,1)$ is the standard Gaussian probability on $\mathbb{R}$ with density
$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad x \in \mathbb{R}.$$
- The probability of any interval $[a,b]$ is given by
$$\gamma([a,b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

We are going to introduce two basic differential operators. For any $f \in C^1(\mathbb{R})$ we define:

- Derivative operator: $\quad Df(x) = f'(x)$.
- Divergence operator: $\quad \delta f(x) = xf(x) - f'(x)$.

Denote by $C_p^k(\mathbb{R}^m)$ the space of functions $f : \mathbb{R}^m \to \mathbb{R}$, which are k times continuously differentiable, and such that for some $N \geq 1$, $|f^{(k)}(x)| \leq C(1+|x|^N)$.

Lemma 1.1. *The operators D and δ are adjoint with respect to the measure γ. That means, for any $f, g \in C_p^1(\mathbb{R})$, we have*

$$\boxed{\langle Df, g\rangle_{L^2(\mathbb{R},\gamma)} = \langle f, \delta g\rangle_{L^2(\mathbb{R},\gamma)}}.$$

Proof. Integrating by parts and using $p'(x) = -xp(x)$ we get

$$\begin{aligned}
\int_{\mathbb{R}} f'(x)g(x)p(x)dx &= -\int_{\mathbb{R}} f(x)(g(x)p(x))'dx \\
&= -\int_{\mathbb{R}} f(x)g'(x)p(x)dx + \int_{\mathbb{R}} f(x)g(x)xp(x)dx \\
&= \int_{\mathbb{R}} f(x)\delta g(x)p(x)dx.
\end{aligned}$$

□

Lemma 1.2 (Heisenberg's commutation relation)**.** *Let $f \in C^2(\mathbb{R})$. Then*

$$\boxed{(D\delta - \delta D)f = f}$$

Proof. We can write

$$D\delta f(x) = D(xf(x) - f'(x)) = f(x) + xf'(x) - f''(x)$$

and, on the other hand,
$$\delta Df(x) = \delta f'(x) = xf'(x) - f''(x).$$

This completes the proof. □

More generally, if $f \in C^n(\mathbb{R})$ for $n \geq 2$, we have

$$\boxed{(D\delta^n - \delta^n D)f = n\delta^{n-1}f}$$

Proof. Using induction on n, we can write

$$\begin{aligned} D\delta^n f &= D\delta(\delta^{n-1}f) = \delta D(\delta^{n-1}f) + \delta^{n-1}f \\ &= \delta\left[\delta^{n-1}Df + (n-1)\delta^{n-2}f\right] + \delta^{n-1}f = \delta^n Df + n\delta^{n-1}f. \end{aligned}$$

□

Next we will introduce the Hermite polynomials. Define $H_0(x) = 1$, and for $n \geq 1$ put $H_n(x) = \delta^n 1$. In particular, for $n = 1, 2, 3$, we have

$$\begin{aligned} H_1(x) &= \delta 1 = x \\ H_2(x) &= \delta x = x^2 - 1 \\ H_3(x) &= \delta(x^2 - 1) = x^3 - 3x. \end{aligned}$$

We have the following formula for the derivatives of the Hermite polynomials:

$$H_n' = nH_{n-1}$$

In fact,

$$H_n' = D\delta^n 1 = \delta^n D1 + n\delta^{n-1}1 = nH_{n-1}.$$

Proposition 1.1. *The sequence of normalized Hermite polynomials* $\{\frac{1}{\sqrt{n!}}H_n, n \geq 0\}$ *form a complete orthonormal system of functions in the Hilbert space* $L^2(\mathbb{R}, \gamma)$.

Proof. For $n, m \geq 0$, we can write

$$\int_{\mathbb{R}} H_n(x)H_m(x)p(x)dx = \begin{cases} n! & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

Indeed, using the properties of Hermite polynomials, we obtain

$$\begin{aligned} \int_{\mathbb{R}} H_n(x)H_m(x)p(x)dx &= \int_{\mathbb{R}} H_n(x)\delta^m 1(x)p(x)dx \\ &= \int_{\mathbb{R}} H_n'(x)\delta^{m-1}1(x)p(x)dx \\ &= n\int_{\mathbb{R}} H_{n-1}(x)H_{m-1}(x)p(x)dx. \end{aligned}$$

To show completeness, it suffices to prove that if $f \in L^2(\mathbb{R}, \gamma)$ is orthogonal to all Hermite polynomials, then $f = 0$. Because the leading coefficient of $H_n(x)$ is 1, we have that f is orthogonal to all monomials x^n. As a consequence, for all $t \in \mathbb{R}$,

$$\int_{\mathbb{R}} f(x)e^{itx}p(x)dx = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \int_{\mathbb{R}} f(x)x^n p(x)dx = 0.$$

We can commute the integral and the series because

$$\begin{aligned}\sum_{n=0}^{\infty}\int_{\mathbb{R}}\frac{|tx|^n}{n!}|f(x)|p(x)dx &= \int_{\mathbb{R}} e^{|tx|}|f(x)|p(x)dx\\ &\leq \left[\int_{\mathbb{R}} f^2(x)p(x)dx \int_{\mathbb{R}} e^{2|tx|}p(x)dx\right]^{\frac{1}{2}} < \infty.\end{aligned}$$

Therefore, the Fourier transform of fp is zero, so $fp = 0$, which implies $f = 0$. This completes the proof. □

For each $a \in \mathbb{R}$, we have the following series expansion, which will play an important role.

$$\sum_{n=0}^{\infty}\frac{a^n}{n!}H_n(x) = e^{ax-\frac{a^2}{2}}. \tag{1}$$

Proof of (1): In fact, taking into account that $H_n = \delta^n 1$ and that δ^n is the adjoint of D^n, we obtain

$$\begin{aligned}e^{ax} &= \sum_{n=0}^{\infty}\frac{1}{n!}\langle e^{a\cdot}, H_n\rangle_{L^2(\mathbb{R},\gamma)}H_n(x)\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\langle e^{a\cdot}, \delta^n 1\rangle_{L^2(\mathbb{R},\gamma)}H_n(x)\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\langle D^n(e^{a\cdot}), 1\rangle_{L^2(\mathbb{R},\gamma)}H_n(x)\\ &= \sum_{n=0}^{\infty}\frac{a^n}{n!}\langle e^{a\cdot}, 1\rangle_{L^2(\mathbb{R},\gamma)}H_n(x).\end{aligned}$$

Finally,

$$\langle e^{a\cdot}, 1\rangle_{L^2(\mathbb{R},\gamma)} = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{ax-\frac{x^2}{2}}dx = e^{\frac{a^2}{2}}.$$

and (1) holds true.

Let us now define the *Ornstein Uhlenbeck operator*, which is a second order differential operator. For $f \in C^2(\mathbb{R})$ we set

$$\boxed{Lf(x) = -xf'(x) + f''(x).}$$

This operator has the following properties.

1. $Lf = -\delta Df$.

Proof:

$$\delta Df(x) = \delta f'(x) = xf'(x) - f''(x).$$

2. $LH_n = -nH_n$, that is, H_n is an eigenvector of L with eigenvalue $-n$.

Proof:

$$LH_n = -\delta DH_n = -\delta H_n' = -n\delta H_{n-1} = -nH_n.$$

The operator L is the infinitesimal generator of the Ornstein-Uhlenbeck semigroup. Consider the semigroup of operators $\{P_t, t \geq 0\}$ on $L^2(\mathbb{R}, \gamma)$, defined by $P_t H_n = e^{-nt} H_n$, that is,

$$\boxed{P_t f = \sum_{n=0}^{\infty} \frac{1}{n!} \langle f, H_n \rangle_{L^2(\mathbb{R},\gamma)} e^{-nt} H_n.}$$

Then, L is the generator of P_t, that is, $\frac{dP_t}{dt} = LP_t$.

Proposition 1.2 (Mehler's formula)**.** *For any function $f \in L^2(\mathbb{R}, \gamma)$, we have the following formula for the Ornstein-Uhlenbeck semigroup:*

$$P_t f(x) = \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1-e^{-2t}}y)p(y)dy = E[f(e^{-t}x + \sqrt{1-e^{-2t}}Y)],$$

where Y is a $N(0,1)$ random variable.

Proof. Set $\widetilde{P}_t f(x) = \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1-e^{-2t}}y)p(y)dy$.

(i) We will first show that P_t and $\widetilde{P}_t$ are contraction operators on $L^2(\mathbb{R}, \gamma)$. Indeed,

$$\|P_t f\|^2_{L^2(\mathbb{R},\gamma)} = \sum_{n=0}^{\infty} \frac{1}{n!} \langle f, H_n \rangle^2_{L^2(\mathbb{R},\gamma)} e^{-2nt} \leq \|f\|^2_{L^2(\mathbb{R},\gamma)},$$

and

$$\begin{aligned} \|\widetilde{P}_t f\|^2_{L^2(\mathbb{R},\gamma)} &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(e^{-t}x + \sqrt{1-e^{-2t}}y)p(y)dy \right)^2 p(x)dx \\ &\leq \int_{\mathbb{R}^2} f^2(e^{-t}x + \sqrt{1-e^{-2t}}y)p(y)p(x)dydx \\ &= E[f^2(e^{-t}X + \sqrt{1-e^{-2t}}Y)] = \|f\|^2_{L^2(\mathbb{R},\gamma)}, \end{aligned}$$

where X and Y are independent $N(0,1)$-random variables.

(ii) The functions $\{e^{ax}, a \in \mathbb{R}\}$ form a total system in $L^2(\mathbb{R}, \gamma)$. So, it suffices to show that $P_t e^{a\cdot} = \widetilde{P}_t e^{a\cdot}$ for each $a \in \mathbb{R}$. We have

$$\begin{aligned} (\widetilde{P}_t e^{a\cdot})(x) &= E\left[e^{axe^{-t} + aY\sqrt{1-e^{-2t}}}\right] = e^{axe^{-t}} e^{\frac{a^2}{2}(1-e^{-2t})} \\ &= e^{\frac{a^2}{2}} e^{axe^{-t} - \frac{1}{2}a^2 e^{-2t}} = e^{\frac{a^2}{2}} \sum_{n=0}^{\infty} \frac{a^n e^{-nt}}{n!} H_n(x) \\ &= e^{\frac{a^2}{2}} P_t \left(\sum_{n=0}^{\infty} \frac{a^n}{n!} H_n \right)(x) = e^{\frac{a^2}{2}} P_t \left(e^{a\cdot - \frac{a^2}{2}} \right)(x) \\ &= (P_t e^{a\cdot})(x). \end{aligned}$$

This completes the proof of the proposition. □

The Ornstein-Uhlenbeck semigroup has the following propreties:

1. $\|P_t f\|_{L^p(\mathbb{R},\gamma)} \leq \|f\|_{L^p(\mathbb{R},\gamma)}$ for any $p \geq 2$.

Proof: Using Mehler's formula and Hölder's inequality, we can write

$$\begin{aligned}
\|P_t f\|^p_{L^p(\mathbb{R},\gamma)} &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1-e^{-2t}}y)p(y)dy \right|^p p(x)dx \\
&\leq \int_{\mathbb{R}^2} |f(e^{-t}x + \sqrt{1-e^{-2t}}y)|^p p(y)p(x)dydx \\
&= E[|f(e^{-t}X + \sqrt{1-e^{-2t}}Y)|^p] = \|f\|^p_{L^p(\mathbb{R},\gamma)}.
\end{aligned}$$

2. $P_0 f = f$ and $P_\infty f = \lim_{t\to\infty} P_t f = \int_{\mathbb{R}} f(y)p(y)dy$.

3. $DP_t f = e^{-t} P_t D f$.

4. $f \geq 0$ implies $P_t f \geq 0$.

5. For any $f \in L^2(\mathbb{R},\gamma)$ we have

$$f(x) - \int_{\mathbb{R}} f d\gamma = -\int_0^\infty LP_t f(x)dt.$$

Proof:

$$\begin{aligned}
f(x) - \int_{\mathbb{R}} f d\gamma &= \sum_{n=1}^\infty \frac{1}{n!}\langle f, H_n\rangle_{L^2(\mathbb{R},\gamma)} H_n(x) \\
&= \sum_{n=1}^\infty \frac{1}{n!}\langle f, H_n\rangle_{L^2(\mathbb{R},\gamma)} \left(\int_0^\infty n e^{-nt} H_n(x)dt \right) \\
&= \int_0^\infty \left(\sum_{n=1}^\infty \frac{1}{n!}\langle f, H_n\rangle_{L^2(\mathbb{R},\gamma)} (-LP_t H_n)(x) \right) dt \\
&= -\int_0^\infty LP_t f(x)dt.
\end{aligned}$$

Proposition 1.3 (First Poincaré inequality)**.** *For any* $f \in C^1_p(\mathbb{R})$,

$$\boxed{\mathrm{Var}(f) \leq \|f'\|^2_{L^2(\mathbb{R},\gamma)}}$$

Proof. Set $\bar{f} = \int_{\mathbb{R}} f d\gamma$. We can write

$$\begin{aligned}
\mathrm{Var}(f) &= \int_{\mathbb{R}} f(x)(f(x) - \bar{f})p(x)dx \\
&= -\int_0^\infty \int_{\mathbb{R}} f(x) LP_t f(x)p(x)dxdt \\
&= \int_0^\infty \int_{\mathbb{R}} f(x)\delta DP_t f(x)p(x)dxdt \\
&= \int_0^\infty e^{-t} \int_{\mathbb{R}} f'(x)P_t f'(x)p(x)dxdt \\
&\leq \int_0^\infty e^{-t} \|f'\|_{L^2(\mathbb{R},\gamma)} \|P_t f'\|_{L^2(\mathbb{R},\gamma)} dt \\
&\leq \|f'\|^2_{L^2(\mathbb{R},\gamma)}.
\end{aligned}$$

□

This result as the following interpretation. If f' is small, f is concentrated around its mean value $\bar{f} = \int_{\mathbb{R}} f(x)p(x)dx$ because

$$\mathrm{Var}(f) = \int_{\mathbb{R}} (f(x) - \bar{f})^2 p(x)dx.$$

The result can be extended to the Sobolev space

$$\mathbb{D}^{1,2} = \{f : f, f' \in L^2(\mathbb{R}, \gamma)\}$$

defined as the completion of $C_p^1(\mathbb{R})$ by the norm $\|f\|_{1,2}^2 = \|f\|_{L^2(\mathbb{R},\gamma)}^2 + \|f'\|_{L^2(\mathbb{R},\gamma)}^2$.

1.1 Finite-dimensional case

We consider now the finite-dimensional case. That is, the probability space $(\Omega, \mathcal{F}, P)$ is such that $\Omega = \mathbb{R}^n$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^n)$ is the Borel σ-field of $\mathbb{R}^n$, and P is the standard Gaussian probability with density $p(x) = (2\pi)^{-n/2}e^{-|x|^2/2}$. In this framework we consider, as before, two differential operators. The first is the *derivative operator*, which is simply the gradient of a differentiable function $F\colon \mathbb{R}^n \to \mathbb{R}$:

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n}\right).$$

The second differential operator is the *divergence operator* and is defined on differentiable vector-valued functions $u\colon \mathbb{R}^n \to \mathbb{R}^n$ as follows:

$$\delta(u) = \sum_{i=1}^n \left(u_i x_i - \frac{\partial u_i}{\partial x_i}\right) = \langle u, x\rangle - \operatorname{div} u.$$

It turns out that δ is the adjoint of the derivative operator with respect to the Gaussian measure ¶. This is the content of the next proposition.

Proposition 1.4. *The operator δ is the adjoint of ∇; that is,*

$$\mathbb{E}(\langle u, \nabla F\rangle) = \mathbb{E}(F\delta(u))$$

if $F\colon \mathbb{R}^n \to \mathbb{R}$ and $u\colon \mathbb{R}^n \to \mathbb{R}^n$ are continuously differentiable functions which, together with their partial derivatives, have at most polynomial growth.

Proof. Integrating by parts, and using $\partial p/\partial x_i = -x_i p$, we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} \langle \nabla F, u\rangle p dx &= \sum_{i=1}^n \int_{\mathbb{R}^n} \frac{\partial F}{\partial x_i} u_i p dx \\ &= \sum_{i=1}^n \left(-\int_{\mathbb{R}^n} F \frac{\partial u_i}{\partial x_i} p dx + \int_{\mathbb{R}^n} F u_i x_i p dx \right) \\ &= \int_{\mathbb{R}^n} F\delta(u) p dx.\end{aligned}$$

This completes the proof. □

2 Malliavin calculus on the Wiener space

2.1 Brownian motion and Wiener space

Brownian motion was named after the botanist Robert Brown, who observed in a microscope the complex and erratic motion of grains of pollen suspended in water. Brownian motion was then rigorously defined and studied by Norbert Wiener; this is why it is also called the Wiener process. The mathematical definition of Brownian motion is the following.

Definition 2.1. *A real-valued stochastic process $B = (B_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ is called a Brownian motion if it satisfies the following conditions:*

1. *Almost surely $B_0 = 0$.*
2. *For all $0 \leq t_1 < \cdots < t_n$ the increments $B_{t_n} - B_{t_{n-1}}, \ldots, B_{t_2} - B_{t_1}$ are independent random variables.*
3. *If $0 \leq s < t$, the increment $B_t - B_s$ is a Gaussian random variable with mean zero and variance $t - s$.*
4. *With probability one, the map $t \to B_t$ is continuous.*

Properties (i), (ii), and (iii) are equivalent to saying that B is a Gaussian process with mean zero and covariance function

$$\Gamma(s,t) = \min(s,t). \tag{2}$$

The existence of Brownian motion can be proved in the following way: The function $\Gamma(s,t) = \min(s,t)$ is symmetric and nonnegative definite because it can be written as

$$\min(s,t) = \int_0^\infty \mathbf{1}_{[0,s]}(r)\mathbf{1}_{[0,t]}(r)dr.$$

Then, for any integer $n \geq 1$ and real numbers $a_1, \ldots, a_n$,

$$\begin{aligned}\sum_{i,j=1}^n a_i a_j \min(t_i, t_j) &= \sum_{i,j=1}^n a_i a_j \int_0^\infty \mathbf{1}_{[0,t_i]}(r)\mathbf{1}_{[0,t_j]}(r)dr \\ &= \int_0^\infty \left(\sum_{i=1}^n a_i \mathbf{1}_{[0,t_i]}(r)\right)^2 dr \geq 0.\end{aligned}$$

Therefore, by Kolmogorov's extension theorem, there exists a Gaussian process with mean zero and covariance function $\min(s,t)$. Moreover, for any $s \leq t$, the increment $B_t - B_s$ has the normal distribution $N(0, t-s)$. This implies that for any natural number k we have

$$\mathbb{E}\left((B_t - B_s)^{2k}\right) = \frac{(2k)!}{2^k k!}(t-s)^k.$$

Therefore, by Kolmogorov's continuity theorem, there exists a version of B with Hölder-continuous trajectories of order γ for any $\gamma < (k-1)/(2k)$ on any interval $[0,T]$. This implies that the paths of this version of the process B are γ-Hölder continuous on $[0,T]$ for any $\gamma < 1/2$ and $T > 0$.

Brownian motion can be defined in the canonical probability space $(\Omega, \mathcal{F}, P)$ known as the Wiener space, where

- Ω is the space of continuous functions $\omega\colon \mathbb{R}_+ \to \mathbb{R}$ vanishing at the origin.

- $\mathcal{F}$ is the Borel σ-field $\mathcal{B}(\Omega)$ for the topology corresponding to uniform convergence on compact sets. One can easily show that $\mathcal{F}$ coincides with the σ-field generated by the collection of cylinder sets

$$C = \{\omega \in \Omega : \omega(t_1) \in A_1, \dots, \omega(t_k) \in A_k\}, \tag{3}$$

  for any integer $k \geq 1$, Borel sets $A_1, \dots, A_k$ in $\mathbb{R}$, and $0 \leq t_1 < \cdots < t_k$.

- P is the Wiener measure. That is, P is defined on a cylinder set of the form (3) by

$$P(C) = \int_{A_1 \times \cdots \times A_k} p_{t_1}(x_1) p_{t_2 - t_1}(x_2 - x_1) \cdots p_{t_k - t_{k-1}}(x_k - x_{k-1})\, dx_1 \cdots dx_k, \tag{4}$$

  where $p_t(x)$ denotes the Gaussian density $p_t(x) = (2\pi t)^{-1/2} e^{-x^2/(2t)}, \quad x \in \mathbb{R}, t > 0$.

The mapping P defined by (4) on cylinder sets can be uniquely extended to a probability measure on $\mathcal{F}$. This fact can be proved as a consequence of the existence of Brownian motion on $\mathbb{R}_+$. Finally, the canonical stochastic process defined as $B_t(\omega) = \omega(t)$, $\omega \in \Omega$, $t \geq 0$, is a Brownian motion.

2.2 Wiener integral

We next define the integral of square integrable functions with respect to Brownian motion, known as the Wiener integral. We consider the set $\mathcal{E}_0$ of step functions

$$\varphi_t = \sum_{j=0}^{n-1} a_j \mathbf{1}_{(t_j, t_{j+1}]}(t), \quad t \geq 0, \tag{5}$$

where $n \geq 1$ is an integer, $a_0, \dots, a_{n-1} \in \mathbb{R}$, and $0 = t_0 < \cdots < t_n$. The Wiener integral of a step function $\varphi \in \mathcal{E}_0$ of the form (5) is defined by

$$\int_0^\infty \varphi_t dB_t = \sum_{j=0}^{n-1} a_j (B_{t_{j+1}} - B_{t_j}).$$

The mapping $\varphi \to \int_0^\infty \varphi_t dB_t$ from $\mathcal{E}_0 \subset L^2(\mathbb{R}_+)$ to $L^2(\Omega)$ is linear and isometric:

$$\mathbb{E}\left(\left(\int_0^\infty \varphi_t dB_t\right)^2\right) = \sum_{j=0}^{n-1} a_j^2 (t_{j+1} - t_j) = \int_0^\infty \varphi_t^2 dt = \|\varphi\|^2_{L^2(\mathbb{R}_+)}.$$

The space $\mathcal{E}_0$ is a dense subspace of $L^2(\mathbb{R}_+)$. Therefore, the mapping

$$\varphi \to \int_0^\infty \varphi_t dB_t$$

can be extended to a linear isometry between $L^2(\mathbb{R}_+)$ and the Gaussian subspace of $L^2(\Omega)$ spanned by the Brownian motion. The random variable $\int_0^\infty \varphi_t dB_t$ is called the Wiener integral of $\varphi \in L^2(\mathbb{R}_+)$ and is denoted by $B(\varphi)$. Observe that it is a Gaussian random variable with mean zero and variance $\|\varphi\|^2_{L^2(\mathbb{R}_+)}$.

2.3 Malliavin derivative

Let $B = (B_t)_{t\geq 0}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}$ is the σ-field generated by B. Set $H = L^2(\mathbb{R}_+)$, and for any $h \in H$, consider the Wiener integral

$$B(h) = \int_0^\infty h(t)dB_t.$$

The Hilbert space H plays a basic role in the definition of the derivative operator. In fact, the derivative of a random variable $F\colon \Omega \to \mathbb{R}$ takes values in H, and $(D_tF)_{t\geq 0}$ is a stochastic process in $L^2(\Omega; H)$.

