Malliavin Calculus and Normal Approximations

by David Nualart

1 Introduction

The Malliavin calculus is a stochastic calculus of variations with respect to the trajectories of the Brownian motion, that was introduced by Paul Malliavin in the 70's to provide a probabilistic proof of Hörmander's hypoellipticity theorem (see [9]).

Basically, the Malliavin calculus is a differential calculus on a probability space equipped with a normal or Gaussian probability measure. The course will cover the main applications ot Mallavin calculus, including the existence and regularity of probability densities and the rate of convergence in normal approximations via Stein's method. We will also discuss some recent applications to the asymptotic behavior of spatial averages of stochastic partial differential equations. These notes are based on the monographs [17, 12] and the article [16].

2 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 2.1. *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 2.2 (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,

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

This completes the proof. $\square$

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

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

$\square$

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 2.1. *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. $\square$

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): 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 infinitesimal generator of P_t in $L^2(\mathbb{R}, \gamma)$, that is, $L = \frac{d}{dt}|_{t=0} P_t$. Moreover, $\frac{dP_t}{dt} = LP_t = P_t L$.

Proposition 2.2 (Mehler's formula). *For any function $f \in L^2(\mathbb{R}, \gamma)$, we have the following formula for the Ornstein-Uhlenbeck semigroup:*

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

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

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

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

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

This completes the proof of the proposition. $\square$

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

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

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 2.3 (First Poincaré inequality)**.** *For any* $f \in C_p^1(\mathbb{R})$*,*

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

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

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

□

This result has 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

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

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 $\|f\|^2_{1,2} = \|f\|^2_{L^2(\mathbb{R},\gamma)} + \|f'\|^2_{L^2(\mathbb{R},\gamma)}$.

2.1 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 $F \colon \mathbb{R}^n \to \mathbb{R}$:

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

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 P. This is the content of the next proposition.

Proposition 2.4. *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. □

Exercises

1. Show that the operator L is the infinitesimal generator of P_t in $L^2(\mathbb{R}, \gamma)$, that is, $Lf = \frac{d}{dt}|_{t=0} P_t f$, in the sense that this limit exists in $L^2(\mathbb{R}, \gamma)$ if and only if

$$\sum_{n=1}^{\infty} \frac{n^2}{n!} \langle f, H_n \rangle^2_{L^2(\mathbb{R},\gamma)} < \infty.$$

   If this holds we say that f belongs to the domain of L.

2. Show that for any $n \geq 0$, $H_{n+1}(x) = xH_n(x) - nH_{n-1}(s)$.

3. Using (1) show that $H_n(0) = 0$ if n is odd and $H_n(0) = \frac{(-1)^{n/2} n!}{2^{n/2}(n/2)!}$ if n is even.

4. Using Exercise 2, the fact that $H_n' = nH_{n-1}$ and an induction argument, show that

$$H_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!(-1)^k}{k!(n-2k)!2^k} x^{n-2k}.$$

5. Let X and Y be two random variables with joint Gaussian distribution such that $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ and $\mathbb{E}(X^2) = \mathbb{E}(Y^2) = 1$. Show that for any $n, m \geq 0$, we have

$$\mathbb{E}(H_n(X)H_m(Y)) = \begin{cases} n![\mathbb{E}(XY)]^n & \text{if } n = m. \\ 0 & \text{if } n \neq m. \end{cases}$$

6. Let $(B_t)_{t\geq 0}$ be a Brownian motion. Consider the linear stochastic differential equation

$$dX_t^x = \sqrt{2}dB_t - X_t^x dt, \quad t \geq 0, \quad X_0^x = x \in \mathbb{R}. \tag{2}$$

The solution of (2) is called an *Ornstein-Uhlenbeck stochastic process.*

(i) For any $x \in \mathbb{R}$, show that the solution of (2) is given by

$$X_t^x = e^{-t}x + \sqrt{2}\int_0^t e^{-(t-s)}dB_t, \quad t \geq 0.$$

(ii) For any $x \in \mathbb{R}$ and any function $f \in L^2(\mathbb{R}, \gamma)$, show that $P_t f(x) = \mathbb{E}[f(X_t^x)]$.

3 Brownian motion

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 3.1. *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:*

(i) Almost surely $B_0 = 0$.

(ii) 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.

(iii) If $0 \leq s < t$, the increment $B_t - B_s$ is a Gaussian random variable with mean zero and variance $t - s$.