We start by defining the derivative in a dense subset of $L^2(\Omega)$. More precisely, consider the set $\mathcal{S}$ of smooth and cylindrical random variables of the form

$$F = f(B(h_1), \ldots, B(h_n)), \tag{6}$$

where $f \in C_p^\infty(\mathbb{R}^n)$ and $h_i \in H$.

Definition 2.2. *If $F \in \mathcal{S}$ is a smooth and cylindrical random variable of the form* (6), *the derivative DF is the H-valued random variable defined by*

$$D_tF = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(B(h_1), \ldots, B(h_n))h_i(t).$$

For instance, $D(B(h)) = h$ and $D(B_{t_1}) = \mathbf{1}_{[0,t_1]}$, for any $t_1 \geq 0$.

The derivative operator can be interpreted as a directional derivative. Consider the Cameron-Martin space $H^1 \subset \Omega$, which is is the set of functions of the form $\psi(t) = \int_0^t h(s)ds$, where $h \in H$. Then, For ant $h \in H$, $\langle DF, h\rangle_H$ is the derivative of F in the direction of $\int_0^\cdot h(s)ds$:

$$\langle DF, h\rangle_H = \int_0^T h_t D_t F dt = \frac{d}{d\epsilon} F\left(\omega + \epsilon \int_0^\cdot h_s ds\right)|_{\epsilon=0}.$$

For example, if $F = B_{t_1}$, then

$$F\left(\omega + \epsilon \int_0^\cdot h_s ds\right) = \omega(t_1) + \epsilon \int_0^{t_1} h_s ds,$$

so, $\langle DF, h\rangle_H = \int_0^{t_1} h_s ds$, and $D_tF = \mathbf{1}_{[0,t_1]}(t)$.

The operator D defines a linear and unbounded operator from $\mathcal{S} \subset L^2(\Omega)$ into $L^2(\Omega; H)$. Let us now introduce the divergence operator. Denote by $\mathcal{S}_H$ the class of smooth and cylindrical stochastic processes $u = (u_t)_{t\geq 0}$ of the form

$$u_t = \sum_{j=1}^n F_j h_j(t), \tag{7}$$

where $F_j \in \mathcal{S}$ and $h_j \in H$.

Definition 2.3. *We define the divergence of an element u of the form (7) as the random variable given by*

$$\delta(u) = \sum_{j=1}^n F_j B(h_j) - \sum_{j=1}^n \langle DF_j, h_j\rangle_H.$$

In particular, for any $h \in H$ we have $\delta(h) = B(h)$.

As in the finite-dimensional case, the divergence is the adjoint of the derivative operator, as is shown in the next proposition.

Proposition 2.1. *Let $F \in \mathcal{S}$ and $u \in \mathcal{S}_H$. Then*

$$\mathbb{E}(F\delta(u)) = \mathbb{E}(\langle DF, u\rangle_H).$$

Proof. We can assume that $F = f(B(h_1)\ldots, B(h_n))$ and

$$u = \sum_{j=1}^{n} g_j(B(h_1)\ldots, B(h_n))h_j,$$

where $h_1, \ldots, h_n$ are orthonormal elements in H. In this case, the duality relationship reduces to the finite-dimensional case proved in Proposition 1.4. □

We will make use of the notation $D_hF = \langle DF, h\rangle_H$ for any $h \in H$ and $F \in \mathcal{S}$. The following proposition states the basic properties of the derivative and divergence operators on smooth and cylindrical random variables.

Proposition 2.2. *Suppose that $u, v \in \mathcal{S}_H$, $F \in \mathcal{S}$, and $h \in H$. Then, if $(e_i)_{i\geq 1}$ is a complete orthonormal system in H, we have*

$$\mathbb{E}(\delta(u)\delta(v)) = \mathbb{E}(\langle u, v\rangle_H) + \mathbb{E}\left(\sum_{i,j=1}^{\infty} D_{e_i}\langle u, e_j\rangle_H D_{e_j}\langle v, e_i\rangle_H\right), \quad (8)$$

$$D_h(\delta(u)) = \delta(D_hu) + \langle h, u\rangle_H, \quad (9)$$

$$\delta(Fu) = F\delta(u) - \langle DF, u\rangle_H. \quad (10)$$

Property (8) can also be written as

$$\mathbb{E}(\delta(u)\delta(v)) = \mathbb{E}\left(\int_0^\infty u_t v_t dt\right) + \mathbb{E}\left(\int_0^\infty\int_0^\infty D_s u_t D_t v_s ds dt\right).$$

Proof of Proposition 2.2. We first show property (9). Consider $u = \sum_{j=1}^n F_j h_j$, where $F_j \in \mathcal{S}$ and $h_j \in H$ for $j = 1, \ldots, n$. Then, using $D_h(B(h_j)) = \langle h, h_j\rangle_H$, we obtain

$$\begin{aligned} D_h(\delta(u)) &= D_h\left(\sum_{j=1}^n F_j B(h_j) - \sum_{j=1}^n \langle DF_j, h_j\rangle_H\right) \\ &= \sum_{j=1}^n F_j\langle h, h_j\rangle_H + \sum_{j=1}^n (D_hF_jB(h_j) - \langle D_h(DF_j), h_j\rangle_H) \\ &= \langle u, h\rangle_H + \delta(D_hu). \end{aligned}$$

To show property (8), using the duality formula (Proposition 2.1) and property (9), we get

$$\begin{aligned} \mathbb{E}(\delta(u)\delta(v)) &= \mathbb{E}(\langle v, D(\delta(u))\rangle_H) \\ &= \mathbb{E}\left(\sum_{i=1}^\infty \langle v, e_i\rangle_H\, D_{e_i}(\delta(u))\right) \\ &= \mathbb{E}\left(\sum_{i=1}^\infty \langle v, e_i\rangle_H\, \left(\langle u, e_i\rangle_H + \delta(D_{e_i}u)\right)\right) \\ &= \mathbb{E}(\langle u, v\rangle_H) + \mathbb{E}\left(\sum_{i,j=1}^\infty D_{e_i}\langle u, e_j\rangle_H\, D_{e_j}\langle v, e_i\rangle_H\right). \end{aligned}$$

Finally, to prove property (10) we choose a smooth random variable $G \in \mathcal{S}$ and write, using the duality relationship (Proposition 2.1),

$$\begin{aligned}\mathbb{E}(\delta(Fu)G) &= \mathbb{E}(\langle DG, Fu\rangle_H) = \mathbb{E}(\langle u, D(FG) - GDF\rangle_H) \\ &= \mathbb{E}((\delta(u)F - \langle u, DF\rangle_H)G),\end{aligned}$$

which implies the result because $\mathcal{S}$ is dense in $L^2(\Omega)$. □

2.4 Sobolev spaces

The next proposition will play a basic role in extending the derivative to suitable Sobolev spaces of random variables.

Proposition 2.3. *The operator D is closable from $L^p(\Omega)$ to $L^p(\Omega; H)$ for any $p \geq 1$.*

Proof. Assume that the sequence $F_N \in \mathcal{S}$ satisfies

$$F_N \overset{L^p(\Omega)}{\longrightarrow} 0 \quad \text{and} \quad DF_N \overset{L^p(\Omega;H)}{\longrightarrow} \eta,$$

as $N \to \infty$. Then $\eta = 0$. Indeed, for any $u = \sum_{j=1}^N G_j h_j \in \mathcal{S}_H$ such that $G_j B(h_j)$ and DG_j are bounded, by the duality formula (Proposition 2.1), we obtain

$$\begin{aligned}\mathbb{E}(\langle \eta, u\rangle_H) &= \lim_{N\to\infty} \mathbb{E}(\langle DF_N, u\rangle_H) \\ &= \lim_{N\to\infty} \mathbb{E}(F_N \delta(u)) = 0.\end{aligned}$$

This implies that $\eta = 0$, since the set of $u \in \mathcal{S}_H$ with the above properties is dense in $L^p(\Omega; H)$ for all $p \geq 1$. □

We consider the closed extension of the derivative, which we also denote by D. The domain of this operator is defined by the following Sobolev spaces. For any $p \geq 1$, we denote by $\mathbb{D}^{1,p}$ the closure of $\mathcal{S}$ with respect to the seminorm

$$\|F\|_{1,p} = \left(\mathbb{E}(|F|^p) + \mathbb{E}\left(\left|\int_0^\infty (D_t F)^2 dt\right|^{p/2}\right)\right)^{1/p}.$$

In particular, F belongs to $\mathbb{D}^{1,p}$ if and only if there exists a sequence $F_n \in \mathcal{S}$ such that

$$F_n \overset{L^p(\Omega)}{\longrightarrow} F \quad \text{and} \quad DF_n \overset{L^p(\Omega;H)}{\longrightarrow} DF,$$

as $n \to \infty$. For $p = 2$, the space $\mathbb{D}^{1,2}$ is a Hilbert space with scalar product

$$\langle F, G\rangle_{1,2} = \mathbb{E}(FG) + \mathbb{E}\left(\int_0^\infty D_t F D_t G dt\right).$$

In the same way we can introduce spaces $\mathbb{D}^{1,p}(H)$ by taking the closure of $\mathcal{S}_H$. The corresponding seminorm is denoted by $\|\cdot\|_{1,p,H}$.

The Malliavin derivative satisfies the following chain rule.

Proposition 2.4. *Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be a continuous differentiable function such that $|\varphi'(x)| \leq C(1 + |x|^\alpha)$ for some $\alpha \geq 0$. Let $F \in \mathbb{D}^{1,p}$ for some $p \geq \alpha + 1$. Then, $\varphi(F)$ belongs to $\mathbb{D}^{1,q}$, where $q = p/(\alpha + 1)$, and*

$$D(\varphi(F)) = \varphi'(F)DF.$$

Proof. Notice that $|\varphi(x)| \le C'(1+|x|^{\alpha+1})$, for some constant C', which implies that $\varphi(F) \in L^q(\Omega)$ and, by Hölder's inequality, $\varphi'(F)DF \in L^q(\Omega; H)$. Then, to show the proposition it suffices to approximate F by smooth and cylindrical random variables, and φ by $\varphi * \alpha_n$, where α_n is an approximation of the identity. □

We next define the domain of the divergence operator. We identify the Hilbert space $L^2(\Omega; H)$ with $L^2(\Omega \times \mathbb{R}_+)$.

Definition 2.4. *The domain of the divergence operator* $\operatorname{Dom}\delta$ *in* $L^2(\Omega)$ *is the set of processes* $u \in L^2(\Omega \times \mathbb{R}_+)$ *such that there exists* $\delta(u) \in L^2(\Omega)$ *satisfying the duality relationship*

$$\mathbb{E}(\langle DF, u\rangle_H) = \mathbb{E}(\delta(u)F),$$

for any $F \in \mathbb{D}^{1,2}$.

Observe that δ is a linear operator such that $\mathbb{E}(\delta(u)) = 0$. Moreover, δ is closed; that is, if the sequence $u_n \in \mathcal{S}_H$ satisfies

$$u_n \overset{L^2(\Omega;H)}{\longrightarrow} u \quad \text{and} \quad \delta(u_n) \overset{L^2(\Omega)}{\longrightarrow} G,$$

as $n \to \infty$, then u belongs to $\operatorname{Dom}\delta$ and $\delta(u) = G$.

Proposition 2.2 can be extended to random variables in suitable Sobolev spaces. Property (8) holds for $u, v \in \mathbb{D}^{1,2}(H) \subset \operatorname{Dom}\delta$ and, in this case, for any $u \in \mathbb{D}^{1,2}(H)$ we can write

$$\mathbb{E}(\delta(u)^2) \le \mathbb{E}\left(\int_0^\infty (u_t)^2 dt\right) + \mathbb{E}\left(\int_0^\infty \int_0^\infty (D_s u_t)^2 ds dt\right) = \|u\|_{1,2,H}^2.$$

Property (9) holds if $u \in \mathbb{D}^{1,2}(H)$ and $D_h u \in \operatorname{Dom}\delta$. Finally, property (10) holds if $F \in \mathbb{D}^{1,2}$, $Fu \in L^2(\Omega; H)$, $u \in \operatorname{Dom}\delta$, and the right-hand side is square integrable.

We can also introduce iterated derivatives and the corresponding Sobolev spaces. The kth derivative $D^k F$ of a random variable $F \in \mathcal{S}$ is the k-parameter process obtained by iteration:

$$D^k_{t_1,\dots t_k} F = \sum_{i_1,\dots,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} (B(h_1), \dots, B(h_n)) h_{i_1}(t_1) \cdots h_{i_k}(t_k).$$

For any $p \ge 1$, the operator D^k is closable from $L^p(\Omega)$ into $L^p(\Omega; H^{\otimes k})$, and we denote by $\mathbb{D}^{k,p}$ the closure of $\mathcal{S}$ with respect to the seminorm

$$\|F\|_{k,p} = \left(\mathbb{E}(|F|^p) + \mathbb{E}\left(\sum_{j=1}^{k} \left| \int_{\mathbb{R}_+^j} (D^j_{t_1,\dots,t_j} F)^2 dt_1 \cdots dt_j \right|^{p/2}\right)\right)^{1/p}.$$

For any $k \ge 1$, we set $\mathbb{D}^{k,\infty} := \cap_{p\ge 2} \mathbb{D}^{k,p}$, $\mathbb{D}^{\infty,2} := \cap_{k\ge 1} \mathbb{D}^{k,2}$, and $\mathbb{D}^\infty := \cap_{k\ge 1} \mathbb{D}^{k,\infty}$. Similarly, we can introduce the spaces $\mathbb{D}^{k,p}(H)$.

2.5 The divergence as a tochastic integral

The Malliavin derivative is a local operator in the following sense. Let $[a, b] \subset \mathbb{R}_+$ be fixed. We denote by $\mathcal{F}_{[a,b]}$ the σ-field generated by the random variables $\{B_s - B_a, s \in [a, b]\}$.

Lemma 2.5. *Let F be a random variable in $\mathbb{D}^{1,2} \cap L^2(\Omega, \mathcal{F}_{[a,b]}, P)$. Then $D_t F = 0$ for almost all $(t, \omega) \in [a, b]^c \times \Omega$.*

Proof. If F belongs to $\mathcal{S} \cap L^2(\Omega, \mathcal{F}_{[a,b]}, P)$ then this property is clear. The general case follows by approximation. □

The following result says that the divergence operator is an extension of Itô's integral. For any $t \geq 0$ we denote by $\mathcal{F}_t$ the σ-algebra generated by the null sets and the random variables $B_s, s \in [0, t]$.

Theorem 2.6. *Any process u in $L^2(\Omega \times \mathbb{R}_+)$ which is adapted (for each $t \geq 0$, u_t is $\mathcal{F}_t$-measurable) belongs to* Dom δ *and $\delta(u)$ coincides with Itô's stochastic integral*

$$\delta(u) = \int_0^\infty u_t dB_t.$$

Proof. Consider a simple process u of the form

$$u_t = \sum_{j=0}^{n-1} \phi_j \mathbf{1}_{(t_j, t_{j+1}]}(t),$$

where $0 \leq t_0 < t_1 < \cdots < t_n$ and the random variables $\phi_j \in \mathcal{S}$ are $\mathcal{F}_{t_j}$-measurable. Then $\delta(u)$ coincides with the Itô integral of u because, by (10),

$$\delta(u) = \sum_{j=0}^{n-1} \phi_j (B_{t_{j+1}} - B_{t_j}) - \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} D_t \phi_j dt = \sum_{j=0}^{n-1} \phi_j (B_{t_{j+1}} - B_{t_j}),$$

taking into account that $D_t \phi_j = 0$ if $t > t_j$ by Lemma 2.5. Then the result follows by approximating any process in $L^2(\mathcal{P})$ by simple processes, and approximating any $\phi_j \in L^2(\Omega, \mathcal{F}_{t_j}, P)$ by $\mathcal{F}_{t_j}$-measurable smooth and cylindrical random variables. □

If u is not adapted, $\delta(u)$ coincides with an anticipating stochastic integral introduced by Skorohod. Using techniques of Malliavin calculus, Nualart and Pardoux developed a stochastic calculus for the Skorohod integral.

If u and v are adapted then, for $s < t$, $D_t v_s = 0$ and, for $s > t$, $D_s u_t = 0$. As a consequence, property (8) leads to the isometry property of Itô's integral for adapted processes $u, v \in \mathbb{D}^{1,2}(H)$:

$$\mathbb{E}(\delta(u)\delta(v)) = \mathbb{E}\left(\int_0^\infty u_t v_t dt \right).$$

If u is an adapted process in $\mathbb{D}^{1,2}(H)$ then, from property (9), we obtain

$$D_t \left(\int_0^\infty u_s dB_s \right) = u_t + \int_t^\infty D_t u_s dB_s, \tag{11}$$

because $D_t u_s = 0$ if $t > s$.

2.6 Isonormal Gaussian processes

So far, we have developed the Malliavin calculus with respect to Brownian motion. In this case, the Wiener integral $B(h) = \int_0^\infty h(t)dB_t$ gives rise to a centered Gaussian family indexed by the Hilbert space $H = L^2(\mathbb{R}_+)$. More generally, consider a separable Hilbert space H with scalar product $\langle \cdot, \cdot \rangle_H$. An *isonormal Gaussian process* is a centered Gaussian family $\mathfrak{H}_1 = \{W(h), h \in H\}$ satisfying

$$\mathbb{E}(W(h)W(g)) = \langle h, g \rangle_H,$$

for any $h, g \in H$. Observe that $\mathfrak{H}_1$ is a Gaussian subspace of $L^2(\Omega)$.

The Malliavin calculus can be developed in the framework of an isonormal Gaussian process, and all the notions and properties that do not depend on the fact that $H = L^2(\mathbb{R}_+)$ can be extended to this more general context.

3 Multiple stochastic integrals. Wiener chaos

In this section we present the Wiener chaos expansion, which provides an orthogonal decomposition of random variables in $L^2(\Omega)$ in terms of multiple stochastic integrals. We then compute the derivative and the divergence operators on the Wiener chaos expansion.

3.1 Multiple Stochastic Integrals

Recall that $B = (B_t)_{t\geq 0}$ is a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}$ is generated by B. Let $L^2_s(\mathbb{R}^n_+)$ be the space of symmetric square integrable functions $f : \mathbb{R}^n_+ \to \mathbb{R}$. If $f : \mathbb{R}^n_+ \to \mathbb{R}$, we define its symmetrization by

$$\tilde{f}(t_1, \dots, t_n) = \frac{1}{n!} \sum_\sigma f(t_{\sigma(1)}, \dots, t_{\sigma(n)}),$$

where the sum runs over all permutations σ of $\{1, 2, \dots, n\}$. Observe that

$$\|\tilde{f}\|_{L^2(\mathbb{R}^n_+)} \leq \|f\|_{L^2(\mathbb{R}^n_+)}.$$

Definition 3.1. *The multiple stochastic integral of $f \in L^2_s(\mathbb{R}^n_+)$ is defined as the iterated stochastic integral*

$$I_n(f) = n! \int_0^\infty \int_0^{t_n} \cdots \int_0^{t_2} f(t_1, \dots, t_n) dB_{t_1} \cdots dB_{t_n}.$$

Note that if $f \in L^2(\mathbb{R}_+)$, $I_1(f) = B(f)$ is the Wiener integral of f.

If $f \in L^2(\mathbb{R}^n_+)$ is not necessarily symmetric, we define

$$I_n(f) = I_n(\tilde{f}).$$

Using the properties of Itô's stochastic integral, one can easily check the following isometry property: for all $n, m \geq 1$, $f \in L^2(\mathbb{R}^n_+)$, and $g \in L^2(\mathbb{R}^m_+)$,

$$\mathbb{E}(I_n(f)I_m(g)) = \begin{cases} 0 & \text{if } n \neq m, \\ n!\langle \tilde{f}, \tilde{g} \rangle_{L^2(\mathbb{R}^n_+)} & \text{if } n = m. \end{cases} \quad (12)$$

Next, we want to compute the product of two multiple integrals. Let $f \in L^2_s(\mathbb{R}^n_+)$ and $g \in L^2_s(\mathbb{R}^m_+)$. For any $r = 0, \dots, n \wedge m$, we define the *contraction* of f and g of order r to be the element of $L^2(\mathbb{R}^{n+m-2r}_+)$ defined by

$$\begin{aligned}&(f \otimes_r g)\,(t_1, \dots, t_{n-r}, s_1, \dots, s_{m-r})\\&\quad= \int_{\mathbb{R}^r_+} f(t_1, \dots, t_{n-r}, x_1, \dots, x_r) g(s_1, \dots, s_{m-r}, x_1, \dots, x_r) dx_1 \cdots dx_r.\end{aligned}$$

We denote by $f \tilde{\otimes}_r g$ the symmetrization of $f \otimes_r g$. Then, the product of two multiple stochastic integrals satisfies the following formula:

$$I_n(f) I_m(g) = \sum_{r=0}^{n \wedge m} r! \binom{n}{r} \binom{m}{r} I_{n+m-2r}(f \otimes_r g). \tag{13}$$

The next result gives the relation between multiple stochastic integrals and Hermite polynomials.

Proposition 3.1. *For any $g \in L^2(\mathbb{R}_+)$, we have*

$$I_n(g^{\otimes n}) = \|g\|^n_{L^2(\mathbb{R}_+)} H_n\left(\frac{B(g)}{\|g\|_{L^2(\mathbb{R}_+)}} \right),$$

where $g^{\otimes n}(t_1, \dots, t_n) = g(t_1) \cdots g(t_n)$.

Proof. We can assume that $\|g\|_{L^2(\mathbb{R}_+)} = 1$. We proceed by induction over n. The case $n = 1$ is immediate. We then assume that the result holds for $1, \dots, n$. Using the product rule (13), the induction hypothesis, and the recursive relation for the Hermite polynomials, we get

$$\begin{aligned}I_{n+1}(g^{\otimes(n+1)}) &= I_n(g^{\otimes n}) I_1(g) - n I_{n-1}(g^{\otimes(n-1)})\\&= H_n(B(g)) B(g) - n H_{n-1}(B(g))\\&= H_{n+1}(B(g)),\end{aligned}$$

which concludes the proof. □

The next result is the *Wiener chaos expansion.*

Theorem 3.2. *Every $F \in L^2(\Omega)$ can be uniquely expanded into a sum of multiple stochastic integrals as follows:*

$$F = \mathbb{E}(F) + \sum_{n=1}^{\infty} I_n(f_n),$$

where $f_n \in L^2_s(\mathbb{R}^n_+)$.

For any $n \geq 1$, we denote by $\mathcal{H}_n$ the closed subspace of $L^2(\Omega)$ formed by all multiple stochastic integrals of order n. For $n = 0$, $\mathcal{H}_0$ is the space of constants. Observe that $\mathcal{H}_1$ coincides with the isonormal Gaussian process $\{B(f), f \in L^2(\mathbb{R}_+)\}$. Then Theorem 3.2 can be reformulated by saying that we have the orthogonal decomposition

$$L^2(\Omega) = \oplus_{n=0}^{\infty} \mathcal{H}_n.$$

Proof of Theorem 3.2. It suffices to show that if a random variable $G \in L^2(\Omega)$ is orthogonal to $\oplus_{n=0}^{\infty} \mathcal{H}_n$ then $G = 0$. This assumption implies that G is orthogonal to all random variables of the form $B(g)^k$, where $g \in L^2(\mathbb{R}_+)$, $k \geq 0$. This in turn implies that G is orthogonal to all the exponentials $\exp(B(h))$, which form a total set in $L^2(\Omega)$. So $G = 0$. □

3.2 Derivative operator on the Wiener chaos

Let us compute the derivative of a multiple stochastic integral.

Proposition 3.2. *Let $f \in L_s^2(\mathbb{R}_+^n)$. Then $I_n(f) \in \mathbb{D}^{1,2}$ and*

$$D_t I_n(f) = n I_{n-1}(f(\cdot, t)).$$

Proof. Assume that $f = g^{\otimes n}$, with $\|g\|_{L^2(\mathbb{R}_+)} = 1$. Then, using Proposition 3.1 and the properties of Hermite polynomials, we have

$$\begin{aligned} D_t I_n(f) &= D_t(H_n(B(g))) = H_n'(B(g))D_t(B(g)) = nH_{n-1}(B(g))g(t) \\ &= ng(t)I_{n-1}(g^{\otimes(n-1)}) = nI_{n-1}(f(\cdot,t)). \end{aligned}$$

The general case follows using linear combinations and a density argument. This finishes the proof. □

Moreover, applying (12), we have

$$\begin{aligned} \mathbb{E}\left(\int_{\mathbb{R}_+} (D_t I_n(f))^2 dt\right) &= n^2 \int_{\mathbb{R}_+} \mathbb{E}(I_{n-1}(f(\cdot,t))^2) dt \\ &= n^2 (n-1)! \int_{\mathbb{R}_+} \|f(\cdot,t)\|^2_{L^2(\mathbb{R}_+^{n-1})} dt \\ &= n n! \|f\|^2_{L^2(\mathbb{R}_+^n)} \\ &= n\mathbb{E}(I_n(f)^2). \end{aligned} \tag{14}$$

As a consequence of Proposition 3.2 and (14), we deduce the following result.

Proposition 3.3. *Let $F \in L^2(\Omega)$ with Wiener chaos expansion $F = \sum_{n=0}^\infty I_n(f_n)$. Then $F \in \mathbb{D}^{1,2}$ if and only if*

$$\mathbb{E}(\|DF\|_H^2) = \sum_{n=1}^\infty n n! \|f_n\|^2_{L^2(\mathbb{R}_+^n)} < \infty,$$

and in this case

$$D_t F = \sum_{n=1}^\infty n I_{n-1}(f_n(\cdot, t)).$$

Similarly, if $k \geq 2$, one can show that $F \in \mathbb{D}^{k,2}$ if and only if

$$\sum_{n=1}^\infty n^k n! \|f_n\|^2_{L^2(\mathbb{R}_+^n)} < \infty,$$

and in this case

$$D^k_{t_1,\dots,t_k} F = \sum_{n=k}^\infty n(n-1)\cdots(n-k+1) I_{n-k}(f_n(\cdot, t_1, \dots, t_k)),$$

where the series converges in $L^2(\Omega \times \mathbb{R}_+^k)$. As a consequence, if $F \in \mathbb{D}^{\infty,2}$ then the following formula, due to Stroock, allows us to compute explicitly the kernels in the Wiener chaos expansion of F:

$$f_n = \frac{1}{n!}\mathbb{E}(D^n F). \tag{15}$$

Example 3.3. *Consider $F = B_1^3$. Then*

$$\begin{aligned} f_1(t_1) &= \mathbb{E}(D_{t_1} B_1^3) = 3\mathbb{E}(B_1^2)\mathbf{1}_{[0,1]}(t_1) = 3\mathbf{1}_{[0,1]}(t_1), \\ f_2(t_1,t_2) &= \tfrac{1}{2}\mathbb{E}(D^2_{t_1,t_2} B_1^3) = 3\mathbb{E}(B_1)\mathbf{1}_{[0,1]}(t_1 \vee t_2) = 0, \\ f_3(t_1,t_2,t_3) &= \tfrac{1}{6}\mathbb{E}(D^3_{t_1,t_2,t_3} B_1^3) = \mathbf{1}_{[0,1]}(t_1 \vee t_2 \vee t_3), \end{aligned}$$

and we obtain the Wiener chaos expansion

$$B_1^3 = 3B_1 + 6\int_0^1 \int_0^{t_1} \int_0^{t_2} dB_{t_1} dB_{t_2} dB_{t_3}.$$

Proposition 3.3 implies the following characterization of the space $\mathbb{D}^{1,2}$.

Proposition 3.4. *Let $F \in L^2(\Omega)$. Assume that there exists an element $u \in L^2(\Omega; H)$ such that, for all $G \in \mathcal{S}$ and $h \in H$, the following duality formula holds:*

$$\mathbb{E}(\langle u, h\rangle_H G) = \mathbb{E}(F\delta(Gh)). \tag{16}$$

Then $F \in \mathbb{D}^{1,2}$ and $DF = u$.

Proof. Let $F = \sum_{n=0}^{\infty} I_n(f_n)$, where $f_n \in L_s^2(\mathbb{R}_+^n)$. By the duality formula (Proposition 2.1) and Proposition 3.2, we obtain

$$\begin{aligned} \mathbb{E}(F\delta(Gh)) &= \sum_{n=0}^{\infty} \mathbb{E}(I_n(f_n)\delta(Gh)) = \sum_{n=0}^{\infty} \mathbb{E}(\langle D(I_n(f_n)), h\rangle_H G) \\ &= \sum_{n=1}^{\infty} \mathbb{E}(\langle nI_{n-1}(f_n(\cdot,t)), h\rangle_H G). \end{aligned}$$

Then, by (16), we get

$$\sum_{n=1}^{\infty} \mathbb{E}(\langle nI_{n-1}(f_n(\cdot,t)), h\rangle_H G) = \mathbb{E}(\langle u, h\rangle_H G),$$

which implies that the series $\sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot,t))$ converges in $L^2(\Omega; H)$ and its sum coincides with u. Proposition 3.3 allows us to conclude the proof. □

Corollary 3.4. *Let $(F_n)_{n\geq 1}$ be a sequence of random variables in $\mathbb{D}^{1,2}$ that converges to F in $L^2(\Omega)$ and is such that*

$$\sup_n \mathbb{E}(\|DF_n\|_H^2) < \infty.$$

Then F belongs to $\mathbb{D}^{1,2}$ and the sequence of derivatives $(DF_n)_{n\geq 1}$ converges to DF in the weak topology of $L^2(\Omega; H)$.

Proof. The assumptions imply that there exists a subsequence $(F_{n(k)})_{k\geq 1}$ such that the sequence of derivatives $(DF_{n(k)})_{k\geq 1}$ converges in the weak topology of $L^2(\Omega; H)$ to some element $\alpha \in L^2(\Omega; H)$. By Proposition 3.4, it suffices to show that, for all $G \in \mathcal{S}$ and $h \in H$,

$$\mathbb{E}(\langle \alpha, h\rangle_H G) = \mathbb{E}(F\delta(Gh)). \tag{17}$$

By the duality formula (Proposition 2.1), we have

$$\mathbb{E}(\langle DF_{n(k)}, h\rangle_H G) = \mathbb{E}(F_{n(k)}\delta(Gh)).$$

Then, taking the limit as k tends to infinity, we obtain (17), which concludes the proof. □

The next proposition shows that the indicator function of a set $A \in \mathcal{F}$ such that $0 < P(A) < 1$ does not belong to $\mathbb{D}^{1,2}$.

Proposition 3.5. *Let $A \in \mathcal{F}$ and suppose that the indicator function of A belongs to the space $\mathbb{D}^{1,2}$. Then, $P(A)$ is zero or one.*

Proof. Consider a continuously differentiable function φ with compact support, such that $\varphi(x) = x^2$ for each $x \in [0,1]$. Then, by Proposition 2.4, we can write

$$D\mathbf{1}_A = D[(\mathbf{1}_A)^2] = D[\varphi(\mathbf{1}_A)] = 2\mathbf{1}_A D\mathbf{1}_A.$$

Therefore $D\mathbf{1}_A = 0$ and, from Proposition 3.3, we deduce that $\mathbf{1}_A = P(A)$, which completes the proof. □

3.3 Divergence on the Wiener chaos

We now compute the divergence operator on the Wiener chaos expansion. A square integrable stochastic process $u = (u_t)_{t\geq 0} \in L^2(\Omega \times \mathbb{R}_+)$ has an orthogonal expansion of the form

$$u_t = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)),$$

where $f_0(t) = \mathbb{E}(u_t)$ and, for each $n \geq 1$, $f_n \in L^2(\mathbb{R}_+^{n+1})$ is a symmetric function in the first n variables.

Proposition 3.6. *The process u belongs to the domain of δ if and only if the series*

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \tag{18}$$

converges in $L^2(\Omega)$.

Proof. Suppose that $G = I_n(g)$ is a multiple stochastic integral of order $n \geq 1$, where g is symmetric. Then

$$\begin{aligned}
\mathbb{E}(\langle u, DG\rangle_H) &= \int_{\mathbb{R}_+} \mathbb{E}\left(I_{n-1}(f_{n-1}(\cdot,t)) n I_{n-1}(g(\cdot,t))\right) dt \\
&= n(n-1)! \int_{\mathbb{R}_+} \langle f_{n-1}(\cdot,t), g(\cdot,t)\rangle_{L^2(\mathbb{R}_+^{n-1})}\, dt \\
&= n! \langle f_{n-1}, g\rangle_{L^2(\mathbb{R}_+^n)} = n! \langle \tilde{f}_{n-1}, g\rangle_{L^2(\mathbb{R}_+^n)} \\
&= \mathbb{E}(I_n(\tilde{f}_{n-1}) I_n(g)) = \mathbb{E}(I_n(\tilde{f}_{n-1}) G).
\end{aligned}$$

If $u \in \operatorname{Dom}\delta$, we deduce that

$$\mathbb{E}(\delta(u)G) = \mathbb{E}(I_n(\tilde{f}_{n-1})G)$$

for every $G \in \mathcal{H}_n$. This implies that $I_n(\tilde{f}_{n-1})$ coincides with the projection of $\delta(u)$ on the nth Wiener chaos. Consequently, the series in (18) converges in $L^2(\Omega)$ and its sum is equal to $\delta(u)$. The converse can be proved by similar arguments. □

4 Ornstein-Uhlenbeck semigroup. Meyer inequalities

In this section we describe the main properties of the Ornstein–Uhlenbeck semigroup and its generator. We then give the relationship between the Malliavin derivative, the divergence operator, and the Ornstein–Uhlenbeck semigroup generator.

4.1 Mehler's formula

Let $B = (B_t)_{t\geq 0}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}$ is generated by B. Let F be a random variable in $L^2(\Omega)$ with the Wiener chaos decomposition $F = \sum_{n=0}^{\infty} I_n(f_n)$, $f_n \in L^2_s(\mathbb{R}^n_+)$.

Definition 4.1. *The Ornstein–Uhlenbeck semigroup is the one-parameter semigroup $(T_t)_{t\geq 0}$ of operators on $L^2(\Omega)$ defined by*

$$T_t(F) = \sum_{n=0}^{\infty} e^{-nt} I_n(f_n).$$

An alternative and useful expression for the Ornstein–Uhlenbeck semigroup is *Mehler's formula*:

Proposition 4.1. *Let $B' = (B'_t)_{t\geq 0}$ be an independent copy of B. Then, for any $t \geq 0$ and $F \in L^2(\Omega)$, we have*

$$T_t(F) = \mathbb{E}'(F(e^{-t}B + \sqrt{1-e^{-2t}}B')), \tag{19}$$

where $\mathbb{E}'$ denotes the mathematical expectation with respect to B'.

Proof. Both T_t in Definition 4.1 and the right-hand side of (19) give rise to linear contraction operators on $L^p(\Omega)$, for all $p \geq 1$. For the first operator, this is clear. For the second, using Jensen's inequality it follows that, for any $p \geq 1$,

$$\begin{aligned}\mathbb{E}(|T_t(F)|^p) &= \mathbb{E}(|\mathbb{E}'(F(e^{-t}B + \sqrt{1-e^{-2t}}B'))|^p) \\ &\leq \mathbb{E}(\mathbb{E}'(|F(e^{-t}B + \sqrt{1-e^{-2t}}B')|^p)) = \mathbb{E}(|F|^p).\end{aligned}$$

Thus, it suffices to show (19) for random variables of the form $F = \exp\left(\lambda B(h) - \frac{1}{2}\lambda^2\right)$, where $B(h) = \int_{\mathbb{R}_+} h_t dB_t$, $h \in H$, is an element of norm one, and $\lambda \in \mathbb{R}$. We have, using formula (1),

$$\begin{aligned}&\mathbb{E}'\left(\exp\left(e^{-t}\lambda B(h) + \sqrt{1-e^{-2t}}\lambda B'(h) - \tfrac{1}{2}\lambda^2\right)\right) \\ &\quad = \exp\left(e^{-t}\lambda B(h) - \tfrac{1}{2}e^{-2t}\lambda^2\right) = \sum_{n=0}^{\infty} e^{-nt}\frac{\lambda^n}{n!} H_n\left(B(h)\right) = T_t F,\end{aligned}$$

because

$$F = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} H_n\left(B(h)\right)$$

and $H_n(B(h)) = I_n(h^{\otimes n})$ (see Proposition 3.1). This completes the proof. □

Mehler's formula implies that the operator T_t is nonnegative. Moreover, T_t is symmetric, that is,

$$\mathbb{E}(GT_t(F)) = \mathbb{E}(FT_t(G)) = \sum_{n=0}^{\infty} e^{-nt}\mathbb{E}(I_n(f_n)I_n(g_n)),$$

where $F = \sum_{n=0}^{\infty} I_n(f_n)$ and $G = \sum_{n=0}^{\infty} I_n(g_n)$.

The Ornstein–Uhlenbeck semigroup has the following hypercontractivity property

Theorem 4.2. *Let $F \in L^p(\Omega)$, $p > 1$, and $q(t) = e^{2t}(p-1)+1 > p$, $t > 0$. Then*

$$\|T_tF\|_{q(t)} \leq \|F\|_p.$$

As a consequence of the hypercontractivity property, for any $1 < p < q < \infty$ the norms $\|\cdot\|_p$ and $\|\cdot\|_q$ are equivalent on any Wiener chaos $\mathcal{H}_n$. In fact, putting $q = e^{2t}(p-1)+1 > p$ with $t > 0$, we obtain, for every $F \in \mathcal{H}_n$,

$$e^{-nt}\|F\|_q = \|T_tF\|_q \leq \|F\|_p,$$

which implies that

$$\|F\|_q \leq \left(\frac{q-1}{p-1}\right)^{n/2} \|F\|_p. \tag{20}$$

Moreover, for any $n \geq 1$ and $1 < p < \infty$, the orthogonal projection onto the nth Wiener chaos J_n is bounded in $L^p(\Omega)$, and

$$\|J_nF\|_p \leq \begin{cases} (p-1)^{n/2}\|F\|_p & \text{if } p > 2, \\ (p-1)^{-n/2}\|F\|_p & \text{if } p < 2. \end{cases} \tag{21}$$

In fact, suppose first that $p > 2$ and let $t > 0$ be such that $p - 1 = e^{2t}$. Using the hypercontractivity property with exponents p and 2, we obtain

$$\|J_nF\|_p = e^{nt}\|T_tJ_nF\|_p \leq e^{nt}\|J_nF\|_2 \leq e^{nt}\|F\|_2 \leq e^{nt}\|F\|_p.$$

If $p < 2$, we have

$$\|J_nF\|_p = \sup_{\|G\|_q \leq 1} \mathbb{E}((J_nF)G) \leq \|F\|_p \sup_{\|G\|_q \leq 1} \|J_nG\|_q \leq e^{nt}\|F\|_p,$$

where q is the conjugate of p, and $q - 1 = e^{2t}$.

As an application we can establish the following lemma.

Lemma 4.3. *Fix an integer $k \geq 1$ and a real number $p > 1$. Then, there exists a constant $c_{p,k}$ such that, for any random variable $F \in \mathbb{D}^{k,2}$,*

$$\|\mathbb{E}(D^kF)\|_{H^{\otimes k}} \leq c_{p,k}\|F\|_p.$$

Proof. Suppose that $F = \sum_{n=0}^{\infty} I_n(f_n)$. Then, by Stroock's formula (15), $\mathbb{E}(D^kF) = k!f_k$. Therefore,

$$\|\mathbb{E}(D^kF)\|_{H^{\otimes k}} = k!\|f_k\|_{H^{\otimes k}} = \sqrt{k!}\|J_kF\|_2.$$

From (20) we obtain

$$\|J_kF\|_2 \leq \big((p-1) \wedge 1\big)^{-k/2}\|J_kF\|_p.$$

Finally, applying (21) we get

$$\|J_kF\|_p \leq (p-1)^{\text{sign}(p-2)k/2}\|F\|_p,$$

which concludes the proof. □

The next result can be regarded as a regularizing property of the Ornstein–Uhlenbeck semigroup.

Proposition 4.2. *Let $F \in L^p(\Omega)$ for some $p > 1$. Then, for any $t > 0$, we have that $T_tF \in \mathbb{D}^{1,p}$ and there exists a constant c_p such that*

$$\|DT_tF\|_{L^p(\Omega;H)} \leq c_p t^{-1/2} \|F\|_p. \tag{22}$$

Proof. Consider a sequence of smooth and cylindrical random variables $F_n \in \mathcal{S}$ which converges to F in $L^p(\Omega)$. We know that T_tF_n converges to T_tF in $L^p(\Omega)$. We have $T_tF_n \in \mathbb{D}^{1,p}$, and using Mehler's formula (19), we can write

$$\begin{aligned} D(T_t(F_n - F_m)) &= D\left(\mathbb{E}'((F_n - F_m)(e^{-t}B + \sqrt{1-e^{-2t}})B')\right) \\ &= \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \mathbb{E}'\left(D'((F_n - F_m)(e^{-t}B + \sqrt{1-e^{-2t}})B'))\right). \end{aligned}$$

Then, Lemma 4.3 implies that

$$\|D(T_t(F_n - F_m))\|_{L^p(\Omega;H)} \leq c_{p,1} \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \|F_n - F_m\|_p.$$

Hence, DT_tF_n is a Cauchy sequence in $L^p(\Omega; H)$. Therefore, $T_tF \in \mathbb{D}^{1,p}$ and DT_tF is the limit in $L^p(\Omega; H)$ of DT_tF_n. The estimate (22) follows by the same arguments. □

With the above ingredients, we can show an extension of Corollary 3.4 to any $p > 1$.

Proposition 4.3. *Let $F_n \in \mathbb{D}^{1,p}$ be a sequence of random variables converging to F in $L^p(\Omega)$ for some $p > 1$. Suppose that*

$$\sup_n \|F_n\|_{1,p} < \infty.$$

Then $F \in \mathbb{D}^{1,p}$.

Proof. The assumptions imply that there exists a subsequence $(F_{n(k)})_{k\geq 1}$ such that the sequence of derivatives $(DF_{n(k)})_{k\geq 1}$ converges in the weak topology of $L^q(\Omega; H)$ to some element $\alpha \in L^q(\Omega; H)$, where $1/p + 1/q = 1$. By Proposition 4.2, for any $t > 0$, we have that T_tF belongs to $\mathbb{D}^{1,p}$ and $DT_tF_{n(k)}$ converges to DT_tF in $L^p(\Omega; H)$. Then, for any $\beta \in L^q(\Omega; H)$, we can write

$$\begin{aligned} \mathbb{E}(\langle DT_tF, \beta\rangle_H) &= \lim_{k\to\infty} \mathbb{E}(\langle DT_tF_{n(k)}, \beta\rangle_H) = \lim_{k\to\infty} e^{-t}\mathbb{E}(\langle T_tDF_{n(k)}, \beta\rangle_H) \\ &= \lim_{k\to\infty} e^{-t}\mathbb{E}(\langle DF_{n(k)}, T_t\beta\rangle_H) = e^{-t}\mathbb{E}(\langle \alpha, T_t\beta\rangle_H) \\ &= \mathbb{E}(\langle e^{-t}T_t\alpha, \beta\rangle_H). \end{aligned}$$

Therefore, $DT_tF = e^{-t}T_t\alpha$. This implies that DT_tF converges to α as $t \downarrow 0$ in $L^p(\Omega; H)$. Using that D is a closed operator, we conclude that $F \in \mathbb{D}^{1,p}$ and $DF = \alpha$. □

4.2 Generator of the Ornstein-Uhlenbeck semigroup

The generator of the Ornstein–Uhlenbeck semigroup in $L^2(\Omega)$ is the operator given by

$$LF = \lim_{t \downarrow 0} \frac{T_t F - F}{t},$$

and the domain of L is the set of random variables $F \in L^2(\Omega)$ for which the above limit exists in $L^2(\Omega)$. It is easy to show that a random variable $F = \sum_{n=0}^{\infty} I_n(f_n)$, $f_n \in L^2_s(\mathbb{R}^n_+)$, belongs to the domain of L if and only if

$$\sum_{n=1}^{\infty} n^2 \|I_n(f_n)\|_2^2 < \infty;$$

and, in this case, $LF = \sum_{n=1}^{\infty} -nI_n(f_n)$. Thus, $\operatorname{Dom} L$ coincides with the space $\mathbb{D}^{2,2}$.

We also define the operator L^{-1}, which is the pseudo-inverse of L, as follows. For every $F \in L^2(\Omega)$, set

$$LF = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n).$$

Note that L^{-1} is an operator with values in $\mathbb{D}^{2,2}$ and that $LL^{-1}F = F - \mathbb{E}(F)$, for any $F \in L^2(\Omega)$, so L^{-1} acts as the inverse of L for centered random variables.

The next proposition explains the relationship between the operators D, δ, and L.

Proposition 4.4. *Let $F \in L^2(\Omega)$. Then, $F \in \operatorname{Dom} L$ if and only if $F \in \mathbb{D}^{1,2}$ and $DF \in \operatorname{Dom} \delta$ and, in this case, we have*

$$\delta DF = -LF.$$

Proof. Let $F = \sum_{n=0}^{\infty} I_n(f_n)$. Suppose first that $F \in \mathbb{D}^{1,2}$ and $DF \in \operatorname{Dom} \delta$. Then, for any random variable $G = I_m(g_m)$, we have, using the duality relationship (Proposition 2.1),

$$\mathbb{E}(G\delta DF) = \mathbb{E}(\langle DG, DF \rangle_H) = mm! \langle g_m, f_m \rangle_{L^2(\mathbb{R}^m_+)} = \mathbb{E}(GmI_m(f_m)).$$

Therefore, the projection of δDF onto the mth Wiener chaos is equal to $mI_m(f_m)$. This implies that the series $\sum_{n=1}^{\infty} nI_n(f_n)$ converges in $L^2(\Omega)$ and its sum is δDF. Therefore, $F \in \operatorname{Dom} L$ and $LF = -\delta DF$.