(iv) With probability one, the map $t \to B_t$ is continuous.

More generally, a d-dimensional Brownian motion is defined as an $\mathbb{R}^d$-valued stochastic process $B = (B_t)_{t\geq 0}$, $B_t = (B_t^1, \ldots, B_t^d)$, where $B^1, \ldots, B^d$ are d independent Brownian motions.

Proposition 3.1. *Properties (i), (ii), and (iii) are equivalent to saying that B is a Gaussian process with mean zero and covariance function*

$$\Gamma(s,t) = \min(s,t). \tag{3}$$

Proof. Suppose (i), (ii), and (iii) hold. The probability distribution of the random vector $(B_{t_1}, \ldots, B_{t_n})$, for $0 < t_1 < \cdots < t_n$, is normal because this vector is a linear transformation of the vector $\left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}\right)$, which has a normal distribution, because its components are independent and normal. The mean $m(t)$ and the covariance function $\Gamma(s,t)$ are given by

$$\begin{aligned} m(t) = \mathbb{E}(B_t) &= 0, \\ \Gamma(s,t) = \mathbb{E}(B_s B_t) &= \mathbb{E}(B_s(B_t - B_s + B_s)) \\ &= \mathbb{E}(B_s(B_t - B_s)) + \mathbb{E}(B_s^2) = s = \min(s,t), \end{aligned}$$

if $s \leq t$. The converse is also easy to show. □

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

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

In fact, 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

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

As a consequence, 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 $T > 0$.

Remark 3.2. *As we have seen, the trajectories of Brownian motion on an interval $[0,T]$ are Hölder continuous of order γ for any $\gamma < \frac{1}{2}$. However, the trajectories are not Hölder continuous of order $\frac{1}{2}$. More precisely, the following property holds:*

$$P\left(\sup_{s,t \in [0,1]} \frac{|B_t - B_s|}{\sqrt{|t-s|}} = +\infty \right) = 1.$$

The exact modulus of continuity of Brownian motion was obtained by Paul Lévy:

$$\limsup_{\delta \downarrow 0} \sup_{s,t \in [0,1], |t-s| < \delta} \frac{|B_t - B_s|}{\sqrt{2|t-s| \log |t-s|}} = 1, \quad \text{a.s.}$$

Brownian motion can also be regarded as the limit in distribution of a symmetric random walk. Indeed, fix a time interval $[0,T]$. Consider n independent and identically distributed random variables $\xi_1, \ldots, \xi_n$ with zero mean and variance $\frac{T}{n}$. Define the partial sums

$$R_k = \xi_1 + \cdots + \xi_k, \quad k = 1, \ldots, n.$$

By the central limit theorem the sequence R_n converges in distribution, as n tends to infinity, to the normal distribution $N(0,T)$.

Consider the continuous stochastic process $S_n(t)$ defined by linear interpolation from the values

$$S_n\left(\frac{kT}{n}\right) = R_k, \quad k = 0, \ldots, n.$$

Then, a functional version of the central limit theorem, known as the Donsker invariance principle, says that the sequence of stochastic processes $S_n(t)$ converges in law to Brownian motion on $[0,T]$. This means that for any continuous and bounded function $\varphi\colon C([0,T]) \to \mathbb{R}$, we have

$$\mathbb{E}(\varphi(S_n)) \to \mathbb{E}(\varphi(B)),$$

as n tends to infinity.

3.1 Quadratic variation

Brownian motion satisfies $\mathbb{E}(|B_t - B_s|^2) = t - s$ for all $s \leq t$. This means that when $t-s$ is small, $B_t - B_s$ is of order $\sqrt{t-s}$ and $(B_t - B_s)^2$ is of order $t-s$. Moreover, the quadratic variation of a Brownian motion on $[0,t]$ equals t in $L^2(\Omega)$, as is proved in the following proposition.

Proposition 3.2. *Fix a time interval $[0,t]$ and consider a subdivision π of this interval*

$$0 = t_0 < t_1 < \cdots < t_n = t.$$

The norm of the subdivision π is defined as $|\pi| = \max_{0\leq j\leq n-1}(t_{j+1} - t_j)$. The following convergence holds in $L^2(\Omega)$

$$\lim_{|\pi|\to 0} \sum_{j=0}^{n-1} \left(B_{t_{j+1}} - B_{t_j}\right)^2 = t. \tag{4}$$

Proof. Set $\xi_j = (B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)$. The random variables ξ_j are independent and centered. Thus,

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{j=0}^{n-1} \left(B_{t_{j+1}} - B_{t_j}\right)^2 - t\right)^2\right] &= \mathbb{E}\left[\left(\sum_{j=0}^{n-1} \xi_j\right)^2\right] = \sum_{j=0}^{n-1} \mathbb{E}\left[\xi_j^2\right] \\
&= \sum_{j=0}^{n-1} \left[3\left(t_{j+1} - t_j\right)^2 - 2\left(t_{j+1} - t_j\right)^2 + \left(t_{j+1} - t_j\right)^2\right] \\
&= 2\sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right)^2 \leq 2t|\pi| \overset{|\pi|\to 0}{\longrightarrow} 0,
\end{aligned}$$

which proves the result. $\square$

As a consequence, we have the following result.

Proposition 3.3. *The total variation of Brownian motion on an interval $[0,t]$, defined by*

$$V = \sup_{\pi} \sum_{j=0}^{n-1} \left|B_{t_{j+1}} - B_{t_j}\right|,$$

where $\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$, is infinite with probability one.

Proof. Using the continuity of the trajectories of Brownian motion, we have

$$\begin{aligned}\sum_{j=1}^{n-1}\left(B_{t_{j+1}}-B_{t_j}\right)^2 &\leq \sup_j\left|B_{t_{j+1}}-B_{t_j}\right|\left(\sum_{j=0}^{n-1}\left|B_{t_{j+1}}-B_{t_j}\right|\right)\\ &\leq V\sup_j\left|B_{t_{j+1}}-B_{t_j}\right|\stackrel{|\pi|\to 0}{\longrightarrow}0,\end{aligned}$$

if $V<\infty$, which contradicts the fact that $\sum_{j=0}^{n-1}(B_{t_{j+1}}-B_{t_j})^2$ converges in mean square to t as $|\pi|\to 0$. Therefore, $P(V<\infty)=0$. $\square$

3.2 Wiener space

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

$$C=\{\omega\in\Omega:\omega(t_1)\in A_1,\dots,\omega(t_k)\in A_k\},\tag{5}$$

  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 (5) by

$$P(C)=\int_{A_1\times\cdots\times A_k}p_{t_1}(x_1)p_{t_2-t_1}(x_2-x_1)\cdots p_{t_k-t_{k-1}}(x_k-x_{k-1})\,dx_1\cdots dx_k,\tag{6}$$

  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 (6) 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.

3.3 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

$$h_t=\sum_{j=0}^{n-1}a_j\mathbf{1}_{(t_j,t_{j+1}]}(t),\quad t\geq 0,\tag{7}$$

where $n \geq 1$ is an integer, $a_0, \ldots, a_{n-1} \in \mathbb{R}$, and $0 = t_0 < \cdots < t_n$. The Wiener integral of a step function $h \in \mathcal{E}_0$ of the form (7) is defined by

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

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

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

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

$$h \to \int_0^\infty h_t dB_t$$

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 $\int_0^\infty h_t dB_t$ is called the Wiener integral of $h \in L^2(\mathbb{R}_+)$ and is denoted by $B(h)$. Observe that it is a Gaussian random variable with mean zero and variance $\|h\|^2_{L^2(\mathbb{R}_+)}$. Notice that $B_t = B(\mathbf{1}_{[0,t]})$.

3.4 Isonormal Gaussian space

The Wiener integral $B(h) = \int_0^\infty h(t) dB_t$ gives rise to a centered Gaussian family of random variables, indexed by the Hilbert space $H = L^2(\mathbb{R}_+)$, such that

$$\mathbb{E}[B(h)B(g)] = \langle h, g \rangle_{L^2(\mathbb{R}_+)}.$$

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 $\mathcal{H}_1 = \{W(h), h \in H\}$ satisfying

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

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

Examples:

(A) Let $(T, \mathcal{B}, \mu)$ be a measure space, where μ is a σ-finite measure. Consider an isonormal Gaussian process W on the Hilbert space $H = L^2(T, \mathcal{B}, \mu)$. Define $W(A) = W(\mathbf{1}_A)$, for $A \in \mathcal{B}$ such that $\mu(A) < +\infty$. Then, $\{W(A), A \in \mathcal{B}, \mu(A) < +\infty\}$, is a random measure with covariance

$$\mathbb{E}\left(W(A) \cap W(B)\right) = \mu(A \cap B),$$

which is called a *white noise* on $(T, \mathcal{B}, \mu)$.