Conversely, suppose that $F \in \operatorname{Dom} L$. Clearly, $F \in \mathbb{D}^{1,2}$. Then, for any random variable $G \in \mathbb{D}^{1,2}$ with Wiener chaos expansion $G = \sum_{n=0}^{\infty} I_n(g_n)$, we have

$$\mathbb{E}(\langle DG, DF \rangle_H) = \sum_{n=1}^{\infty} nn! \langle g_n, f_n \rangle_{L^2(\mathbb{R}^n_+)} = -\mathbb{E}(GLF).$$

As a consequence, DF belongs to the domain of δ and $\delta DF = -LF$. □

The operator L behaves as a second-order differential operator on smooth random variables.

Proposition 4.5. *Suppose that $F = (F^1, \dots, F^m)$ is a random vector whose components belong to $\mathbb{D}^{2,4}$. Let φ be a function in $C^2(\mathbb{R}^m)$ with bounded first and second partial derivatives. Then, $\varphi(F) \in \operatorname{Dom} L$ and*

$$L(\varphi(F)) = \sum_{i,j=1}^{m} \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(F) \langle DF^i, DF^j \rangle_H + \sum_{i=1}^{m} \frac{\partial \varphi}{\partial x_i}(F) LF^i.$$

Proof. By the chain rule (see Proposition 2.4), $\varphi(F)$ belongs to $\mathbb{D}^{1,2}$ and

$$D(\varphi(F)) = \sum_{i=1}^{m} \frac{\partial \varphi}{\partial x_i}(F) DF^i.$$

Moreover, by Proposition 4.4, $\varphi(F)$ belongs to $\operatorname{Dom} L$ and $L(\varphi(F)) = -\delta(D(\varphi(F)))$. Using the factorization property of the divergence operator yields the result. □

In the finite-dimensional case ($\Omega = \mathbb{R}^n$ equipped with the standard Gaussian law), $L = \Delta - x \cdot \nabla$ coincides with the generator of the Ornstein–Uhlenbeck process $(X_t)_{t\geq 0}$ in $\mathbb{R}^n$, which is the solution to the stochastic differential equation

$$dX_t = \sqrt{2} dB_t - X_t dt,$$

where $(B_t)_{t\geq 0}$ is an n-dimensional Brownian motion.

4.3 Meyer's inequality

The next theorem provides an estimate for the $L^p(\Omega)$-norm of the divergence operator for any $p > 1$. It was proved by Pisier, using the boundedness in $L^p(\mathbb{R})$ of the Riesz transform.

Theorem 4.4. *For any $p > 1$, there exists a constant $c_p > 0$ such that for any $u \in \mathbb{D}^{1,p}(H)$,*

$$\mathbb{E}(|\delta(u)|^p) \leq c_p \left(\mathbb{E}(\|Du\|^p_{L^2(\mathbb{R}^2_+)}) + \|\mathbb{E}(u)\|^p_H \right). \tag{23}$$

As a consequence of Theorem 4.4, the divergence operator is continuous from $\mathbb{D}^{1,p}(H)$ to $L^p(\Omega)$, and so we have *Meyer's inequality*:

$$\mathbb{E}(|\delta(u)|^p) \leq c_p \left(\mathbb{E}(\|Du\|^p_{L^2(\mathbb{R}^2_+)}) + \mathbb{E}(\|u\|^p_H) \right) = c_p \|u\|^p_{1,p,H}. \tag{24}$$

This result can be extended as follows.

Theorem 4.5. *For any $p > 1$, $k \geq 1$, and $u \in \mathbb{D}^{k,p}(H)$,*

$$\|\delta(u)\|_{k-1,p} \leq c_{k,p} \left(\mathbb{E}(\|D^k u\|^p_{L^2(\mathbb{R}^{k+1}_+)}) + \mathbb{E}(\|u\|^p_H) \right) = c_{k,p} \|u\|^p_{k,p,H}.$$

This implies that the operator δ is continuous from $\mathbb{D}^{k,p}(H)$ into $\mathbb{D}^{k-1,p}(H)$.

5 Stochastic integral representations. Clark-Ocone formula

This section deals with the following problem. Given a random variable F in $L^2(\Omega)$, with $\mathbb{E}(F) = 0$, find a stochastic process u in $\operatorname{Dom} \delta$ such that $F = \delta(u)$. We present two different answers to this question, both integral representations. The first is the Clark–Ocone formula, in which u is required to be adapted. Therefore, the process u is unique and its expression involves a conditional expectation of the Malliavin derivative of F. The second uses the inverse of the Ornstein–Uhlenbeck generator. We then present some applications of these integral representations.

5.1 Clark-Ocone formula

Let $B = (B_t)_{t\geq 0}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}$ is generated by B, equipped with its Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$. The next result expresses the integrand of the integral representation theorem of a square integrable random variable in terms of the conditional expectation of its Malliavin derivative.

Theorem 5.1 (Clark–Ocone formula)**.** *Let $F \in \mathbb{D}^{1,2} \cap L^2(\Omega, \mathcal{F}_T, P)$. Then F admits the following representation:*

$$F = \mathbb{E}(F) + \int_0^T \mathbb{E}(D_t F|\mathcal{F}_t)dB_t.$$

Proof. By the Itô integral representation theorem, there exists a unique process adated process $u \in L^2(\Omega \times [0,T])$ such that $F \in L^2(\Omega, \mathcal{F}_T, P)$ admits the stochastic integral representation

$$F = \mathbb{E}(F) + \int_0^T u_t dB_t.$$

It suffices to show that $u_t = \mathbb{E}(D_t F|\mathcal{F}_t)$ for almost all $(t,\omega) \in [0,T] \times \Omega$. Consider a process $v \in L^2_T(\mathcal{P})$. On the one hand, the isometry property yields

$$\mathbb{E}(\delta(v)F) = \int_0^T \mathbb{E}(v_s u_s)ds.$$

On the other hand, by the duality relationship (Proposition 2.1), and taking into account that v is progressively measurable,

$$\mathbb{E}(\delta(v)F) = \mathbb{E}\left(\int_0^T v_t D_t F dt\right) = \int_0^T \mathbb{E}(v_s \mathbb{E}(D_t F|\mathcal{F}_t))dt.$$

Therefore, $u_t = \mathbb{E}(D_t F|\mathcal{F}_t)$ for almost all $(t,\omega) \in [0,T] \times \Omega$, which concludes the proof. □

Consider the following simple examples of the application of this formula.

Example 5.2. *Suppose that $F = B_t^3$. Then $D_s F = 3B_t^2 \mathbf{1}_{[0,t]}(s)$ and*

$$\mathbb{E}(D_s F|\mathcal{F}_s) = 3\mathbb{E}((B_t - B_s + B_s)^2|\mathcal{F}_s) = 3(t - s + B_s^2).$$

Therefore

$$B_t^3 = 3\int_0^t (t - s + B_s^2)dB_s. \tag{25}$$

This formula should be compared with Itô's formula,

$$B_t^3 = 3\int_0^t B_s^2 dB_s + 3\int_0^t B_s ds. \tag{26}$$

Notice that equation (25) contains only a stochastic integral but it is not a martingale, because the integrand depends on t, whereas (26) contains two terms and one is a martingale. Moreover, the integrand in (25) is unique.

Example 5.3. *Consider the Brownian motion local time $(L_t^x)_{t\geq 0, x\in\mathbb{R}}$. For any $\varepsilon > 0$, we set*

$$p_\varepsilon(x) = (2\pi\varepsilon)^{-1/2} e^{-x^2/(2\varepsilon)}.$$

We have that, as $\epsilon \to 0$,

$$F_\varepsilon = \int_0^t p_\varepsilon(B_s - x)ds \overset{L^2(\Omega)}{\longrightarrow} L_t^x. \tag{27}$$

Applying the derivative operator yields

$$D_r F_\epsilon = \int_0^t p'_\epsilon(B_s - x) D_r B_s ds = \int_r^t p'_\epsilon(B_s - x)ds.$$

Thus

$$\begin{aligned} \mathbb{E}(D_r F_\epsilon | \mathcal{F}_r) &= \int_r^t \mathbb{E}(p'_\epsilon(B_s - B_r + B_r - x)|\mathcal{F}_r)ds \\ &= \int_r^t p'_{\epsilon+s-r}(B_r - x)ds. \end{aligned}$$

As a consequence, taking the limit as $\epsilon \to 0$, we obtain the following integral representation of the Brownian local time:

$$L_t^x = \mathbb{E}(L_t^x) + \int_0^t \varphi(t - r, B_r - x) dB_r,$$

where

$$\varphi(r, y) = \int_0^r p'_s(y)ds.$$

5.2 Second integral representation

Recall that L is the generator of the Ornstein–Uhlenbeck semigroup.

Proposition 5.1. *Let F be in $\mathbb{D}^{1,2}$ with $\mathbb{E}(F) = 0$. Then the process*

$$u = -DL^{-1}F$$

belongs to $\operatorname{Dom}\delta$ *and satisfies $F = \delta(u)$. Moreover $u \in L^2(\Omega; H)$ is unique among all square integrable processes with a chaos expansion*

$$u_t = \sum_{q=0}^{\infty} I_q(f_q(t))$$

such that $f_q(t, t_1, \ldots, t_q)$ is symmetric in all $q + 1$ variables $t, t_1, \ldots, t_q$.

Proof. By Proposition 4.4,

$$F = LL^{-1}F = -\delta(DL^{-1}F).$$

Clearly, the process $u = -DL^{-1}F$ has a Wiener chaos expansion with functions symmetric in all their variables. To show uniqueness, let $v \in L^2(\Omega; H)$ with a chaos expansion $v_t = \sum_{q=0}^{\infty} I_q(g_q(t))$, such that the function $g_q(t, t_1, \ldots, t_q)$ is symmetric in all $q + 1$ variables

$t, t_1, \dots, t_q$ and such that $\delta(v) = F$. Then, there exists a random variable $G \in \mathbb{D}^{1,2}$ such that $DG = v$. Indeed, it suffices to take

$$G = \sum_{q=0}^{\infty} \frac{1}{q+1} I_{q+1}(g_q).$$

We claim that $G = -L^{-1}F$. This follows from $LG = -\delta DG = -\delta(v) = -F$. The proof is now complete. □

It is important to notice that, unlike the Clark–Ocone formula, which requires that the underlying process is a Brownian motion, the representation provided in Proposition 5.1 holds in the context of a general Gaussian isonormal process.

6 Existence and regularity of densities. Density formulas

In this Section we apply Malliavin calculus to derive explicit formulas for the densities of random variables on Wiener space and to establish criteria for their regularity.

6.1 Analysis of densities in the one-dimensional case

We recall that $B = (B_t)_{t \geq 0}$ is a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}$ is generated by B. The topological support of the law of a random variable F is defined as the set of points $x \in \mathbb{R}$ such that $P(|x - F| < \epsilon) > 0$ for all $\epsilon > 0$.

Our first result says that if a random variable F belongs to the Sobolev space $\mathbb{D}^{1,2}$ then the topological support of the law of F is a closed interval.

Proposition 6.1. *Let $F \in \mathbb{D}^{1,2}$. Then, the topological support of the law of F is a closed interval.*

Proof. Clearly the topological support of the law of F is a closed set. Then, it suffices to show that it is connected. We show this by contradiction. If the topological support of the law of F is not connected, there exists a point $a \in \mathbb{R}$ and $\epsilon > 0$ such that $P(a - \epsilon < F < a + \epsilon) = 0$, $P(F \geq a + \epsilon) < 1$, and $P(F \leq a - \epsilon) < 1$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function such that $\varphi(x) = 0$ if $x \leq a - \epsilon$ and $\varphi(x) = 1$ if $x \geq a + \epsilon$. By Proposition 2.4, $\varphi(F) \in \mathbb{D}^{1,2}$ but, almost surely, $\varphi(F) = \mathbf{1}_{\{F \geq a+\epsilon\}}$. Therefore, by Proposition 3.5, we must have $P(F \geq a + \epsilon) = 0$ or $P(F \geq a + \epsilon) = 1$, which leads to a contradiction. □

If a random variable F belongs to $\mathbb{D}^{1,2}$, and its derivative is not degenerate, then F has a density.

Proposition 6.2. *Let F be a random variable in the space $\mathbb{D}^{1,2}$ such that $\|DF\|_H > 0$ almost surely. Then, the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

Proof. Replacing F by $\arctan F$, we may assume that F takes values in $(-1, 1)$. It suffices to show that, for any measurable function $g : (-1, 1) \to [0, 1]$ such that $\int_{-1}^{1} g(y)dy = 0$, we have $\mathbb{E}(g(F)) = 0$. We can find a sequence of continuous functions $g_n : (-1, 1) \to [0, 1]$ such that, as n tends to infinity, $g_n(y)$ converges to $g(y)$ for almost all y with respect to the measure $P \circ F^{-1} + \ell$, where ℓ denotes the Lebesgue measure on $\mathbb{R}$. Set

$$\psi_n(x) = \int_{-\infty}^{x} g_n(y)dy.$$

Then, $\psi_n(F)$ converges to 0 almost surely and in $L^2(\Omega)$ because g_n converges almost everywhere to g, with respect to the Lebesgue measure, and $\int_{-1}^{1} g(y)dy = 0$. Furthermore, by the chain rule (Proposition 2.4), $\psi_n(F) \in \mathbb{D}^{1,2}$ and

$$D(\psi_n(F)) = g_n(F)DF,$$

which converges almost surely and in $L^2(\Omega)$ to $g(F)DF$. Because D is closed, we conclude that $g(F)DF = 0$. Our hypothesis $\|DF\|_H > 0$ implies that $g(F) = 0$ almost surely, and this finishes the proof. □

The following result is an expression for the density of a random variable in the Sobolev space $\mathbb{D}^{1,2}$, assuming that $\|DF\|_H > 0$ a.s.

Proposition 6.3. *Let F be a random variable in the space $\mathbb{D}^{1,2}$ such that $\|DF\|_H > 0$ a.s. Suppose that $DF/\|DF\|_H^2$ belongs to the domain of the operator δ in $L^2(\Omega)$. Then the law of F has a continuous and bounded density, given by*

$$p(x) = \mathbb{E}\left(\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right). \tag{28}$$

Proof. Let ψ be a nonnegative function in $C_0^\infty(\mathbb{R})$, and set $\varphi(y) = \int_{-\infty}^{y} \psi(z)dz$. Then, by the chain rule (Proposition 2.4), $\varphi(F)$ belongs to $\mathbb{D}^{1,2}$ and we can write

$$\langle D(\varphi(F)), DF\rangle_H = \psi(F)\|DF\|_H^2.$$

Using the duality formula (Proposition 2.1), we obtain

$$\mathbb{E}(\psi(F)) = \mathbb{E}\left(\left\langle D(\varphi(F))\,, \frac{DF}{\|DF\|_H^2}\right\rangle_H\right) = \mathbb{E}\left(\varphi(F)\,\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right). \tag{29}$$

By an approximation argument, equation (29) holds for $\psi(y) = \mathbf{1}_{[a,b]}(y)$, where $a < b$. As a consequence, we can apply Fubini's theorem to get

$$\begin{aligned} P(a \le F \le b) &= \mathbb{E}\left(\left(\int_{-\infty}^{F} \psi(x)dx\right)\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right) \\ &= \int_a^b \mathbb{E}\left(\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right)dx, \end{aligned}$$

which implies the desired result. □

Remark 6.1. *Equation (28) still holds under the hypotheses $F \in \mathbb{D}^{1,p}$ and $DF/\|DF\|_H^2 \in \mathbb{D}^{1,p'}(H)$ for some $p, p' > 1$. Sufficient conditions for these hypotheses are $F \in \mathbb{D}^{2,\alpha}$ and $\mathbb{E}(\|DF\|^{-2\beta}) < \infty$ with $1/\alpha + 1/\beta < 1$.*

Example 6.2. *Let $F = B(h)$. Then $DF = h$ and*

$$\delta\left(\frac{DF}{\|DF\|_H^2}\right) = B(h)\|h\|_H^{-2}.$$

As a consequence, formula (28) yields

$$p(x) = \|h\|_H^{-2}\,\mathbb{E}\left(\mathbf{1}_{\{F>x\}}F\right),$$

which is true because $p(x)$ is the density of the distribution $N(0, \|h\|_H^2)$.

Applying equation (28) we can derive density estimates. Notice first that (28) holds if $\mathbf{1}_{\{F>x\}}$ is replaced by $\mathbf{1}_{\{F<x\}}$, because the divergence has zero expectation. Fix p and q such that $1/p + 1/q = 1$. Then, by Hölder's inequality, we obtain

$$p(x) \le (P(|F| > |x|))^{1/q} \left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_p, \tag{30}$$

for all $x \in \mathbb{R}$. Applying (30) and Meyer's inequality (24), we can deduce the following result.

Proposition 6.4. *Let q, α, β be three positive real numbers such that $1/q + 1/\alpha + 1/\beta = 1$. Let F be a random variable in the space $\mathbb{D}^{2,\alpha}$, such that $\mathbb{E}(\|DF\|_H^{-2\beta}) < \infty$. Then, the density $p(x)$ of F can be estimated as follows:*

$$\begin{aligned} p(x) \le \quad & c_{q,\alpha,\beta}\,(P(|F| > |x|))^{1/q} \\ & \times\left(\mathbb{E}(\|DF\|_H^{-1}) + \left\|D^2F\right\|_{L^\alpha(\Omega;L^2(\mathbb{R}_+^2))}\left\|\|DF\|_H^{-2}\right\|_\beta\right). \end{aligned} \tag{31}$$

6.2 Existence and smoothness of densities for random vectors

Let $F = (F^1, \dots, F^m)$ be such that $F^i \in \mathbb{D}^{1,2}$ for $i = 1, \dots, m$. We define the *Malliavin matrix* of F as the random symmetric nonnegative definite matrix

$$\gamma_F = (\langle DF^i, DF^j\rangle_H)_{1\le i,j\le m}. \tag{32}$$

In the one-dimensional case, $\gamma_F = \|DF\|_H^2$. The following theorem is a multidimensional version of Proposition 6.2.

Theorem 6.3. *If $\det\gamma_F > 0$ a.s. then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.*

This theorem was proved by Bouleau and Hirsch using the co-area formula and techniques of geometric measure theory, and we omit the proof. As a consequence, the measure $(\det\gamma_F \times P) \circ F^{-1}$ is always absolutely continuous; that is,

$$P(F \in B\,, \det\gamma_F > 0) = 0,$$

for any Borel set $B \in \mathcal{B}(\mathbb{R}^m)$ of zero Lebesgue measure.

Definition 6.4. *We say that a random vector $F = (F^1, \dots, F^m)$ is nondegenerate if $F^i \in \mathbb{D}^{1,2}$ for $i = 1, \dots, m$ and*

$$\mathbb{E}((\det\gamma_F)^{-p}) < \infty,$$

for all $p \ge 2$.

Set $\partial_i = \partial/\partial x_i$ and, for any multi-index $\alpha \in \{1, \ldots, m\}^k$, $k \geq 1$, we denote by ∂_α the partial derivative $\partial^k/(\partial x_{\alpha_1} \cdots \partial x_{\alpha_k})$.

Lemma 6.5. *Let γ be an $m \times m$ random matrix such that $\gamma^{ij} \in \mathbb{D}^{1,\infty}$ for all i, j and $\mathbb{E}(|\det \gamma|^{-p}) < \infty$ for all $p \geq 2$. Then, $\left(\gamma^{-1}\right)^{ij}$ belongs to $\mathbb{D}^{1,\infty}$ for all i, j, and*

$$D\left(\gamma^{-1}\right)^{ij} = -\sum_{k,\ell=1}^{m} \left(\gamma^{-1}\right)^{ik} \left(\gamma^{-1}\right)^{\ell j} D\gamma^{k\ell}. \tag{33}$$

Proof. It can be proved that $P(\det \gamma > 0)$ is zero or one. So, we can assume that $\det \gamma > 0$ almost surely. For any $\epsilon > 0$, we define $\gamma_\epsilon^{-1} = (\det \gamma + \epsilon)^{-1} A(\gamma)$, where $A(\gamma)$ is the adjoint matrix of γ. Then, the entries of γ_ϵ^{-1} belong to $\mathbb{D}^{1,\infty}$ and converge in $L^p(\Omega)$, for all $p \geq 2$, to those of γ^{-1} as ϵ tends to zero. Moreover, the entries of γ_ϵ^{-1} satisfy

$$\sup_{\epsilon \in (0,1]} \|(\gamma_\epsilon^{-1})^{ij}\|_{1,p} < \infty,$$

for all $p \geq 2$. Therefore, by Proposition 4.3 the entries of γ_ϵ^{-1} belong to $\mathbb{D}^{1,p}$ for any $p \geq 2$. Finally, from the expression $\gamma_\epsilon^{-1}\gamma = (\det \gamma/(\det \gamma + \epsilon)) I_m$, where I_m denotes the identity matrix of order m, we deduce (33) on applying the derivative operator and letting ϵ tend to zero. □

The following result can be regarded as an integration-by-parts formula and plays a fundamental role in the proof of the regularity of densities.

Proposition 6.5. *Let $F = (F^1, \ldots, F^m)$ be a nondegenerate random vector. Fix $k \geq 1$ and suppose that $F^i \in \mathbb{D}^{k+1,\infty}$ for $i = 1, \ldots, m$. Let $G \in \mathbb{D}^\infty$ and let $\varphi \in C_p^\infty(\mathbb{R}^m)$. Then, for any multi-index $\alpha \in \{1, \ldots, m\}^k$, there exists an element $H_\alpha(F, G) \in \mathbb{D}^\infty$ such that*

$$\mathbb{E}(\partial_\alpha \varphi(F) G) = \mathbb{E}(\varphi(F) H_\alpha(F, G)), \tag{34}$$

where the elements $H_\alpha(F, G)$ are recursively given by

$$H_{(i)}(F, G) = \sum_{j=1}^{m} \delta\left(G\left(\gamma_F^{-1}\right)^{ij} DF^j\right)$$

and, for $\alpha = (\alpha_1, \ldots, \alpha_k)$, $k \geq 2$, we set

$$H_\alpha(F, G) = H_{\alpha_k}(F, H_{(\alpha_1, \ldots, \alpha_{k-1})}(F, G)).$$

Proof. By the chain rule (Proposition 2.4), we have

$$\langle D(\varphi(F)), DF^j \rangle_H = \sum_{i=1}^{m} \partial_i \varphi(F) \langle DF^i, DF^j \rangle_H = \sum_{i=1}^{m} \partial_i \varphi(F) \gamma_F^{ij}$$

and, consequently,

$$\partial_i \varphi(F) = \sum_{j=1}^{m} \langle D(\varphi(F)), DF^j \rangle_H (\gamma_F^{-1})^{ji}.$$

Taking expectations and using the duality relationship (Proposition 2.1) yields

$$\mathbb{E}(\partial_i \varphi(F) G) = \mathbb{E}(\varphi(F) H_{(i)}(F, G)),$$

where $H_{(i)} = \sum_{j=1}^{m} \delta\left(G\left(\gamma_F^{-1}\right)^{ij} DF^j\right)$. Notice that Meyer's inequality (Theorem 4.4) and Lemma 6.5 imply that $H_{(i)}$ belongs to $L^p(\Omega)$ for any $p \geq 2$. We finish the proof with a recurrence argument. □

One can show that, for any $p > 1$, there exist constants β, $\gamma > 1$ and integers n, m such that

$$\|H_\alpha(F, G)\|_p \leq c_{p,q} \left\|\det \gamma_F^{-1}\right\|_\beta^m \|DF\|_{k,\gamma}^n \|G\|_{k,q}. \tag{35}$$

The proof of this inequality is based on Meyer's and Hölder's inequalities.