(B) Let $X = (X_t)_{t\in[0,T]}$ be a continuous centered Gaussian process with covariance function

$$R(s,t) = \mathbb{E}\left(X_t X_s\right).$$

The Gaussian subspace generated by X can be identified with an isonormal Gaussian process $\mathcal{H}_1 = \{X(h), h \in H\}$, where the separable Hilbert space H is defined as follows. We denote by $\mathcal{E}$ the set of step functions on $[0,T]$. Let H be the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the scalar product

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle_H = R(t,s).$$

The mapping $\mathbf{1}_{[0,t]} \mapsto X_t$ can be extended to an isometry $h \to X(h)$ between H and $\mathcal{H}_1$.

An isonormal Gaussian process on a separable Hilbert space H can de defined as follows. Suppose that $\{e_i, i \geq 1\}$ is a complete orthonormal system on H. Consider a sequence of independent random variables $\{\xi_i, i \geq 1\}$ and define

$$W(h) = \sum_{i=1}^{\infty} \langle h, e_i \rangle_H \xi_i.$$

Exercises

1. Let Z be a random variable with law $N(0,1)$. Consider the process $X_t = \sqrt{t}Z$, $t \geq 0$. Is $(X_t)_{t\geq 0}$ a Brownian motion? Which properties of Definition 3.1 hold and which don't?

2. Let B be a d-dimensional Brownian motion. Consider an orthogonal $d \times d$ matrix U (that is, $UU^T = I_d$, where I_d denotes the identity matrix of order d). Show that the process

$$X_t = UB_t, \quad t \geq 0,$$

is a d-dimensional Brownian motion.

3. Compute the mean and covariance function of the following stochastic processes related to Brownian motion:

   (a) *Brownian bridge*: $X_t = B_t - tB_1$, $t \in [0,1]$.

   (b) *Brownian motion with drift*: $X_t = \sigma B_t + \mu t$, $t \geq 0$, where $\sigma > 0$ and $\mu \in \mathbb{R}$ are constants.

   (c) *Geometric Brownian motion*: $X_t = e^{\sigma B_t + \mu t}$, $t \geq 0$, where $\sigma > 0$ and $\mu \in \mathbb{R}$ are constants.

   Which of them are Gaussian processes?

4. Define $X_t = \int_0^t B_s ds$, where $B = (B_t)_{t \geq 0}$ is a Brownian motion. Show that X_t is a Gaussian random variable. Compute its mean and its variance.

5. Let B be a Brownian motion. Show that for any $a > 0$, the process $(a^{-\frac{1}{2}} B_{at})_{t \geq 0}$ is a Brownian motion (*Selfsimilarity property*).

6. Let B be a Brownian motion. Show that for any $h > 0$, the process $(B_{t+h} - B_h)_{t \geq 0}$ is a Brownian motion.

7. Let B be a Brownian motion. Show that the process $(-B_t)_{t \geq 0}$ is a Brownian motion.

8. Let B be a Brownian motion. Show that $\lim_{t \to \infty} B_t / t = 0$ almost surely and that the process
$$X_t = \begin{cases} tB_{1/t}, & t > 0, \\ 0, & t = 0, \end{cases}$$
is a Brownian motion.

9. Let B be a Brownian motion. Using the Borel-Cantelli lemma, show that if $(\pi^n)_{n \geq 1}$, $\pi^n = \{0 = t_0^n < \cdots < t_{k_n}^n = t\}$, is a sequence of partitions of $[0, t]$ such that $\sum_n |\pi^n| < \infty$, then $\sum_{j=0}^{k_n - 1} (B_{t_{j+1}^n} - B_{t_j^n})^2$ converges almost surely to t.

4 Stochastic calculus

The first aim of this section is to construct Itô's stochastic integrals of the form $\int_0^\infty u_t dB_t$, where $B = (B_t)_{t\geq 0}$ is a Brownian motion and $u = (u_t)_{t\geq 0}$ is an adapted process. We then prove Itô's formula, which is a change of variables formula for Itô's stochastic integrals that plays a crucial role in the applications of stochastic calculus. We then discuss some consequences of Itô's formula including the integral representation theorem for square integrable random variables.