The following result is a multidimensional version of the density formula (28).

Proposition 6.6. *Let $F = (F^1, \ldots, F^m)$ be a nondegenerate random vector such that $F^i \in \mathbb{D}^{m+1,\infty}$ for $i = 1, \ldots, m$. Then F has a continuous and bounded density given by*

$$p(x) = \mathbb{E}(\mathbf{1}_{\{F > x\}} H_\alpha(F, 1)), \tag{36}$$

where $\alpha = (1, 2, \ldots, m)$.

Proof. Recall that, for $\alpha = (1, 2, \ldots, m)$

$$\begin{aligned} & H_\alpha(F, 1) \\ & = \sum_{j_1, \ldots, j_m = 1}^{m} \delta\left((\gamma_F^{-1})^{1 j_1} DF^{j_1} \delta\left((\gamma_F^{-1})^{2 j_2} DF^{j_2} \cdots \delta\left((\gamma_F^{-1})^{m j_m} DF^{j_m}\right) \cdots\right)\right). \end{aligned}$$

Then, equality (34) applied to the multi-index $\alpha = (1, 2, \ldots, m)$ yields, for any $\varphi \in C_p^\infty(\mathbb{R}^m)$,

$$\mathbb{E}(\partial_\alpha \varphi(F)) = \mathbb{E}(\varphi(F) H_\alpha(F, 1)).$$

Notice that

$$\varphi(F) = \int_{-\infty}^{F^1} \cdots \int_{-\infty}^{F^m} \partial_\alpha \varphi(x) dx.$$

Hence, by Fubini's theorem we can write

$$\mathbb{E}(\partial_\alpha \varphi(F)) = \int_{\mathbb{R}^m} \partial_\alpha \varphi(x) \mathbb{E}(\mathbf{1}_{\{F > x\}} H_\alpha(F, 1)) dx. \tag{37}$$

Given any function $\psi \in C_0^\infty(\mathbb{R}^m)$, we can take $\varphi \in C_p^\infty(\mathbb{R}^m)$ such that $\psi = \partial_\alpha \varphi$, and (37) yields

$$\mathbb{E}(\psi(F)) = \int_{\mathbb{R}^m} \psi(x) \mathbb{E}(\mathbf{1}_{\{F > x\}} H_\alpha(F, 1)) dx,$$

which implies the result. □

The following theorem is the basic criterion for the smoothness of densities.

Theorem 6.6. *Let $F = (F^1, \ldots, F^m)$ be a nondegenerate random vector such that $F^i \in \mathbb{D}^\infty$ for all $i = 1, \ldots, m$. Then the law of F possesses an infinitely differentiable density.*

Proof. For any multi-index β and any $\varphi \in C_p^\infty(\mathbb{R}^m)$, we have, taking $\alpha = (1, 2, \dots, m)$,

$$\begin{aligned}\mathbb{E}\left(\partial_\beta \partial_\alpha \varphi(F)\right) &= \mathbb{E}\left(\varphi(F) H_\beta(F, H_\alpha(F, 1))\right) \\ &= \int_{\mathbb{R}^m} \partial_\alpha \varphi(x)\, \mathbb{E}\left(\mathbf{1}_{\{F>x\}} H_\beta(F, H_\alpha(F, 1))\right) dx.\end{aligned}$$

Hence, for any $\xi \in C_0^\infty(\mathbb{R}^m)$,

$$\int_{\mathbb{R}^m} \partial_\beta \xi(x) p(x) dx = \int_{\mathbb{R}^m} \xi(x)\, \mathbb{E}\left(\mathbf{1}_{\{F>x\}} H_\beta(F, H_\alpha(F, 1))\right) dx.$$

Therefore, $p(x)$ is infinitely differentiable and, for any multi-index β, we have

$$\partial_\beta p(x) = (-1)^{|\beta|}\, \mathbb{E}\left(\mathbf{1}_{\{F>x\}} H_\beta(F, (H_\alpha(F, 1))\right).$$

This completes the proof. □

6.3 Density formula using the Riesz transform

In this section we present a method for obtaining a density formula using the Riesz transform, following the methodology introduced by Malliavin and extensively studied by Bally and Caremellino. In contrast with (36), here we only need two derivatives, instead of $m+1$.

Let Q_m be the fundamental solution to the Laplace equation $\Delta Q_m = \delta_0$ on $\mathbb{R}^m$, $m \geq 2$. That is,

$$Q_2(x) = a_2^{-1} \ln \frac{1}{|x|}, \quad Q_m(x) = a_m^{-1} |x|^{2-m}, \;\; m > 2,$$

where a_m is the area of the unit sphere in $\mathbb{R}^m$. We know that, for any $1 \leq i \leq m$,

$$\partial_i Q_m(x) = -c_m \frac{x_i}{|x|^m}, \tag{38}$$

where $c_m = 2(m-2)/a_m$ if $m > 2$ and $c_2 = 2/a_2$. Notice that any function φ in $C_0^1(\mathbb{R}^m)$ can be written as

$$\varphi(x) = \nabla \varphi * \nabla Q_m(x) = \sum_{i=1}^m \int_{\mathbb{R}^m} \partial \varphi(x-y) \partial_i Q_m(y) dy. \tag{39}$$

Indeed,

$$\nabla \varphi * \nabla Q_m(x) = \varphi * \Delta Q_m(x) = \varphi(x).$$

Theorem 6.7. *Let F be an m-dimensional nondegenerate random vector whose components are in $\mathbb{D}^{2,\infty}$. Then, the law of F admits a continuous and bounded density p given by*

$$p(x) = \sum_{i=1}^m \mathbb{E}(\partial_i Q_m(F-x) H_{(i)}(F, 1)),$$

where

$$H_{(i)}(F, 1) = \sum_{j=1}^m \delta\left((\gamma_F^{-1})^{ij} DF^j\right).$$

Proof. Let $\varphi \in C_0^1(\mathbb{R}^m)$. Applying (39), we can write

$$\mathbb{E}(\varphi(F)) = \sum_{i=1}^{m} \mathbb{E}\left(\int_{\mathbb{R}^m} \partial_i Q_m(y)(\partial_i \varphi(F-y))dy \right).$$

Assume that the support of φ is included in the ball $B_R(0)$ for some $R > 1$. Then, using (38) we obtain

$$\begin{aligned}
\mathbb{E}\left(\int_{\mathbb{R}^m} |\partial_i Q_m(y)\partial_i\varphi(F-y)|\, dy \right) &\le \|\partial_i \varphi\|_\infty \mathbb{E}\left(\int_{\{y:|F|-R\le|y|\le|F|+R\}} |\partial_i Q_m(y)| dy \right) \\
&\le c_m \mathrm{Vol}(B_1(0)) \|\partial_i\varphi\|_\infty \mathbb{E}\left(\int_{|F|-R}^{|F|+R} \frac{r}{r^m} r^{m-1} dr \right) \\
&= 2c_m \mathrm{Vol}(B_1(0)) \|\partial_i\varphi\|_\infty R\mathbb{E}(|F|) < \infty.
\end{aligned}$$

As a consequence, Fubini's theorem and (34) yield

$$\begin{aligned}
\mathbb{E}(\varphi(F)) &= \sum_{i=1}^{m} \int_{\mathbb{R}^m} \partial_i Q_m(y) \mathbb{E}(\partial_i \varphi(F-y)) dy \\
&= \sum_{i=1}^{m} \int_{\mathbb{R}^m} \partial_i Q_m(y) \mathbb{E}(\varphi(F-y) H_{(i)}(F,1)) dy \\
&= \sum_{i=1}^{m} \int_{\mathbb{R}^m} \varphi(y) \mathbb{E}(\partial_i Q_m(F-y) H_{(i)}(F,1)) dy.
\end{aligned}$$

This completes the proof. □

The approach based on the Riesz transform can also be used to obtain the following uniform estimate for densities, due to Stroock.

Lemma 6.8. *Under the assumptions of Theorem 6.7, for any $p > m$ there exists a constant c depending only on m and p such that*

$$\|p\|_\infty \le c \left(\max_{1\le i\le m} \|H_{(i)}(F,1)\|_p \right)^m .$$

Proof. From

$$p(x) = \sum_{i=1}^{m} \mathbb{E}(\partial_i Q_m(F-x) H_{(i)}(F,1)),$$

applying Hölder's inequality with $1/p + 1/q = 1$ and the estimate (see (38))

$$|\partial_i Q_m(F-x)| \le c_m |F-x|^{1-m}$$

yields

$$p(x) \le mc_m A \left(\mathbb{E}\left(|F-x|^{(1-m)q} \right) \right)^{1/q}, \tag{40}$$

where $A = \max_{1\le i\le m} \|H_{(i)}(F,1)\|_p$.

Suppose first that p is bounded and let $M = \sup_{x\in\mathbb{R}} p(x)$. We can write, for any $\epsilon > 0$,

$$\mathbb{E}(|F-x|^{(1-m)q}) \le \epsilon^{(1-m)q} + \int_{|z-x|\le\epsilon} |z-x|^{(1-m)q} p(x)dx \\ \le \epsilon^{(1-m)q} + C_{m,p}\epsilon^{(p-m)/(p-1)} M. \tag{41}$$

Therefore, substituting (41) into (40), we get

$$M \le Amc_m \left(\epsilon^{1-m} + C_{m,p}^{1/q}\epsilon^{(p-m)/p} M^{1/q}\right).$$

Now we minimize with respect to ϵ and obtain $M \le AC_{m,p}M^{1-1/m}$, for some constant $C_{m,p}$, which implies that $M \le C_{m,p}^m A^m$. If p is not bounded, we apply the procedure to $p * \psi_\delta$, where ψ_δ is an approximation of the identity, and let δ tend to zero at the end. □

7 Malliavin Differentiability of Diffusion Processes. Proof of Hörmander's theorem

Suppose that $B = (B_t)_{t\ge0}$, with $B_t = (B_t^1, \ldots, B_t^d)$, is a d-dimensional Brownian motion. Consider the m-dimensional stochastic differential equation

$$dX_t = \sum_{j=1}^d \sigma_j(X_t)dB_t^j + b(X_t)dt, \tag{42}$$

with initial condition $X_0 = x_0 \in \mathbb{R}^m$, where the coefficients $\sigma_j, b\colon \mathbb{R}^m \to \mathbb{R}^m$, $1 \le j \le d$ are measurable functions.

By definition, a solution to equation (42) is an adapted process $X = (X_t)_{t\ge0}$ such that, for any $T > 0$ and $p \ge 2$,

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t|^p\right) < \infty$$

and X satisfies the integral equation

$$X_t = x_0 + \sum_{j=1}^d \int_0^t \sigma_j(X_s)dB_s^j + \int_0^t b(X_s)ds. \tag{43}$$

The following result is well known.

Theorem 7.1. *Suppose that the coefficients $\sigma_j, b\colon \mathbb{R}^m \to \mathbb{R}^m$, $1 \le j \le d$, satisfy the Lipschitz condition: for all $x, y \in \mathbb{R}^m$,*

$$\max_j (|\sigma_j(x) - \sigma_j(y)|, |b(x) - b(y)|) \le K|x-y|. \tag{44}$$

Then there exists a unique solution X to Equation (43).

When the coefficients in equation (42) are continuously differentiable, the components of the solution are differentiable in the Malliavin calculus sense.

Proposition 7.1. *Suppose that the coefficients σ_j, b are in $C^1(\mathbb{R}^m;\mathbb{R}^m)$ and have bounded partial derivatives. Then, for all $t \geq 0$ and $i = 1,\dots,m$, $X_t^i \in \mathbb{D}^{1,\infty}$, and for $r \leq t$ and $j = 1,\dots,d$,*

$$D_r^j X_t = \sigma_j(X_r) + \sum_{k=1}^m \sum_{\ell=1}^d \int_r^t \partial_k \sigma_\ell(X_s) D_r^j X_s^k dB_s^\ell + \sum_{k=1}^m \int_r^t \partial_k b(X_s) D_r^j X_s^k ds. \tag{45}$$

Proof. To simplify, we assume that $b = 0$. Consider the Picard approximations given by $X_t^{(0)} = x_0$ and

$$X_t^{(n+1)} = x_0 + \sum_{j=1}^d \int_0^t \sigma_j(X_s^{(n)}) dB_s^j,$$

if $n \geq 0$. We will prove the following claim by induction on n:

Claim: $X_t^{(n),i} \in \mathbb{D}^{1,\infty}$ for all $i = 1,\dots,m, t \geq 0$. Moreover, for all $p > 1$ and $t \geq 0$,

$$\psi_n(t) := \sup_{0\leq r\leq t} \mathbb{E}\left(\sup_{s\in[r,t]} |D_r X_s^{(n)}|^p\right) < \infty \tag{46}$$

and, for all $T > 0$ and $t \in [0,T]$,

$$\psi_{n+1}(t) \leq c_1 + c_2 \int_0^t \psi_n(s) ds, \tag{47}$$

for some constants c_1, c_2 depending on T.

Clearly, the claim holds for $n = 0$. Suppose that it is true for n. Applying property (11) of the divergence operator and the chain rule (Proposition 2.4), for any $r \leq t$, $i = 1,\dots,m$, and $\ell = 1,\dots,d$, we get

$$\begin{aligned} D_r^\ell X_t^{(n+1),i} &= D_r^\ell \left(\sum_{j=1}^m \int_0^t \sigma_j^i(X_s^{(n)}) dB_s^j \right) \\ &= \sum_{j=1}^m \left(\delta_{\ell,j} \sigma_\ell^i(X_r^{(n)}) + \int_r^t D_r^\ell \left(\sigma_j^i(X_s^{(n)}) \right) dB_s^j \right) \\ &= \sum_{j=1}^m \left(\delta_{\ell,j} \sigma_\ell^i(X_r^{(n)}) + \sum_{k=1}^m \int_r^t \partial_k \sigma_j(X_s^{(n)}) D_r^\ell X_s^{(n),k} dB_s^j \right). \end{aligned}$$

From these equalities and condition (46) we see that $X_t^{(n+1),i} \in \mathbb{D}^{1,\infty}$ and we obtain, using the Burkholder–David–Gundy inequality and Hölder's inequality,

$$\mathbb{E}\left(\sup_{r\leq s\leq t} |D_r X_s^{(n+1)}|^p\right) \leq c_p \left(\gamma_p + T^{(p-1)/2} K^p \int_r^t \mathbb{E}\left(|D_r^j X_s^{(n)}|^p\right) ds\right), \tag{48}$$

where

$$\gamma_p = \sup_{n,j} \mathbb{E}\left(\sup_{0\leq t\leq T} |\sigma_j(X_t^{(n)})|^p\right) < \infty.$$

So (46) and (47) hold for $n+1$ and the claim is proved.

We know that

$$\mathbb{E}\left(\sup_{s\le T}|X_s^{(n)}-X_s|^p\right)\longrightarrow 0$$

as n tends to infinity. By Gronwall's lemma applied to (47) we deduce that the derivatives of the sequence $X_t^{(n),i}$ are bounded in $L^p(\Omega;H)$ uniformly in n for all $p\ge 2$. This implies that the random variables X_t^i belong to $\mathbb{D}^{1,\infty}$. Finally, applying the operator D to equation (43) we deduce the linear stochastic differential equation (45) for the derivative of X_t^i.

This completes the proof of the proposition. □

Example 7.2. *Consider the diffusion process in* $\mathbb{R}$

$$dX_t=\sigma(X_t)dB_t+b(X_t)dt,\qquad X_0=x_0,$$

where σ and b are globally Lipschitz functions in $C^1(\mathbb{R})$. Then, for all $t>0$, X_t belongs to $\mathbb{D}^{1,\infty}$ and the Malliavin derivative $(D_rX_t)_{r\le t}$ satisfies the following linear equation:

$$D_rX_t=\sigma(X_r)+\int_r^t\sigma'(X_s)D_r(X_s)dB_s+\int_r^t b'(X_s)D_r(X_s)ds.$$

Therefore, by Itô's formula,

$$D_rX_t=\sigma(X_t)\exp\left(\int_r^t\sigma'(X_s)dB_s+\int_r^t(b(X_s)-\tfrac{1}{2}(\sigma')^2(X_s))ds\right).$$

Consider the $m\times m$ matrix-valued process defined by

$$Y_t=I_m+\sum_{l=1}^d\int_0^t\partial\sigma_\ell(X_s)Y_sdB_s^\ell+\int_0^t\partial b(X_s)Y_sds,$$

where I_m denotes the identity matrix of order m and $\partial\sigma_\ell$ denotes the $m\times m$ Jacobian matrix of the function σ_ℓ; that is,

$$(\partial\sigma_\ell)_j^i=\partial_j\sigma_\ell^i.$$

In the same way, ∂b denotes the $m\times m$ Jacobian matrix of b. If the coefficients of equation (43) are of class $C^{1+\alpha}$, $\alpha>0$, then there is a version of the solution $X_t(x_0)$ to this equation that is continuously differentiable in x_0, and for which Y_t is the Jacobian matrix $\partial X_t/\partial x_0$:

$$Y_t=\frac{\partial X_t}{\partial x_0}.$$

Proposition 7.2. *For any $t\in[0,T]$ the matrix Y_t is invertible. Its inverse Z_t satisfies*

$$\begin{aligned}Z_t &= I_m-\sum_{\ell=1}^d\int_0^t Z_s\partial\sigma_\ell(X_s)dB_s^\ell\\ &\quad -\int_0^t Z_s\left(\partial b(X_s)-\sum_{\ell=1}^d\partial\sigma_\ell(X_s)\partial\sigma_\ell(X_s)\right)ds.\end{aligned}$$

Proof. By means of Itô's formula, one can check that $Z_tY_t = Y_tZ_t = I_m$, which implies that $Z_t = Y_t^{-1}$. In fact,

$$
\begin{aligned}
Z_tY_t &= I_m + \sum_{\ell=1}^{d} \int_0^t Z_s \; \partial\sigma_\ell(X_s)Y_s dB_s^\ell + \int_0^t Z_s \partial b(X_s)Y_s ds \\
&\quad - \sum_{\ell=1}^{d} \int_0^t Z_s \partial\sigma_l(X_s)Y_s dB_s^\ell \\
&\quad - \int_0^t Z_s \left(\partial b(X_s) - \sum_{\ell=1}^{d} \partial\sigma_\ell(X_s)\partial\sigma_\ell(X_s) \right) Y_s ds \\
&\quad - \int_0^t Z_s \left(\sum_{\ell=1}^{d} \partial\sigma_\ell(X_s)\partial\sigma_\ell(X_s) \right) Y_s ds = I_m.
\end{aligned}
$$

Similarly, we can show that $Y_tZ_t = I_m$. □

Lemma 7.3. *The $m \times d$ matrix $(D_rX_t)^i_j = D^j_rX^i_t$ can be expressed as*

$$D_rX_t = Y_tY_r^{-1}\sigma(X_r), \tag{49}$$

where σ denotes the $m \times d$ matrix with columns $\sigma_1, \dots, \sigma_d$.

Proof. It suffices to check that the process $\Phi_{t,r} := Y_tY_r^{-1}\sigma(X_r)$, $t \geq r$ satisfies

$$\Phi_{t,r} = \sigma(X_r) + \sum_{\ell=1}^{d} \int_r^t \partial\sigma_\ell(X_s)\Phi_{s,r} dB_s^\ell + \int_r^t \partial b(X_s)\Phi_{s,r} ds.$$

In fact,

$$
\begin{aligned}
&\sigma(X_r) + \sum_{\ell=1}^{d} \int_r^t \partial\sigma_\ell(X_s)(Y_sY_r^{-1}\sigma(X_r)) dB_s^\ell \\
&+ \int_r^t \partial b(X_s)(Y_sY_r^{-1}\sigma(X_r)) ds \\
&\quad = \sigma(X_r) + (Y_t - Y_r)Y_r^{-1}\sigma(X_r) = Y_tY_r^{-1}\sigma(X_r).
\end{aligned}
$$

This completes the proof. □

Consider the Malliavin matrix of X_t, denoted by $\gamma_{X_t} := Q_t$ and given by

$$Q_t^{i,j} = \sum_{\ell=1}^{d} \int_0^t D_s^\ell X_t^i D_s^\ell X_t^j ds.$$

That is, $Q_t = \int_0^t (D_sX_t)(D_sX_t)^T ds$. Equation (49) leads to

$$Q_t = Y_tC_tY_t^T, \tag{50}$$

where

$$C_t = \int_0^t Y_s^{-1}\sigma\sigma^T(X_s)(Y_s^{-1})^T ds.$$

Taking into account that Y_t is invertible, the nondegeneracy of the matrix Q_t will depend only on the nondegeneracy of the matrix C_t, which is called the *reduced Malliavin matrix*.

7.1 Absolute continuity under ellipticity conditions

Consider the stopping time defined by

$$S = \inf\{t > 0 : \det \sigma\sigma^T(X_t) \neq 0\}.$$

Theorem 7.4. *Let $(X_t)_{t\geq 0}$ be a diffusion process with $C^{1+\alpha}$ and Lipschitz coefficients. Then, for any $t > 0$, the law of X_t conditioned by $\{t > S\}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.*

Proof. It suffices to show that $\det C_t > 0$ a.s. on the set $\{S < t\}$. Suppose that $t > S$. For any $u \in \mathbb{R}^m$ with $|u| = 1$ we can write

$$\begin{aligned} u^T C_t u &= \int_0^t u^T Y_s^{-1} \sigma\sigma^T(X_s)(Y_s^{-1})^T u ds \\ &\geq \int_0^t \inf_{|v|=1} \left(v^T \sigma\sigma^T(X_s)v\right) |(Y_s^{-1})^T u|^2 ds. \end{aligned}$$

Notice that $\inf_{|v|=1} \left(v^T \sigma\sigma^T(X_s)v\right)$ is the smallest eigenvalue of $\sigma\sigma^T(X_s)$, which is strictly positive in an open interval contained in $[0, t]$ by the definition of the stopping time S and because $t > S$.

Furthermore, $|(Y_s^{-1})^T u| \geq |u|\, |Y_s|^{-1}$. Therefore we obtain

$$u^T C_t u \geq k|u|^2,$$

for some positive random variable $k > 0$, which implies that the matrix C_t is invertible. This completes the proof. □

Example 7.5. *Assume that $\sigma(x_0) \neq 0$ in Example 7.2. Then, for any $t > 0$, the law of X_t is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}$.*

7.2 Regularity of the density under Hörmander's conditions

We need the following regularity result, whose proof is similar to that of Proposition 7.1 and is thus omitted.

Proposition 7.3. *Suppose that the coefficients σ_j, $1 \leq j \leq m$, and b of equation (42) are infinitely differentiable with bounded derivatives of all orders. Then, for all $t \geq 0$ and $i = 1, \ldots, m$, X_t^i belong to $\mathbb{D}^\infty$.*

Consider the following vector fields on $\mathbb{R}^m$:

$$\begin{aligned} \sigma_j &= \sum_{i=1}^m \sigma_j^i(x) \frac{\partial}{\partial x_i}, \quad j = 1, \ldots, d, \\ b &= \sum_{i=1}^m b^i(x) \frac{\partial}{\partial x_i}. \end{aligned}$$

The Lie bracket between the vector fields σ_j and σ_k is defined by

$$[\sigma_j, \sigma_k] = \sigma_j \sigma_k - \sigma_k \sigma_j = \sigma_j^\nabla \sigma_k - \sigma_k^\nabla \sigma_j,$$

where

$$\sigma_j^\nabla \sigma_k = \sum_{i,\ell=1}^m \sigma_j^\ell \partial_\ell \sigma_k^i \frac{\partial}{\partial x_i}.$$

Set

$$\sigma_0 = b - \frac{1}{2}\sum_{\ell=1}^d \sigma_\ell^\nabla \sigma_\ell.$$

The vector field σ_0 appears when we write the stochastic differential equation (43) in terms of the Stratonovich integral (see Section 2.7) instead of Itô's integral:

$$X_t = x_0 + \sum_{j=1}^d \int_0^t \sigma_j(X_s) \circ dB_s^j + \int_0^t \sigma_0(X_s)ds.$$

Let us introduce the nondegeneracy condition required for the smoothness of the density.
(HC) *Hörmander's condition:* The vector space spanned by the vector fields

$$\sigma_1, \dots, \sigma_d,\ [\sigma_i, \sigma_j], 0 \le i \le d, 1 \le j \le d,\ [\sigma_i, [\sigma_j, \sigma_k]], 0 \le i, j, k \le d, \dots$$

at the point x_0 is $\mathbb{R}^m$.

For instance, if $m = d = 1$, $\sigma_1^1(x) = a(x)$ and $\sigma_0^1(x) = a_0(x)$; then Hörmander's condition means that $a(x_0) \neq 0$ or $a^n(x_0)a_0(x_0) \neq 0$ for some $n \ge 1$.

Theorem 7.6. *Assume that Hörmander's condition holds. Then, for any $t > 0$, the random vector X_t has an infinitely differentiable density.*

This result can be considered as a probabilistic version of Hörmander's theorem on the hypoellipticity of second-order differential operators. In fact, the density p_t of X_t satisfies the Fokker–Planck equation

$$\Big(-\frac{\partial}{\partial t} + L^*\Big)p_t = 0,$$

where

$$L = \frac{1}{2}\sum_{i,j=1}^m (\sigma\sigma^T)^{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b^i \frac{\partial}{\partial x_i}.$$

Then, $p_t \in C^\infty(\mathbb{R}^m)$ means that $\partial/\partial t - L^*$ is hypoelliptic (Hörmander's theorem).

For the proof of Theorem 7.6 we need several technical lemmas.

Lemma 7.7. *Let C be an $m \times m$ symmetric nonnegative definite random matrix. Assume that the entries C^{ij} have moments of all orders and that for any $p \ge 2$ there exists $\epsilon_0(p)$ such that, for all $\epsilon \le \epsilon_0(p)$,*

$$\sup_{|v|=1} P\left(v^T C v \le \epsilon\right) \le \epsilon^p.$$

Then $\mathbb{E}((\det C)^{-p}) < \infty$ for all $p \ge 2$.

Proof. Let $\lambda = \inf_{|v|=1} v^T C v$ be the smallest eigenvalue of C. We know that $\lambda^m \le \det C$. Thus, it suffices to show that $\mathbb{E}(\lambda^{-p}) < \infty$ for all $p \ge 2$. Set $|C| = \left(\sum_{i,j=1}^m (C^{ij})^2\right)^{\frac{1}{2}}$. Fix

$\epsilon > 0$, and let $v_1, \ldots, v_N$ be a finite set of unit vectors such that the balls with their center in these points and radius $\epsilon^2/2$ cover the unit sphere S^{m-1}. Then, we have

$$\begin{aligned} P(\lambda < \epsilon) &= P\left(\inf_{|v|=1} v^T C v < \epsilon\right) \\ &\leq P\left(\inf_{|v|=1} v^T C v < \epsilon, |C| \leq \frac{1}{\epsilon}\right) + P\left(|C| > \frac{1}{\epsilon}\right). \end{aligned} \tag{51}$$

Assume that $|C| \leq 1/\epsilon$ and $v_k^T C v_k \geq 2\epsilon$ for any $k = 1, \ldots, N$. For any unit vector v, there exists a v_k such that $|v - v_k| \leq \epsilon^2/2$ and we can deduce the following inequalities:

$$\begin{aligned} v^T C v &\geq v_k^T C v_k - |v^T C v - v_k^T C v_k| \\ &\geq 2\epsilon - (|v^T C v - v^T C v_k| + |v^T C v_k - v_k^T C v_k|) \\ &\geq 2\epsilon - 2|C|\,|v - v_k| \geq \epsilon. \end{aligned}$$

As a consequence, (51) implies that

$$P(\lambda < \epsilon) \leq P\left(\bigcup_{k=1}^N \{v_k^T C v_k < 2\epsilon\}\right) + P\left(|C| > \frac{1}{\epsilon}\right) \leq N(2\epsilon)^{p+2m} + \epsilon^p \mathbb{E}(|C|^p)$$

if $\epsilon \leq \frac{1}{2}\epsilon_0(p + 2m)$. The number N depends on ϵ but is bounded by a constant times ϵ^{-2m}. Therefore, we obtain $P(\lambda < \epsilon) \leq C\epsilon^p$ for all $\epsilon \leq \epsilon_1(p)$ and for all $p \geq 2$. This implies that λ^{-1} has moments of all orders, which completes the proof of the lemma. □

Lemma 7.8. *Let $(Z_t)_{t\geq 0}$ be a real-valued, adapted, continuous process such that $Z_0 = z_0 \neq 0$. Suppose that there exist $\alpha > 0$ and $t_0 > 0$ such that, for all $p \geq 1$ and $t \in [0, t_0]$,*

$$\mathbb{E}\left(\sup_{0\leq s\leq t} |Z_s - z_0|^p\right) \leq C_p t^{p\alpha}.$$

Then, for all $p \geq 1$ and $t \geq 0$,

$$\mathbb{E}\left(\left(\int_0^t |Z_s| ds\right)^{-p}\right) < \infty.$$

Proof. We can assume that $t \in [0, t_0]$. For any $0 < \epsilon < t|z_0|/2$, we have

$$\begin{aligned} P\left(\int_0^t |Z_s| ds < \epsilon\right) &\leq P\left(\int_0^{2\epsilon/|z_0|} |Z_s| ds < \epsilon\right) \\ &\leq P\left(\sup_{0\leq s\leq 2\epsilon/|z_0|} |Z_s - z_0| > \frac{|z_0|}{2}\right) \\ &\leq \frac{2^p C_p}{|z_0|^p}\left(\frac{2\epsilon}{|z_0|}\right)^{p\alpha}, \end{aligned}$$

which implies the desired result. □

The next lemma was proved by Norris , following the ideas of Stroock, and is the basic ingredient in the proof of Theorem 7.6.

Lemma 7.9 (Norris's lemma). *Consider a continuous semimartingale of the form*

$$Y_t = y + \int_0^t a_s ds + \sum_{i=1}^d \int_0^t u_s^i dB_s^i,$$

where

$$a(t) = \alpha + \int_0^t \beta_s ds + \sum_{i=1}^d \int_0^t \gamma_s^i dB_s^i$$

and $c = \mathbb{E}\left(\sup_{0\le t\le T}(|\beta_t| + |\gamma_t| + |a_t| + |u_t|)^p\right) < \infty$ *for some* $p \ge 2$.

Fix $q > 8$. *Then, for all* $r < (q-8)/27$ *there exists an* ϵ_0 *such that, for all* $\epsilon \le \epsilon_0$, *we have*

$$P\left(\int_0^T Y_t^2 dt < \epsilon^q, \int_0^T (|a_t|^2 + |u_t|^2)dt \ge \epsilon\right) \le c_1 \epsilon^{rp}.$$

Proof of Theorem 7.6. The proof will be carried out in several steps:

Step 1 We need to show that, for all $t > 0$ and all $p \ge 2$, $\mathbb{E}((\det Q_t)^{-p}) < \infty$, where Q_t is the Malliavin matrix of X_t. Taking into account that

$$\mathbb{E}\left(|\det Y_t^{-1}|^p + |\det Y_t|^p\right) < \infty,$$

it suffices to show that $\mathbb{E}((\det C_t)^{-p}) < \infty$ for all $p \ge 2$.

Step 2 Fix $t > 0$. Using Lemma 7.7, the problem reduces to showing that, for all $p \ge 2$, we have

$$\sup_{|v|=1} P\left(v^T C_t v \le \epsilon\right) \le \epsilon^p,$$

for any $\epsilon \le \epsilon_0(p)$, where the quadratic form associated with the matrix C_t is given by

$$v^T C_t v = \sum_{j=1}^d \int_0^t \langle v, Y_s^{-1}\sigma_j(X_s)\rangle^2 ds. \tag{52}$$

Step 3 Fix a smooth function V and use Itô's formula to compute the differential of $Y_t^{-1}V(X_t)$:

$$\begin{aligned} d\left(Y_t^{-1}V(X_t)\right) &= Y_t^{-1}\sum_{k=1}^d [\sigma_k, V](X_t)dB_t^k \\ &\quad + Y_t^{-1}\left([\sigma_0, V] + \tfrac{1}{2}\sum_{k=1}^d [\sigma_k, [\sigma_k, V]]\right)(X_t)dt. \end{aligned} \tag{53}$$

Step 4 We introduce the following sets of vector fields:

$$\begin{aligned} \Sigma_0 &= \{\sigma_1, \dots, \sigma_d\}, \\ \Sigma_n &= \{[\sigma_k, V], k = 0, \dots, d, V \in \Sigma_{n-1}\} \quad \text{if } n \ge 1, \\ \Sigma &= \cup_{n=0}^\infty \Sigma_n \end{aligned}$$

and

$$
\begin{aligned}
\Sigma_0' &= \Sigma_0, \\
\Sigma_n' &= \Big\{ [\sigma_k, V], k = 1, \dots, d, V \in \Sigma_{n-1}'; \\
&\qquad [\sigma_0, V] + \tfrac{1}{2} \sum_{j=1}^{d} [\sigma_j, [\sigma_j, V]], V \in \Sigma_{n-1}' \Big\} \quad \text{if } n \geq 1, \\
\Sigma' &= \cup_{n=0}^{\infty} \Sigma_n'.
\end{aligned}
$$

We denote by $\Sigma_n(x)$ (resp. $\Sigma_n'(x)$) the subset of $\mathbb{R}^m$ obtained by freezing the variable x in the vector fields of Σ_n (resp. Σ_n'). Clearly, the vector spaces spanned by $\Sigma(x_0)$ or by $\Sigma'(x_0)$ coincide and, under Hörmander's condition, this vector space is $\mathbb{R}^m$. Therefore, there exists an integer $j_0 \geq 0$ such that the linear span of the set of vector fields $\bigcup_{j=0}^{j_0} \Sigma_j'(x)$ at point x_0 has dimension m.

As a consequence there exist constants $R > 0$ and $c > 0$ such that

$$
\sum_{j=0}^{j_0} \sum_{V \in \Sigma_j'} \langle v, V(y) \rangle^2 \geq c, \tag{54}
$$

for all v and y with $|v| = 1$ and $|y - x_0| < R$.

Step 5 For any $j = 0, 1, \dots, j_0$ we put $m(j) = 2^{-4j}$ and define the set

$$
E_j = \left\{ \sum_{V \in \Sigma_j'} \int_0^t \langle v, Y_s^{-1} V(X_s) \rangle^2 ds \leq \epsilon^{m(j)} \right\}.
$$

Notice that $\{v^T C_t v \leq \epsilon\} = E_0$ because $m(0) = 1$. Consider the decomposition

$$
E_0 \subset (E_0 \cap E_1^c) \cup (E_1 \cap E_2^c) \cup \cdots \cup (E_{j_0 - 1} \cap E_{j_0}^c) \cup F,
$$

where $F = E_0 \cap E_1 \cap \cdots \cap E_{j_0}$. Then, for any unit vector v, we have

$$
P(v^T C_t v \leq \epsilon) = P(E_0) \leq P(F) + \sum_{j=0}^{j_0 - 1} P(E_j \cap E_{j+1}^c).
$$

We will now estimate each term in this sum.

Step 6 Let us first estimate $P(F)$. By the definition of F we obtain

$$
P(F) \leq P\left(\sum_{j=0}^{j_0} \sum_{V \in \Sigma_j'} \int_0^t \langle v, Y_s^{-1} V(X_s) \rangle^2 ds \leq (j_0 + 1) \epsilon^{m(j_0)} \right).
$$

Then, taking into account (54), we can apply Lemma 7.8 to the process

$$
Z_s = \inf_{|v|=1} \sum_{j=0}^{j_0} \sum_{V \in \Sigma_j'} \langle v, Y_s^{-1} V(X_s) \rangle^2,
$$

and we obtain

$$\mathbb{E}\left(\left|\inf_{|v|=1}\sum_{j=0}^{j_0}\sum_{V\in\Sigma'_j}\int_0^t \langle v, Y_s^{-1}V(X_s)\rangle^2 ds\right|^{-p}\right) < \infty.$$

Therefore, for any $p \geq 1$, there exists ϵ_0 such that

$$P(F) \leq \epsilon^p$$

for any $\epsilon < \epsilon_0$.

Step 7 For any $j = 0, \ldots, j_0$, the probability of the event $E_j \cap E_{j+1}^c$ is bounded by the sum with respect to $V \in \Sigma'_j$ of the probability that the two following events happen:

$$\int_0^t \langle v, Y_s^{-1}V(X_s)\rangle^2 ds \leq \epsilon^{m(j)}$$

and

$$\sum_{k=1}^d \int_0^t \langle v, Y_s^{-1}[\sigma_k, V](X_s)\rangle^2 ds$$
$$+\int_0^t \left\langle v, Y_s^{-1}\left([\sigma_0, V] + \tfrac{1}{2}\sum_{j=1}^d [\sigma_j, [\sigma_j, V]]\right)(X_s)\right\rangle^2 ds > \frac{\epsilon^{m(j+1)}}{n(j)},$$

where $n(j)$ denotes the cardinality of the set Σ'_j.

Consider the continuous semimartingale $(\langle v, Y_s^{-1}V(X_s)\rangle)_{s\geq 0}$. From (53) we see that the quadratic variation of this semimartingale is equal to

$$\sum_{k=1}^d \int_0^s \langle v, Y_r^{-1}[\sigma_k, V](X_r)\rangle^2 dr,$$

and the bounded variation component is

$$\int_0^s \left\langle v, Y_r^{-1}\left([\sigma_0, V] + \tfrac{1}{2}\sum_{j=1}^d [\sigma_j, [\sigma_j, V]]\right)(X_r)\right\rangle dr.$$

Taking into account that $8m(j+1) < m(j)$, from Norris's lemma (Lemma 7.9) applied to the semimartingale $Y_s = v^T Y_s^{-1} V(X_s)$, we get that, for any $p \geq 1$, there exists an $\epsilon_0 > 0$ such that

$$P(E_j \cap E_{j+1}^c) \leq \epsilon^p,$$

for all $\epsilon \leq \epsilon_0$. The proof of the theorem is now complete. □

8 Stein's method for normal approximation

The following lemma is a characterization of the standard normal distribution on the real line.

Lemma 8.1 (Stein's lemma)**.** *A random variable X such that $\mathbb{E}(|X|) < \infty$ has the standard normal distribution $N(0,1)$ if and only if, for any function $f \in C_b^1(\mathbb{R})$, we have*

$$\mathbb{E}(f'(X) - f(X)X) = 0. \tag{55}$$

Proof. Suppose first that X has the standard normal distribution $N(0,1)$. Then, equality (55) follows integrating by parts and using that the density $p(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$ satisfies the differential equation

$$p'(x) = -xp(x).$$

Conversely, let $\varphi(\lambda) = \mathbb{E}(e^{i\lambda X})$, $\lambda \in \mathbb{R}$, be the characteristic function of X. Because X is integrable, we know that φ is differentiable and $\varphi'(\lambda) = i\mathbb{E}(Xe^{i\lambda X})$. By our assumption, this is equal to $-\lambda\varphi(\lambda)$. Therefore, $\varphi(\lambda) = \exp(-\lambda^2/2)$, which concludes the proof. □

If the expectation $\mathbb{E}(f'(X) - f(X)X)$ is small for functions f in some large set, we might conclude that the distribution of X is close to the normal distribution. This is the main idea of Stein's method for normal approximations and the goal is to quantify this assertion in a proper way. To do this, consider a random variable X with the $N(0,1)$ distribution and fix a measurable function $h\colon \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}(|h(X)|) < \infty$. Stein's equation associated with h is the linear differential equation

$$f_h'(x) - xf_h(x) = h(x) - \mathbb{E}(h(X)), \quad x \in \mathbb{R}. \tag{56}$$

Definition 8.2. *A solution to equation (56) is an absolutely continuous function f_h such that there exists a version of the derivative f_h' satisfying (56) for every $x \in \mathbb{R}$.*

The next result provides the existence of a unique solution to Stein's equation.

Proposition 8.1. *The function*

$$f_h(x) = e^{x^2/2} \int_{-\infty}^{x} (h(y) - \mathbb{E}(h(X)))e^{-y^2/2} dy \tag{57}$$

is the unique solution of Stein's equation (56) *satisfying*

$$\lim_{x\to\pm\infty} e^{-x^2/2} f_h(x) = 0. \tag{58}$$

Proof. Equation (56) can be written as

$$e^{x^2/2} \frac{d}{dx}\left(e^{-x^2/2} f_h(x)\right) = h(x) - \mathbb{E}(h(X)).$$

This implies that any solution to equation (56) is of the form

$$f_h(x) = ce^{x^2/2} + e^{x^2/2} \int_{-\infty}^{x} (h(y) - \mathbb{E}(h(X)))e^{-y^2/2} dy,$$

for some $c \in \mathbb{R}$. Taking into account that

$$\lim_{x\to\pm\infty} \int_{-\infty}^{x} (h(y) - \mathbb{E}(h(X)))e^{-y^2/2} dy = 0,$$

the asymptotic condition (58) is satisfied if and only if $c = 0$. □

Notice that, since $\int_{\mathbb{R}}(h(y)-\mathbb{E}(h(X)))e^{-y^2/2}dy=0$, we have

$$\int_{-\infty}^{x}(h(y)-\mathbb{E}(h(X)))e^{-y^2/2}dy=-\int_{x}^{\infty}(h(y)-\mathbb{E}(h(X)))e^{-y^2/2}dy. \tag{59}$$

Proposition 8.2. *Let $h\colon\mathbb{R}\to[0,1]$ be a measurable function. Then the solution to Stein's equation f_h given by* (57) *satisfies*

$$\|f_h\|_\infty\le\sqrt{\frac{\pi}{2}}\quad\text{and}\quad\|f_h'\|_\infty\le 2. \tag{60}$$

Proof. Taking into account that $|h(x)-\mathbb{E}(h(X))|\le 1$, where X has law $N(0,1)$, we obtain

$$|f_h(x)|\le e^{x^2/2}\int_{|x|}^{\infty}e^{-y^2/2}dy=\sqrt{\frac{\pi}{2}},$$

because the function $x\to e^{x^2/2}\int_{|x|}^{\infty}e^{-y^2/2}dy$ attains its maximum at $x=0$.

To prove the second estimate, observe that, in view of (59), we can write

$$\begin{aligned}f_h'(x)&=h(x)-\mathbb{E}(h(X))+xe^{x^2/2}\int_{-\infty}^{x}(h(y)-\mathbb{E}(h(X)))e^{-y^2/2}dy\\&=h(x)-\mathbb{E}(h(X))-xe^{x^2/2}\int_{x}^{\infty}(h(y)-\mathbb{E}(h(X)))e^{-y^2/2}dy,\end{aligned}$$

for every $x\in\mathbb{R}$. Therefore

$$|f_h'(x)|\le 1+|x|e^{x^2/2}\int_{|x|}^{\infty}e^{-y^2/2}dy=2.$$

This completes the proof. □

8.1 Total variation and convergence in law

Let F_n be a sequence of random variables defined in a probability space $(\Omega,\mathcal{F},P)$.

Definition 8.3. *We say that $F_n\xrightarrow{\mathcal{L}}F$ if $E[g(F_n)]\to E[g(F)]$ for any $g:\mathbb{R}\to\mathbb{R}$ continuous and bounded.*

We know that $F_n\xrightarrow{\mathcal{L}}F$ if and only if $P(F_n\le z)\to P(F\le z)$ for any point $z\in\mathbb{R}$ of continuity of the distribution function of F.

The total variation distance between two probabilities ν_1 and ν_2 on $\mathbb{R}$ is defined as

$$\boxed{d_{TV}(\nu_1,\nu_2)=\sup_{B\in\mathcal{B}(\mathbb{R})}|\nu_1(B)-\nu_2(B)|}$$

Then, $d_{TV}(P\circ F_n^{-1},P\circ F^{-1})\to 0$ is strictly stronger that the convergence in law $F_n\xrightarrow{\mathcal{L}}F$.

Using the Stein's method, we can prove the following result.

Proposition 8.3. *Let ν be a probability on $\mathbb{R}$. Then,*

$$d_{TV}(\nu,\gamma) \le \sup_{\phi\in\mathscr{F}_{TV}} \left| \int_{\mathbb{R}} [\phi'(x) - x\phi(x)]\nu(dx) \right|,$$

where

$$\mathscr{F}_{TV} = \{\phi \in C^1(\mathbb{R}) : \|\phi\|_\infty \le \sqrt{\frac{\pi}{2}}, \|\phi'\|_\infty \le 2\}.$$

Proof. Let $h : \mathbb{R} \to [0,1]$ be a continuous function and let ϕ_h be the solution to the Stein's equation associated with h, that is,

$$h(x) - E[h(Z)] = \phi_h'(x) - x\phi_h(x).$$

Integrating with respect to ν yields

$$\begin{aligned} \left| \int_{\mathbb{R}} h d\nu - \int_{\mathbb{R}} h d\gamma \right| &= \left| \int_{\mathbb{R}} [\phi_h'(x) - x\phi_h(x)]\nu(dx) \right| \\ &\le \sup_{\phi\in C^1(\mathbb{R}): \|\phi\|_\infty \le \sqrt{\frac{\pi}{2}}, \|\phi'\|_\infty \le 2} \left| \int_{\mathbb{R}} [\phi'(x) - x\phi(x)]\nu(dx) \right|. \end{aligned}$$

This inequality holds for any $h : \mathbb{R} \to [0,1]$ measurable, because we can approximate h by continuous functions almost everywhere with respect to the measure $\nu + \gamma$. Taking $h = \mathbf{1}_B$, we obtain the result. □

9 Central limit theorems and Malliavin calculus

Let $(B_t)_{\in[0,T]}$ be a Brownian motion defined on the Wiener space $(\Omega, \mathcal{F}, P)$. The following results connects Stein's method with Malliavin calculus.