4.1 Stochastic integrals

Recall that $B = (B_t)_{t\geq 0}$ is a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$. For any time $t \geq 0$, we define the σ-field $\mathcal{F}_t$ generated by the random variables $(B_s)_{0\leq s\leq t}$ and the events in $\mathcal{F}$ of probability zero. That is, $\mathcal{F}_t$ is the smallest σ-field that contains the sets of the form

$$\{B_s \in A\} \cup N,$$

where $0 \leq s \leq t$, A is a Borel subset of $\mathbb{R}$, and $N \in \mathcal{F}$ is such that $P(N) = 0$. Notice that $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$; that is, $(\mathcal{F}_t)_{t\geq 0}$ is a non-decreasing family of σ-fields. We say that $(\mathcal{F}_t)_{t\geq 0}$ is the *natural filtration* of Brownian motion on the probability space $(\Omega, \mathcal{F}, P)$.

We say that a stochastic process $(u_t)_{t\geq 0}$ is *adapted* if for any $t \geq 0$, u_t is $\mathcal{F}_t$-measurable.

We know that the trajectories of Brownian motion have infinite variation on any finite interval. So, in general, we cannot define the integral

$$\int_0^\infty u_t(\omega) dB_t(\omega)$$

as a pathwise integral. We are going to construct the integral $\int_0^\infty u_t dB_t$ by means of a global probabilistic approach for a class of processes satisfying some adaptability and integrability conditions specified below.

Definition 4.1. *We say that a stochastic process $u = (u_t)_{t\geq 0}$ is progressively measurable if, for any $t \geq 0$, the restriction of u to $\Omega \times [0, t]$ is $\mathcal{F}_t \times \mathcal{B}([0, t])$-measurable.*

Remark 4.2. *If u is adapted and measurable (the mapping $(\omega, s) \longrightarrow u_s(\omega)$ is measurable on the product space $\Omega \times \mathbb{R}_+$ with respect to the product σ-field $\mathcal{F} \times \mathcal{B}(\mathbb{R}_+)$), then there is a version of u which is progressively measurable. Progressive measurability will guarantee that random variables of the form $\int_0^t u_s ds$ are $\mathcal{F}_t$-measurable.*

Let $\mathcal{P}$ be the σ-field of sets $A \subset \Omega \times \mathbb{R}_+$ such that $\mathbf{1}_A$ is progressively measurable. We denote by $L^2(\mathcal{P})$ the Hilbert space $L^2(\Omega \times \mathbb{R}_+, \mathcal{P}, P \times \ell)$, where ℓ is the Lebesgue measure, equipped with the norm

$$\|u\|^2_{L^2(\mathcal{P})} = \mathbb{E}\left(\int_0^\infty u_s^2 ds\right) = \int_0^\infty \mathbb{E}\left(u_s^2\right) ds,$$

where the last equality follows from Fubini's theorem.

In this section we define the stochastic integral $\int_0^\infty u_t dB_t$ of a processes u in $L^2(\mathcal{P})$ as the limit in mean square of the integral of simple processes.

Definition 4.3. *A process $u = (u_t)_{t\geq 0}$ is called a simple process if it is of the form*

$$u_t = \sum_{j=0}^{n-1} \phi_j \mathbf{1}_{(t_j, t_{j+1}]}(t), \tag{8}$$

where $0 \leq t_0 < t_1 < \cdots < t_n$, and the ϕ_j are $\mathcal{F}_{t_j}$-measurable random variables such that $\mathbb{E}(\phi_j^2) < \infty$. We denote by $\mathcal{E}$ the space of simple processes.

We define the stochastic integral of a process $u \in \mathcal{E}$ of the from (8) as

$$I(u) := \int_0^\infty u_t dB_t = \sum_{j=0}^{n-1} \phi_j \left(B_{t_{j+1}} - B_{t_j} \right).$$

The stochastic integral defined on the space $\mathcal{E}$ of simple processes has the following properties:

1. *Linearity:* For any $a, b \in \mathbb{R}$, and simple processes $u, v \in \mathcal{E}$,

$$\int_0^\infty (au_t + bv_t)\, dB_t = a \int_0^\infty u_t dB_t + b \int_0^\infty v_t dB_t.$$