Theorem 9.1 (Nourdin-Peccati). *Suppose that $F \in \mathbb{D}^{1,2}$ satisfies $F = \delta(u)$, where u belongs to the domain in L^2 of the divergence operator δ. Then,*

$$\boxed{d_{TV}(P_F,\gamma) \le 2E[|1 - \langle DF, u\rangle_H|].}$$

Proof. It follows from

$$E[F\phi(F)] = E[\delta(u)\phi(F)] = E[\langle u, D[\phi(F)]\rangle_H] = E[\phi'(F)\langle u, DF\rangle_H].$$

Therefore,

$$\begin{aligned} |E[\phi'(F)] - E(F\phi(F))| &= |E[\phi'(F)[1 - \langle DF, u\rangle_H]| \\ &\le 2E[|1 - \langle DF, u\rangle_H|] \end{aligned}$$

for any $\phi \in \mathscr{F}_{TV}$. □

Suppose that $F = \int_0^T u_s dB_s$, where u is an adapted measurable process in $\mathbb{D}^{1,2}(H)$. Then,

$$D_t F = u_t + \int_t^T D_t u_s dB_s,$$

and

$$\langle u, DF\rangle_H = \|u\|_H^2 + \int_0^T \left(\int_t^T D_t u_s dB_s\right) u_t dt.$$

As a consequence,

$$\begin{aligned} d_{TV}(P_F,\gamma) &\le 2E\left(|1-\|u\|_H^2|\right) + 2E\left(\left|\int_0^T \left(\int_t^T D_t u_s dB_s\right) u_t dt\right|\right) \\ &\le 2E\left(|1-\|u\|_H^2|\right) + 2\left[E\int_0^T \left(\int_0^s u_t D_t u_s dt\right)^2 ds\right]^{\frac{1}{2}}. \end{aligned}$$

Proposition 9.1. *A sequence $F_n = \int_0^T u_s^{(n)} dB_s$, where $u^{(n)}$ is progressively measurable and $u^{(n)} \in \mathbb{D}^{1,2}(H)$, converges in total variation to the law $N(0,1)$ if:*

(i) $\|u^{(n)}\|_H^2 \to 1$ *in* $L^1(\Omega)$ *and*

(ii) $E\int_0^T \left(\int_0^s u_t^{(n)} D_t u_s^{(n)} dt\right)^2 ds \to 0.$

Example 1. The previous proposition can be applied to the following example.

$$u_t^{(n)} = \sqrt{2n}t^n \exp(B_t(1-t))\mathbf{1}_{[0,1]}(t).$$

We can take $u = -DL^{-1}F$, because

$$F = LL^{-1}F = -\delta DL^{-1}F,$$

and, we obtain

$$\boxed{d_{TV}(P_F,\gamma) \le 2E[|1-\langle DF, -DL^{-1}F\rangle_H|]}$$

If $E[F^2] = \sigma^2 > 0$ and we take $\gamma_\sigma = N(0,\sigma^2)$, we can derive the following inequality:

$$d_{TV}(F,\gamma_\sigma) \le \frac{2}{\sigma^2}E[|\sigma^2 - \langle DF, u\rangle_H|].$$

Proof.

$$\begin{aligned} d_{TV}(F,\gamma_\sigma) &= \sup_{B\in\mathcal{B}(\mathbb{R})} |P(F\in B) - \gamma_\sigma(B)| \\ &= \sup_{B\in\mathcal{B}(\mathbb{R})} |P(\sigma^{-1}F \in \sigma^{-1}B) - \gamma(\sigma^{-1}B)| \\ &= \sup_{B\in\mathcal{B}(\mathbb{R})} |P(\sigma^{-1}F\in B) - \gamma(B)| \\ &\le \frac{2}{\sigma^2}E[|\sigma^2 - \langle DF, u\rangle_H|]. \end{aligned}$$

□

9.1 Normal approximation on a fixed Wiener chaos

Recall that for any $F \in \mathbb{D}^{1,2}$ such that $E[F]=0$,

$$d_{TV}(P_F,\gamma_\sigma) \le \frac{2}{\sigma^2} E[|\sigma^2 - \langle DF, -DL^{-1}F\rangle_H|].$$

Proposition 9.2. *Suppose $F \in \mathcal{H}_q$ for some $q \ge 2$ and $E(F^2)=\sigma^2$. Then,*

$$\boxed{d_{TV}(P_F,\gamma_\sigma) \le \frac{2}{q\sigma^2}\sqrt{\operatorname{Var}\left(\|DF\|_H^2\right)}}$$

Proof. Using $L^{-1}F = -\frac{1}{q}F$ and $E[\|DF\|_H^2] = q\sigma^2$, we obtain

$$\begin{aligned} E[|\sigma^2 - \langle DF, -DL^{-1}F\rangle_H|] &= E\left[\left|\sigma^2 - \frac{1}{q}\|DF\|_H^2\right|\right] \\ &\le \frac{1}{q}\sqrt{\operatorname{Var}\left(\|DF\|_H^2\right)}. \end{aligned}$$

□

Proposition 9.3. *Suppose that $F = I_q(f) \in \mathcal{H}_q$, $q \ge 2$. Then,*

$$\operatorname{Var}\left(\|DF\|_H^2\right) \le \frac{(q-1)q}{3}(E(F^4) - 3\sigma^4) \le (q-1)\operatorname{Var}\left(\|DF\|_H^2\right).$$

Proof. This proposition is a consequence of the following two formulas:

$$\boxed{\operatorname{Var}\left(\|DF\|_H^2\right) = \sum_{r=1}^{q-1} r^2 (r!)^2 \binom{q}{r}^4 (2q-2r)! \|f \widetilde{\otimes}_r f\|^2_{H^{\otimes(2q-2r)}}} \tag{61}$$

Proof of (61): We have $D_tF = qI_{q-1}(f(\cdot,t))$, and using the product formula for multiple stochastic integrals we obtain

$$\begin{aligned} \|DF\|_H^2 &= q^2 \int_0^T I_{q-1}(f(\cdot,t))^2 dt \\ &= q^2 \sum_{r=0}^{q-1} r! \binom{q-1}{r}^2 I_{2q-2r-2}(f \widetilde{\otimes}_{r+1} f) \\ &= q^2 \sum_{r=1}^{q} (r-1)! \binom{q-1}{r-1}^2 I_{2q-2r}(f \widetilde{\otimes}_r f) \\ &= qq! \|f\|^2_{H^{\otimes q}} + q^2 \sum_{r=1}^{q-1} (r-1)! \binom{q-1}{r-1}^2 I_{2q-2r}(f \widetilde{\otimes}_r f). \end{aligned} \tag{62}$$

Then, (61) follows from the isometry property of multiple integrals.
The second formula is the following one:

$$\boxed{E[F^4] - 3\sigma^4 = \frac{3}{q}\sum_{r=1}^{q-1} r(r!)^2 \binom{q}{r}^4 (2q-2r)! \|f \widetilde{\otimes}_r f\|^2_{H^{\otimes(2q-2r)}}} \tag{63}$$

Proof on (63): Using that $-L^{-1}F = \frac{1}{q}F$ and $L = -\delta D$ we can write

$$\begin{aligned} E[F^4] &= E[F \times F^3] = E[(-\delta DL^{-1}F)F^3] = E[\langle -DL^{-1}F, D(F^3)\rangle_H] \\ &= \frac{1}{q}E[\langle DF, D(F^3)\rangle_H] = \frac{3}{q}E[F^2\|DF\|_H^2]. \end{aligned} \tag{64}$$

By the product formula of multiple integrals,

$$F^2 = I_q(f)^2 = q!\|f\|^2_{H^{\otimes q}} + \sum_{r=0}^{q-1} r!\binom{q}{r}^2 I_{2q-2r}(f\widetilde{\otimes}_r f). \tag{65}$$

Then (63) follows from (64), (65), (62) and the isometry property of multiple integrals. □

9.2 Fourth Moment theorem

Stein's method combined with Malliavin calculus leads to a simple proof of the Fourth Moment theorem:

Theorem 9.2. *Fix $q \geq 2$. Let $F_n = I_q(f_n) \in \mathcal{H}_q$, $n \geq 1$ be such that*

$$\lim_{n\to\infty} E(F_n^2) = \sigma^2.$$

The following conditions are equivalent:

(i) $F_n \overset{\mathcal{L}}{\to} N(0, \sigma^2)$, *as* $n \to \infty$.

(ii) $E(F_n^4) \to 3\sigma^4$, *as* $n \to \infty$.

(iii) $\|DF_n\|_H^2 \to q\sigma^2$ *in* $L^2(\Omega)$, *as* $n \to \infty$.

(iv) For all $1 \leq r \leq q-1$, $f_n \otimes_r f_n \to 0$, *as* $n \to \infty$.

This theorem constitutes a drastic simplification of the method of moments.

Proof. First notice that (i) implies (ii) because for any $p > 2$, the hypercontractivity property of the Ornstein-Uhlenbeck semigroup implies

$$\sup_n \|F_n\|_p \leq (p-1)^{\frac{q}{2}} \sup_n \|F_n\|_2 < \infty.$$

The equivalence of (ii) and (iii)) follows from the previous proposition, and these conditions imply (i), with convergence in total variation. The fact that (iv) implies (ii) and (iii) is a consequence of $\|f_n\widetilde{\otimes}_r f_n\| \leq \|f_n \otimes_r f_n\|$. Let us show that (ii) implies (iv). From (65) we get

$$\begin{aligned} E[F_n^4] &= \sum_{r=0}^{q} (r!)^2 \binom{q}{r}^4 (2q-2r)!\|f_n\tilde{\otimes}_r f_n\|^2_{H^{\otimes(2q-2r)}} \\ &= (2q)!\|f_n\tilde{\otimes} f_n\|^2_{H^{\otimes 2q}} + \sum_{r=1}^{q-1} (r!)^2 \binom{q}{r}^4 (2q-2r)!\|f_n\tilde{\otimes}_r f_n\|^2_{H^{\otimes(2q-2r)}} \\ &\quad +(q!)^2\|f_n\|_H^4. \end{aligned}$$

Then, we use the fact that $(2q)!\|f_n\widetilde{\otimes} f_n\|^2_{H^{\otimes 2q}}$ equals to $2(q!)^2\|f_n\|_H^4$ plus a linear combination of the terms $\|f_n \otimes_r f_n\|^2_{H^{\otimes(2q-2r)}}$, with $1 \leq r \leq q-1$, to conclude that

$$\|f_n \otimes_r f_n\|_{H^{\otimes(2q-2r)}} \to 0, \quad 1 \leq r \leq q-1.$$

□

9.3 Multivariate Gaussian approximation

The next result is as multivariate extension of the fourth moment theorem.

Theorem 9.3 (Peccati-Tudor '05). *Let $d \geq 2$ and $1 \leq q_1 < \cdots < q_d$. Consider random vectors*

$$F_n = (F_n^1, \dots, F_n^d) = (I_{q_1}(f_n^1), \dots, I_{q_d}(f_n^d)),$$

where $f_n^i \in L_s^2([0,T]^{q_i})$. Suppose that, for any $1 \leq i \leq d$,

$$\lim_{n \to \infty} E[(F_n^i)^2] = \sigma_i^2.$$

Then, the following two conditions are equivalent:

(i) $F_n \xrightarrow{\mathcal{L}} N_d(0, \Sigma)$, where Σ is a diagonal matrix such that $\Sigma_{ii} = \sigma_i^2$.

(ii) For every $i = 1, \dots, d$, $F_n^i \xrightarrow{\mathcal{L}} N(0, \sigma_i^2)$.

Note that the convergence of the marginal distributions implies the joint convergence to a random vector with independent components.

9.4 Chaotic Central Limit Theorem

Theorem 9.4. *Let $F_n = \sum_{q=1}^{\infty} I_q(f_{q,n})$, $n \geq 1$. Suppose that:*

(i) For all $q \geq 1$, $q! \|f_{q,n}\|^2 \to \sigma_q^2$ as $n \to \infty$.

(ii) For all $q \geq 2$ and $1 \leq r \leq q-1$, $f_{q,n} \otimes_r f_{q,n} \to 0$ as $n \to \infty$.

(iii) $q! \|f_{q,n}\|^2 \leq \delta_q$, where $\sum_q \delta_q < \infty$.

Then, as n tends to infinity

$$F_n \xrightarrow{\mathcal{L}} N(0, \sigma^2), \quad \text{where} \quad \sigma^2 = \sum_{q=1}^{\infty} \sigma_q^2.$$

Assuming (i), condition (ii) is equivalent to (ii)': $\lim_{n \to \infty} E(I_q(f_{q,n})^4) = 3\sigma_q^4$, $q \geq 2$. The theorem implies the convergence in law of the whole sequence $(I_q(f_{q,n}), q \geq 1)$ to an infinite dimensional Gaussian vector with independent components.

9.5 Breuer-Major theorem

A function $f \in L^2(\mathbb{R}, \gamma)$ has *Hermite rank* $d \geq 1$ if

$$f(x) = \sum_{q=d}^{\infty} a_q H_q(x), \quad a_d \neq 0.$$

For example, $f(x) = |x|^p - \int_{\mathbb{R}} |x|^p d\gamma(x)$ has Hermite rank 1 if $p > 0$ is not an even integer.

Let $X = \{X_k, k \in \mathbb{Z}\}$ be a centered stationary Gaussian sequence with unit variance. Set $\rho(v) = E[X_0 X_v]$ for $v \in \mathbb{Z}$.

Theorem 9.5 (Breuer-Major '83). *Let $f \in L^2(\mathbb{R},\gamma)$ with Hermite rank $d \geq 1$ and assume $\sum_{v\in\mathbb{Z}} |\rho(v)|^d < \infty$. Then,*

$$V_n := \frac{1}{\sqrt{n}} \sum_{k=1}^n f(X_k) \xrightarrow{\mathcal{L}} N(0,\sigma^2),$$

as $n \to \infty$, where $\sigma^2 = \sum_{q=d}^\infty q! a_q^2 \sum_{v\in\mathbb{Z}} \rho(v)^q$.

Proof. From the chaotic Central Limit Theorem, it suffices to consider the case $f = a_q H_q$, $q \geq d$. There exists a sequence $\{e_k, k \geq 1\}$ in $H = L^2([0,T])$ such that

$$\langle e_k, e_j \rangle_H = \rho(k-j).$$

The sequence $\{B(e_k)\}$ has the same law as $\{X_k\}$, and we may replace V_n by

$$G_n = \frac{a_q}{\sqrt{n}} \sum_{k=1}^n H_q(B(e_k)) = I_q(f_{q,n}),$$

where $f_{q,n} = \frac{a_q}{\sqrt{n}} \sum_{k=1}^n e_k^{\otimes q}$. We can write

$$q! \|f_{q,n}\|^2_{H^{\otimes q}} = \frac{q! a_q^2}{n} \sum_{k,j=1}^n \rho(k-j)^q = q! a_q^2 \sum_{v\in\mathbb{Z}} \rho(v)^q \left(1 - \frac{|v|}{n}\right) \mathbf{1}_{\{|v|<n\}},$$

and by the dominated convergence theorem

$$E[G_n^2] = q! \|f_{q,n}\|^2_{H^{\otimes q}} \to q! a_q^2 \sum_{v\in\mathbb{Z}} \rho(v)^q = \sigma^2.$$

Applying the Fourth Moment Theorem, It suffices to show that for $r = 1,\dots,q-1$,

$$f_{q,n} \otimes_r f_{q,n} = \frac{a_q^2}{n} \sum_{k,j=1}^n \rho(k-j)^r e_k^{\otimes(q-r)} \otimes e_j^{\otimes(q-r)} \to 0.$$

We have

$$\|f_{q,n} \otimes_r f_{q,n}\|^2_{H^{\otimes(2q-2r)}} = \frac{a_q^4}{n^2} \sum_{i,j,k,\ell=1}^n \rho(k-j)^r \rho(i-\ell)^r \rho(k-i)^{q-r} \rho(j-\ell)^{q-r}.$$

Using $|\rho(k-j)^r \rho(k-i)^{q-r}| \leq |\rho(k-j)|^q + |\rho(k-i)|^q$, we obtain

$$\begin{aligned} \|f_{q,n} \otimes_r f_{q,n}\|^2_{H^{\otimes(2q-2r)}} \quad &\leq \quad 2a_q^4 \sum_{k\in\mathbb{Z}} |\rho(k)|^q \left(n^{-1+\frac{r}{q}} \sum_{|i|\leq n} |\rho(i)|^r \right) \\ &\quad \times \left(n^{-1+\frac{q-r}{q}} \sum_{|j|\leq n} |\rho(j)|^{q-r} \right). \end{aligned}$$

Then, it suffices to show that for $r = 1,\dots,q-1$,

$$n^{-1+\frac{r}{q}} \sum_{|i|\leq n} |\rho(i)|^r \to 0.$$

This follows from Hölder's inequality. Indeed, for a fixed $\delta \in (0,1)$, we have the estimates

$$n^{-1+\frac{r}{q}} \sum_{|i| \le [n\delta]} |\rho(i)|^r \le n^{-1+\frac{r}{q}} (2[n\delta]+1)^{1-\frac{r}{q}} \left(\sum_{i\in\mathbb{Z}} |\rho(i)|^q \right)^{\frac{r}{q}} \le c\delta^{1-\frac{r}{q}},$$

and

$$n^{-1+\frac{r}{q}} \sum_{[n\delta]<|i|\le n} |\rho(i)|^r \le \left(\sum_{[n\delta]<|i|\le n} |\rho(i)|^q \right)^{\frac{r}{q}}.$$

The first term converges to zero as δ tends to zero and the second one converges to zero for fixed δ as $n \to \infty$. □

9.6 Fractional Brownian motion

The fractional Brownian motion (fBm) $B^H = (B_t^H)_{t\ge 0}$ is a zero mean Gaussian process with covariance

$$E(B_s^H B_t^H) = R_H(s,t) = \frac{1}{2}\left(s^{2H} + t^{2H} - |t-s|^{2H}\right).$$

$H \in (0,1)$ is called the Hurst parameter.

The covariance formula implies $E(B_t^H - B_s^H)^2 = |t-s|^{2H}$. As a consequence, for any $\gamma < H$, with probability one, the trajectories $t \to B_t^H(\omega)$ are Hölder continuous of order γ:

$$|B_t^H(\omega) - B_s^H(\omega)| \le G_{\gamma,T}(\omega)|t-s|^{\gamma}, \quad s,t \in [0,T].$$

For $H = \frac{1}{2}$, $B^{\frac{1}{2}}$ is a Brownian motion.

Properties of the fractional Brownian motion:

1) The fractional Brownian motion has the following self-similarity property. For all $a > 0$, the processes

$$\{a^{-H} B_{at}^H, t \ge 0\}$$

and

$$\{B_t^H, t \ge 0\}$$

have the same probability distribution (they are fractional Brownian motions with Hurst parameter H).

2) Unlike Brownian motion, the fractional Brownian motion has correlated increments. More precisely, For $H \ne \frac{1}{2}$, we can write

$$\begin{aligned} \rho(n) &= E(B_1^H(B_{n+1}^H - B_n^H)) \\ &= \frac{1}{2}\left((n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right) \\ &\sim H(2H-1)n^{2H-2}, \end{aligned}$$

as $n \to \infty$.

(ii) If $H > \frac{1}{2}$, then $\rho(n) > 0$ and $\sum_n \rho(n) = \infty$ (*long memory*).

(iii) If $H < \frac{1}{2}$, then $\rho(n) < 0$ (*intermittency*) and $\sum_n |\rho(n)| < \infty$.

3) The fractional Brownian motion has finite $\frac{1}{H}$-variation: Fix $T > 0$. Set $t_i = \frac{iT}{n}$ for $1 \le i \le n$ and define $\Delta B^H_{t_i} = B^H_{t_i} - B^H_{t_{i-1}}$. Then, as $n \to \infty$,

$$\sum_{i=1}^{n} |\Delta B^H_{t_i}|^{\frac{1}{H}} \overset{L^2(\Omega),a.s.}{\longrightarrow} c_H T,$$

where $c_H = E[|B^H_1|^{\frac{1}{H}}]$.

Proof. By the self-similarity, $\sum_{i=1}^n |\Delta B^H_{t_i}|^{\frac{1}{H}}$ has the same law as

$$\frac{T}{n} \sum_{i=1}^{n} |B^H_i - B^H_{i-1}|^{\frac{1}{H}}.$$

The sequence $\{B^H_i - B^H_{i-1}, i \ge 1\}$ is stationary and ergodic. Therefore, the Ergodic Theorem implies the desired convergence. $\square$

9.6.1 Fractional noise

Let $X_k = B^H_k - B^H_{k-1}$. The sequence $\{X_k, k \ge 1\}$ is Gaussian, stationary and centered with covariance

$$\rho(n) = \frac{1}{2}\left(|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}\right).$$

We have $\rho(n) \sim H(2H-1)n^{2H-2}$ as $n \to \infty$. Then, for any integer $q \ge 2$ such that $H < 1 - \frac{1}{2q}$, we have

$$\sum_{v \in \mathbb{Z}} |\rho(v)|^q < \infty$$

and the Breuer-Major theorem implies

$$\boxed{\frac{1}{\sqrt{n}} \sum_{k=1}^{n} H_q(B^H_k - B^H_{k-1}) \overset{\mathcal{L}}{\to} N(0, \sigma^2_{H,q}),}$$

where $\sigma^2_{H,q} = q! \sum_{v \in \mathbb{Z}} \rho(v)^q$.

9.6.2 CLT for the q-variation of the fBm

For a real $q \ge 1$, set $c_q = E[|Z|^q]$, where $Z \sim N(0,1)$. The Breuer-Major theorem leads to the following convergence:

Theorem 9.6. *Suppose $H < \frac{1}{2}$ and q is not an even integer. As $n \to \infty$ we have*

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left[n^{qH} |B^H_{\frac{k}{n}} - B^H_{\frac{k-1}{n}}|^q - c_q \right] \overset{\mathcal{L}}{\to} N(0, \tilde{\sigma}^2_{H,q}).$$

Proof. Use that $|x|^q - c_q$ has Hermite rank 1. $\square$

9.6.3 Rate of convergence for the quadratic variation

Define for $n \geq 1$,

$$S_n = \sum_{k=1}^{n} (\Delta_{k,n} B^H)^2,$$

where $\Delta_{k,n} B^H = B^H_{\frac{k}{n}} - B^H_{\frac{k-1}{n}}$. Then,

$$n^{2H-1} S_n \stackrel{\text{a.s.}}{\to} 1,$$

n tends to infinity. In fact, by the self-similarity property, $n^{2H-1}S_n$ has the same law as $\frac{1}{n}\sum_{k=1}^{n}(B^H_k - B^H_{k-1})^2$, and the result follows form the Ergodic Theorem. To study the asymptotic normality, consider

$$F_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \left[n^{2H} (\Delta_{k,n} B^H)^2 - 1 \right] \stackrel{\mathcal{L}}{=} \frac{1}{\sigma_n} \sum_{k=1}^{n} \left[(B^H_k - B^H_{k-1})^2 - 1 \right],$$

where σ_n is such that $E[F_n^2] = 1$.