2. *Zero mean:* For any $u \in \mathcal{E}$,

$$\mathbb{E}\left(\int_0^\infty u_t dB_t \right) = 0.$$

In fact, assuming that u is given by (8), and taking into account that the random variables ϕ_j and $B_{t_{j+1}} - B_{t_j}$ are independent, we obtain

$$\mathbb{E}\left(\int_0^\infty u_t dB_t \right) = \sum_{j=0}^{n-1} \mathbb{E}\left[\phi_j \left(B_{t_{j+1}} - B_{t_j} \right) \right] = \sum_{j=0}^{n-1} \mathbb{E}(\phi_j)\mathbb{E}(B_{t_{j+1}} - B_{t_j}) = 0.$$

3. *Isometry property:* For any $u \in \mathcal{E}$,

$$\mathbb{E}\left[\left(\int_0^\infty u_t dB_t \right)^2 \right] = \mathbb{E}\left(\int_0^\infty u_t^2 dt \right).$$

Proof. Assume that u is given by (8). Set $\Delta B_j = B_{t_{j+1}} - B_{t_j}$. Then

$$\mathbb{E}\left(\phi_i \phi_j \Delta B_i \Delta B_j \right) = \begin{cases} 0 & \text{if } \ i \neq j, \\ \mathbb{E}\left(\phi_j^2 \right) (t_{j+1} - t_j) & \text{if } \ i = j, \end{cases}$$

because if $i < j$ the random variables $\phi_i \phi_j \Delta B_i$ and ΔB_j are independent, and if $i = j$ the random variables ϕ_i^2 and $(\Delta B_i)^2$ are independent. So, we obtain

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^\infty u_t dB_t\right)^2\right] &= \sum_{i,j=0}^{n-1} \mathbb{E}\left(\phi_i \phi_j \Delta B_i \Delta B_j\right) = \sum_{i=0}^{n-1} \mathbb{E}\left(\phi_i^2\right)(t_{i+1} - t_i) \\
&= \mathbb{E}\left(\int_0^\infty u_t^2 dt\right).
\end{aligned}$$

□

The extension of the stochastic integral to the class $L^2(\mathcal{P})$ is based on the following density result.

Proposition 4.1. *The space $\mathcal{E}$ of simple processes is dense in $L^2(\mathcal{P})$.*

Proof. We first establish that any $u \in L^2(\mathcal{P})$ can be approximated by processes continuous in $L^2(\Omega)$. Then our result will follow if we show that the simple function are dense in the space of $L^2(\Omega)$-continuous processes.

If u belongs to $L^2(\mathcal{P})$, we define

$$u_t^{(n)} = n \int_{(t-\frac{1}{n})\vee 0}^t u_s ds.$$

The processes $u_t^{(n)}$ are continuous in $L^2(\Omega)$ and satisfy

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^\infty \left|u_t - u_t^{(n)}\right|^2 dt\right) = 0.$$

Indeed, for each ω we have

$$\left|u_t(\omega) - u_t^{(n)}(\omega)\right|^2 \overset{n\to\infty}{\longrightarrow} 0,$$

and we can apply the dominated convergence theorem because

$$\int_0^\infty \left|u_t^{(n)}(\omega)\right|^2 dt \le \int_0^\infty |u_t(\omega)|^2 dt.$$

Suppose $u \in L^2(\mathcal{P})$ is continuous in $L^2(\Omega)$. In this case, we can choose the approximating processes $u_t^{(n,N)} \in \mathcal{E}$ defined by

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

where $t_j = \frac{jN}{n}$. The continuity in mean square of u implies that

$$\mathbb{E}\left(\int_0^\infty \left|u_t - u_t^{(n,N)}\right|^2 dt\right) \le \mathbb{E}\left(\int_N^\infty u_t^2 dt\right) + N \sup_{|t-s|\le N/n} \mathbb{E}\left(|u_t - u_s|^2\right).$$

This converges to zero if we first let $n \to \infty$ and then $N \to \infty$. □

Proposition 4.2. *The stochastic integral can be extended to a linear isometry*

$$I\colon L^2(\mathcal{P}) \to L^2(\Omega).$$

Proof. The stochastic integral of a process u in $L^2(\mathcal{P})$ is defined as the following limit in mean square

$$I(u) = \lim_{n\to\infty} \int_0^\infty u_t^{(n)} dB_t, \tag{9}$$

where $u^{(n)}$ is a sequence of simple processes which converges to u in the norm of $L^2(\mathcal{P})$. Notice that the limit (9) exists because the sequence of random variables $\int_0^\infty u_t^{(n)} dB_t$ is Cauchy in $L^2(\Omega)$, due to the isometry property