Theorem 9.7. *Assume $H < \frac{3}{4}$. Then, $\lim_{n\to\infty} \frac{\sigma^2}{n} = 2\sum_{r\in\mathbb{Z}} \rho^2(r)$ and*

$$d_{TV}(P_{F_n}, \gamma) \leq c_H \times \begin{cases} n^{-\frac{1}{2}} & \text{if } H \in (0, \frac{5}{8}) \\ n^{-\frac{1}{2}} (\log n)^{\frac{3}{2}} & \text{if } H = \frac{5}{8} \\ n^{4H-3} & \text{if } H \in (\frac{5}{8}, \frac{3}{4}). \end{cases}$$

As a consequence,

$$\sqrt{n}(n^{2H-1} S_n - 1) \stackrel{\mathcal{L}}{\to} N\left(0, 2\sum_{r\in\mathbb{Z}} \rho^2(r)\right).$$

The estimator of H given by $\widehat{H}_n = \frac{1}{2} - \frac{\log S_n}{2\log n}$ satisfies $\widehat{H}_n \stackrel{\text{a.s.}}{\to} H$ and

$$\sqrt{n}\log n(\widehat{H}_n - H) \stackrel{\mathcal{L}}{\to} N\left(0, \frac{1}{2}\sum_{r\in\mathbb{Z}} \rho^2(r)\right).$$

Proof. There exists a sequence $\{e_k, k \geq 1\}$ in $H = L^2([0,T])$ such that

$$\langle e_k, e_j \rangle_H = \rho(k-j).$$

The sequence $\{B(e_k)\}$ has the same law as $\{B^H_k - B^H_{k-1}\}$, and we may replace F_n by

$$G_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \left[B(e_k)^2 - 1 \right] = I_2(f_n),$$

where $f_n = \frac{1}{\sigma_n}\sum_{k=1}^{n} e_k \otimes e_k$. By the isometry property of multiple integrals,

$$1 = E[G_n^2] = 2\|f_n\|^2_{L^2([0,T]^2)} = \frac{2}{\sigma_n^2} \sum_{k,j=1}^{n} \rho^2(k-j) = \frac{2n}{\sigma_n^2} \sum_{|r|<n} \left(1 - \frac{|r|}{n}\right) \rho^2(r).$$

Since $\sum_r \rho^2(r) < \infty$, because $H < \frac{3}{4}$, we deduce that $\lim_{n\to\infty} \frac{\sigma^2}{n} = 2\sum_{r\in\mathbb{Z}} \rho^2(r)$. We can write $D_r[I_2(f_n)] = 2I_1(f_n(\cdot, r)]$ and

$$\|D[I_2(f_n)]\|_H^2 = 4\left(I_2(f_n \otimes_1 f_n) + \|f_n\|_H^2\right) = 4I_2(f_n \otimes_1 f_n) + 2.$$

Therefore,

$$\begin{aligned}
\operatorname{Var}\left(\|D[I_2(f_n)]\|_H^2\right) &= 16E\left[(I_2(f_n \otimes_1 f_n))^2\right] \\
&= 8\|f_n \otimes_1 f_n\|_{L^2([0,T]^2)}^2 \\
&= \frac{16}{\sigma_n^4} \sum_{k,j,i,\ell=1}^{n} \rho(k-j)\rho(i-\ell)\rho(k-i)\rho(j-\ell) \\
&\leq \frac{16}{\sigma_n^4} \sum_{i,\ell=1}^{n} (\rho_n * \rho_n)(i-\ell)^2 \\
&\leq \frac{16n}{\sigma_n^4} \sum_{k\in\mathbb{Z}} (\rho_n * \rho_n)(k)^2 = \frac{16n}{\sigma_n^4} \|\rho_n * \rho_n\|_{\ell^2(\mathbb{Z})}^2,
\end{aligned}$$

where $\rho_n(k) = |\rho(k)|\mathbf{1}_{\{|k|\leq n-1\}}$. Applying Young's inequality yields

$$\|\rho_n * \rho_n\|_{\ell^2(\mathbb{Z})}^2 \leq \|\rho_n\|_{\ell^{4/3}(\mathbb{Z})}^4,$$

so that

$$\operatorname{Var}\left(\|D[I_2(f_n)]\|_H^2\right) \leq \frac{16n}{\sigma_n^4} \left(\sum_{|k|<n} |\rho(k)|^{\frac{4}{3}}\right)^3.$$

Thus,

$$d_{TV}(F_n, Z) \leq \frac{4\sqrt{n}}{\sigma_n^2} \left(\sum_{|k|<n} |\rho(k)|^{\frac{4}{3}}\right)^{\frac{3}{2}}$$

and the result follows from $\rho(k) \sim H(2H-1)|k|^{2H-2}$ as $|k| \to \infty$. □

Remark:

Nourdin-Peccati '13 proved the following optimal version of the fourth moment theorem (assuming $E[F_n^2] = 1$):

$$c\mathbf{M}(F_n) \leq d_{TV}(F_n, Z) \leq C\mathbf{M}(F_n),$$

where $\mathbf{M}(F_n) = \max(|E[F_n^3]|, E[F_n^4] - 3)$. As a consequence, the sequence

$$F_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \left[(B_k^H - B_{k-1}^H)^2 - 1\right]$$

satisfies:

$$d_{TV}(P_{F_n}, \gamma) \approx \begin{cases} n^{-\frac{1}{2}} & \text{if } H \in (0, \frac{2}{3}) \\ n^{-\frac{1}{2}}(\log n)^2 & \text{if } H = \frac{2}{3} \\ n^{6H-\frac{9}{2}} & \text{if } H \in (\frac{2}{3}, \frac{3}{4}), \end{cases}$$

where $\approx$ means that we have an upper and lower bounds with some constants $c_{H,1}$ and $c_{H,2}$.

10 Applications of the Malliavin calculus in finance

In this section we present some applications of Malliavin Calculus to mathematical finance. First we discuss a probabilistic method for numerical computations of price sensitivities (Greeks) based on the integration by parts formula. Then, we discuss the use of Clark-Ocone formula to find hedging portfolios in the Black-Scholes model.

10.1 Black-Scholes model

Consider a market consisting of one stock (risky asset) and one bond (risk-less asset). We assume that the price process $(S_t)_{t\geq 0}$ follows a Black-Scholes model with constant coefficients $\sigma > 0$ and μ, that is,

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B_t\right), \tag{66}$$

where $B = (B_t)_{t\in[0,T]}$ is a Brownian motion defined in a complete probability space $(\Omega, \mathcal{F}, P)$. We will denote by $(\mathcal{F}_t)_{\in[0,T]}$ the filtration generated by the Brownian motion and completed by the P-null sets. By Itô's formula we obtain that S_t satisfies a linear stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{67}$$

The coefficient μ is the *mean return rate* and σ is the *volatility.* The price of the bond at time t is e^{rt}, where r is the interest rate.

Consider an investor who starts with some initial endowment $x \geq 0$ and invests in the assets described above. Let α_t be the number of non-risky assets and β_t the number of stocks owned by the investor at time t. The couple $\phi_t = (\alpha_t, \beta_t)$, $t \in [0, T]$, is called a *portfolio*, and we assume that the processes α_t and β_t are measurable and adapted processes such that

$$\int_0^T \beta_t^2 \, dt < \infty, \int_0^T |\alpha_t| \, dt < \infty$$

almost surely. Then the *value* of the portfolio at time t is $V_t(\phi) = \alpha_t e^{rt} + \beta_t S_t$.We say that the portfolio ϕ is *self-financing* if

$$V_t(\phi) = x + r \int_0^t \alpha_s e^{rs} ds + \int_0^t \beta_s dS_s.$$

From now on we will consider only self-financing portfolios. It is easy to check that the discounted value of a self-financing portfolio $\widetilde{V}_t(\phi) = e^{-rt} V_t(\phi)$ satisfies

$$\widetilde{V}_t(\phi) = x + \int_0^t \beta_s d\widetilde{S}_s,$$

where $\widetilde{S}_t = e^{-rt} S_t$. Notice that

$$d\widetilde{S}_t = (\mu - r)\, \widetilde{S}_t dt + \sigma \widetilde{S}_t dB_t.$$

Set $\theta = \frac{\mu - r}{\sigma}$. Consider the martingale measure defined on $\mathcal{F}_T$, by

$$\frac{dQ}{dP} = \exp\left(-\theta B_t - \frac{\theta^2}{2}\right).$$

Under Q the process $W_t = B_t + \theta t$ is a Brownian motion and the discounted price process $\widetilde{S}_t = e^{-rt} S_t$ is a martingale because

$$d\widetilde{S}_t = \sigma \widetilde{S}_t dW_t.$$

Suppose that $F \geq 0$ is an $\mathcal{F}_T$-measurable such that $E_Q(F^2) < \infty$. The random variable F represents the payoff of some derivative. We say that F can ve *replicated* if there exists a self-financing portfolio ϕ such that $V_T(\phi) = F$. The Itô integral representation theorem implies that any derivative is replicable, and this means that the Black-Scholes market is complete. Indeed, if suffices to write

$$e^{-rT} F = E_Q(e^{-rT} F) + \int_0^T u_s dW_s,$$

and take the self-financing portfolio $\phi_t = (\alpha_t, \beta_t)$, where

$$\beta_t = \frac{u_t}{\sigma \widetilde{S}_t}. \tag{68}$$

The price of a derivative with payoff F at time $t \leq T$ is given by the value at time t of a self-financing portfolio which replicates F.Then,

$$V_t(\phi) = e^{-r(T-t)} E_Q(F|\mathcal{F}_t). \tag{69}$$

10.2 Computation of Greeks

In this section we will present a general integration by parts formula and we will apply it to the computation of Greeks in the Black-Schole smodel.

Let $W = \{W(h), h \in H\}$ denote an isonormal Gaussian process associated with the Hilbert space H. We assume that W is defined on a complete probability space $(\Omega, \mathcal{F}, P)$, and that $\mathcal{F}$ is generated by W.

Proposition 10.1. *Let F, G be two random variables such that $F \in \mathbb{D}^{1,2}$. Consider an H-valued random variable u such that $D_u F = \langle DF, u \rangle_H \neq 0$ a.s. and $Gu(D_u F)^{-1} \in \mathrm{Dom}\delta$. Then, for any continuously differentiable function function f with bounded derivative we have*

$$E(f'(F)G) = E(f(F)H(F,G)), \tag{70}$$

where $H(F,G) = \delta(Gu(D_u F)^{-1})$.

Proof: By the chain rule we have

$$D_u(f(F)) = f'(F) D_u F.$$

Hence, by the duality relationship we get

$$\begin{aligned} E(f'(F)G) &= E\left(D_u(f(F))(D_u F)^{-1} G\right) \\ &= E\left(\langle D(f(F)), u(D_u F)^{-1} G \rangle_H\right) \\ &= E(f(F)\delta(Gu(D_u F)^{-1})). \end{aligned}$$

This completes the proof. □

Remark 10.1. *If the law of F is absolutely continuous, we can assume that the function f is Lipschitz.*

Remark 10.2. *Suppose that u is deterministic. Then, for $Gu(D_uF)^{-1} \in \mathrm{Dom}\delta$ it suffices that $G(D_uF)^{-1} \in \mathbb{D}^{1,2}$. Sufficient conditions for this are the following: $G \in \mathbb{D}^{1,4}$, $F \in \mathbb{D}^{2,2}$, $E(G^6) < \infty$, $E((D_uF)^{-12}) < \infty$, and $E(\|DD_uF\|_H^6) < \infty$.*

Remark 10.3. *Suppose we take $u = DF$. In this case*

$$H(F,G) = \delta\left(\frac{GDF}{\|DF\|_H^2}\right),$$

and Equation (66) yields

$$E(f'(F)G) = E\left(f(F)\delta\left(\frac{GDF}{\|DF\|_H^2}\right)\right). \tag{71}$$

10.2.1 Computation of Greeks for European options

A Greek is a derivative of a financial quantity, usually an option price, with respect to any of the parameters of the model. This derivative is useful to measure the stability of this quantity under variations of the parameter. Consider an option with payoff $F \geq 0$ such that $E_Q(F^2) < \infty$. From (69) its price at time $t = 0$ is given by

$$V_0 = E_Q(e^{-rT}F).$$

The most important Greek is the Delta, denoted by Δ, which by definition is the derivative of V_0 with respect to the initial price of the stock S_0.

Suppose that the payoff F only depents on the price of the stock at the maturity time T. That is, $F = \Phi(S_T)$. We call these derivative European options.

Notice that $\frac{\partial S_T}{\partial S_0} = \frac{S_T}{S_0}$. As a consequence, if Φ is a Lipschitz function we can write

$$\Delta = \frac{\partial V_0}{\partial S_0} = E_Q(e^{-rT}\Phi'(S_T)\frac{\partial S_T}{\partial S_0}) = \frac{e^{-rT}}{S_0}E_Q(\Phi'(S_T)S_T).$$

Now we will apply Proposition 10.1 with $u = 1$, $F = S_T$ and $G = S_T$. We have

$$D_uS_T = \int_0^T D_tS_Tdt = \sigma T\ S_T. \tag{72}$$

Hence, all the conditions appearing in Remark 2 above are satisfies in this case and we we have

$$\delta\left(S_T\left(D_uS_T\right)^{-1}\right) = \delta\left(\frac{1}{\sigma T}\right) = \frac{W_T}{\sigma T}.$$

As a consequence,

$$\Delta = \frac{e^{-rT}}{S_0\sigma T}E_Q(\Phi(S_T)W_T). \tag{73}$$

The Gamma, denoted by Γ, is the second derivative of the option price with respect to S_0. As before we obtain

$$\Gamma = \frac{\partial^2 V_0}{\partial S_0^2} = E_Q\left(e^{-rT}\Phi''(S_T)\left(\frac{\partial S_T}{\partial S_0}\right)^2\right) = \frac{e^{-rT}}{S_0^2}E_Q(\Phi''(S_T)S_T^2).$$

Suppose now that Φ' is Lipschitz. We first apply Proposition 10.1 with S_T^2, $F = S_T$ and $u = 1$. From (72) we have

$$\delta\left(S_T^2\,(D_u S_T)^{-1}\right) = \delta\left(\frac{S_T}{\sigma T}\right) = S_T\left(\frac{W_T}{\sigma T} - 1\right),$$

and, as a consequence,

$$E_Q(\Phi''(S_T)S_T^2) = E_Q\left(\Phi'(S_T)S_T\left(\frac{W_T}{\sigma T} - 1\right)\right).$$

Finally, applying again Proposition 10.1 with $G = S_T\left(\frac{W_T}{\sigma T} - 1\right)$, $F = S_T$ and $u = 1$ yields

$$\begin{aligned}
\delta\left(S_T\left(\frac{W_T}{\sigma T} - 1\right)\left(\int_0^T D_t S_T dt\right)^{-1}\right) &= \delta\left(\frac{W_T}{\sigma^2 T^2} - \frac{1}{\sigma T}\right) \\
&= \left(\frac{W_T^2}{\sigma^2 T^2} - \frac{1}{\sigma^2 T} - \frac{W_T}{\sigma T}\right)
\end{aligned}$$

and,

$$E_Q\left(\Phi'(S_T)S_T\left(\frac{W_T}{\sigma T} - 1\right)\right) = E_Q\left(\Phi(S_T)\ \left(\frac{W_T^2}{\sigma^2 T^2} - \frac{1}{\sigma^2 T} - \frac{W_T}{\sigma T}\right)\right).$$

Therefore, we obtain

$$\Gamma = \frac{e^{-rT}}{S_0^2 \sigma T}E_Q\left(\Phi(S_T)\ \left(\frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T\right)\right). \tag{74}$$

The derivative with respect to the volatility is called Vega, and denoted by ϑ:

$$\vartheta = \frac{\partial V_0}{\partial \sigma} = E_Q(e^{-rT}\Phi'(S_T)\frac{\partial S_T}{\partial \sigma}) = \ e^{-rT}E_Q(\Phi'(S_T)S_T\,(W_T - \sigma T)).$$

Applying Proposition 10.1 with $G = S_T W_T$, $F = S_T$ and $u = 1$ yields

$$\begin{aligned}
\delta\left(S_T\,(W_T - \sigma T)\left(\int_0^T D_t S_T dt\right)^{-1}\right) &= \delta\left(\frac{W_T}{\sigma T} - 1\right) \\
&= \left(\frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T\right).
\end{aligned}$$

As a consequence,

$$\vartheta = e^{-rT}E_Q\left(\Phi(S_T)\left(\frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T\right)\right). \tag{75}$$

By means of an approximation procedure these formulas still hold although the function Φ and its derivative are not Lipschitz. We just need Φ to be piecewise continuous with jump discontinuities and with linear growth. In particular, we can apply these formulas to the case of and European call-option ($\Phi(x) = (x-K)^+$), and European put-option ($\Phi(x) = (K-x)^+$), or a digital option ($\Phi(x) = \mathbf{1}_{\{x>K\}}$). For example, using formulas (73), (74), and (75) we can compute the values of Δ, Γ and ϑ for an European Call option with exercise price K and compare the results with those obtained in this case using the explicit expression for the price

$$V_0 = S_0 N(d_+) - K e^{-rT} N(d_-),$$

where N is the distribution function of the law $N(0,1)$, and

$$d_\pm = \frac{\log \frac{S_0}{K} + \left(r \pm \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}}.$$

We can compute the values of the previous derivatives with a Monte Carlo numerical procedure.

10.2.2 Computation of Greeks for exotic options

Consider options whose payoff is a function of the average of the stock price $\frac{1}{T}\int_0^T S_t dt$, that is

$$F = \Phi\left(\frac{1}{T}\int_0^T S_t dt\right).$$

For instance, an Asiatic Call-option with exercise price K, is a derivative of this type, where $F = \left(\frac{1}{T}\int_0^T S_t dt - K\right)^+$. In this case there is no closed formula for the density of the random variable $\frac{1}{T}\int_0^T S_t dt$. From (69) the price of this option at time $t = 0$ is given by

$$V_0 = e^{-rT} E_Q\left(\Phi\left(\frac{1}{T}\int_0^T S_t dt\right)\right).$$

Let us compute the Delta for this type of options. Set $\overline{S}_T = \frac{1}{T}\int_0^T S_t dt$. We have

$$\Delta = \frac{\partial V_0}{\partial S_0} = E_Q(e^{-rT}\Phi'(\overline{S}_T)\frac{\partial \overline{S}_T}{\partial S_0}) = \frac{e^{-rT}}{S_0} E_Q(\Phi'(\overline{S}_T)\overline{S}_T).$$

We are going to apply Proposition 10.1 with $G = \overline{S}_T$, $F = \overline{S}_T$ and $u_t = S_t$. Let us compute

$$D_t F = \frac{1}{T}\int_0^T D_t S_r dr = \frac{\sigma}{T}\int_t^T S_r dr,$$

and

$$\begin{aligned}
\delta\left(\frac{GS.}{\int_0^T S_t D_t F dt}\right) &= \frac{2}{\sigma}\delta\left(\frac{S.}{\int_0^T S_t dt}\right) \\
&= \frac{2}{\sigma}\left(\frac{\int_0^T S_t dW_t}{\int_0^T S_t dt} + \frac{\int_0^T S_t \left(\int_t^T \sigma S_r dr\right) dt}{\left(\int_0^T S_t dt\right)^2}\right) \\
&= \frac{2}{\sigma}\frac{\int_0^T S_t dW_t}{\int_0^T S_t dt} + 1.
\end{aligned}$$

Notice that

$$\int_0^T S_t dW_t = \frac{1}{\sigma}\left(S_T - S_0 - r\int_0^T S_t dt\right).$$

Thus,

$$\delta\left(\frac{GS.}{\int_0^T S_t D_t F dt}\right) = \frac{2\left(S_T - S_0\right)}{\sigma^2 \int_0^T S_t dt} + 1 - \frac{2r}{\sigma^2} = \frac{2}{\sigma^2}\left(\frac{S_T - S_0}{\int_0^T S_t dt} - m\right),$$

where $m = r - \frac{\sigma^2}{2}$. Finally, we obtain the following expression for the Delta

$$\Delta = \frac{2e^{-rT}}{S_0\sigma^2} E_Q\left(\Phi\left(\overline{S}_T\right)\left(\frac{S_T - S_0}{T\overline{S}_T} - m\right)\right).$$

10.3 Application of the Clark-Ocone formula in hedging

In this section we discuss the application of Clark-Ocone formula to find explicit formulas for a replicating portfolio in the Black-Scholes model.

Suppose that $F \in \mathbb{D}^{1,2}$. Then, applying Clark-Ocone's formula, from (68) we obtain

$$\beta_t = \frac{e^{-r(T-t)}}{\sigma S_t} E_Q\left(D_t F | \mathcal{F}_t\right).$$

Consider the particular case of an European option with payoff $F = \Phi(S_T)$. Then

$$\begin{aligned}
\beta_t &= \frac{e^{-r(T-t)}}{\sigma S_t} E_Q\left(\Phi'(S_T)\sigma S_T | \mathcal{F}_t\right) \\
&= e^{-r(T-t)} E_Q\left(\Phi'(\frac{S_T}{S_t}S_t)\frac{S_T}{S_t} | \mathcal{F}_t\right) \\
&= e^{-r(T-t)} E_Q\left(\Phi'(xS_{T-t})S_{T-t}\right)|_{x=S_t}.
\end{aligned}$$

In this way we recover the fact that β_t coincides with $\frac{\partial F}{\partial x}(t, S_t)$, where $F(t,x)$ is the price function.

Consider now an option whose payoff is a function of the average of the stock price $\overline{S}_T = \frac{1}{T}\int_0^T S_t dt$, that is $F = \Phi\left(\overline{S}_T\right)$. In this case we obtain

$$\beta_t = \frac{e^{T-t}}{S_t} E_Q\left(\Phi'(\overline{S}_T)\frac{1}{T}\int_t^T S_r dr | \mathcal{F}_t\right).$$

We can write

$$\overline{S}_T = \frac{t}{T}\overline{S}_t + \frac{1}{T}\int_t^T S_r dr,$$

where $\overline{S}_t = \frac{1}{t}\int_0^t S_r dr$. As a consequence we obtain

$$\beta_t = \frac{e^{-r(T-t)}}{S_t} E_Q\left(\Phi'\left(\frac{tx}{T} + \frac{y(T-t)}{T}\overline{S}_{T-t}\right)\left(\frac{y(T-t)}{T}\overline{S}_{T-t}\right)\right)|_{x=\overline{S}_t, y=S_t}.$$