$$\begin{aligned} \mathbb{E}\left[\left(\int_0^\infty u_t^{(n)} dB_t - \int_0^\infty u_t^{(m)} dB_t\right)^2\right] &= \mathbb{E}\left(\int_0^\infty \left(u_t^{(n)} - u_t^{(m)}\right)^2 dt\right) \\ &\le 2\mathbb{E}\left(\int_0^\infty \left(u_t^{(n)} - u_t\right)^2 dt\right) + 2\mathbb{E}\left(\int_0^\infty \left(u_t - u_t^{(m)}\right)^2 dt\right) \overset{n,m\to\infty}{\longrightarrow} 0. \end{aligned}$$

On the other hand, it is easy to show that the limit (9) does not depend on the approximating sequence $u^{(n)}$. $\square$

The stochastic integral has the following properties: for any $u, v \in L^2(\mathcal{P})$,

$$\mathbb{E}\left(I(u)\right) = 0 \quad \text{and} \quad \mathbb{E}\left(I(u)I(v)\right) = \mathbb{E}\left(\int_0^\infty u_s v_s ds\right).$$

For any $T > 0$, we set

$$\int_0^T u_s dB_s = \int_0^\infty u_s \mathbf{1}_{[0,T]}(s) dB_s, \tag{10}$$

which is the indefinite integral of u with respect to B. Notice that in order to define $\int_0^T u_s dB_s$, we only need that $u \in L_T^2(\mathcal{P})$, where

$$L_T^2(\mathcal{P}) := L^2(\Omega \times [0,T], \mathcal{P}|_{\Omega\times[0,T]}, P \times \ell).$$

Example 4.4. *For any $T > 0$, we have*

$$\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

Indeed, the process B being continuous in mean square, we can choose as approximating sequence

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

where $t_j = \frac{jT}{n}$, and we obtain, using Proposition 3.2,

$$\begin{aligned}\int_0^T B_t dB_t &= \lim_{n\to\infty} \sum_{j=0}^{n-1} B_{t_j}\left(B_{t_{j+1}} - B_{t_j}\right) \\ &= \frac{1}{2}\lim_{n\to\infty}\sum_{j=0}^{n-1}\left(B_{t_{j+1}}^2 - B_{t_j}^2\right) - \frac{1}{2}\lim_{n\to\infty}\sum_{j=0}^{n-1}\left(B_{t_{j+1}} - B_{t_j}\right)^2 \\ &= \frac{1}{2}B_T^2 - \frac{1}{2}T,\end{aligned}$$

where the convergence holds in $L^2(\Omega)$.

4.2 Indefinite stochastic integrals

We denote by $L^2_\infty(\mathcal{P})$ the set of progressively measurable process such that $\mathbb{E}\left(\int_0^t u_s^2 ds\right) < \infty$ for each $t > 0$. For any process $u \in L^2_\infty(\mathcal{P})$, we can define the indefinite integral process

$$\left\{\int_0^t u_s dB_s, t \geq 0\right\}.$$

Indefinite integrals have the following properties:

1. *Additivity*: For any $a \leq b \leq c$, we have

$$\int_a^b u_s dB_s + \int_b^c u_s dB_s = \int_a^c u_s dB_s.$$

2. *Factorization*: If $a < b$, and F is a bounded and $\mathcal{F}_a$-measurable random variable, then

$$\int_a^b F u_s dB_s = F\int_a^b u_s dB_s.$$

3. *Martingale property:*

Proposition 4.3. *Let $u \in L^2_\infty(\mathcal{P})$. Then, the indefinite stochastic integral*

$$M_t = \int_0^t u_s dB_s, \quad t \geq 0,$$

is a square integrable martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ and admits a continuous version.

Proof. We first prove the martingale property. Suppose that $u \in \mathcal{E}$ is a simple process of the form

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

Then, for any $s \le t$,

$$\begin{aligned}
\mathbb{E}\left(\int_0^t u_v dB_v | \mathcal{F}_s\right) &= \sum_{j=0}^{n-1} \mathbb{E}\left(\phi_j (B_{t_{j+1}\wedge t} - B_{t_j \wedge t}) | \mathcal{F}_s\right) \\
&= \sum_{j=0}^{n-1} \mathbb{E}\left(\mathbb{E}\left(\phi_j (B_{t_{j+1}\wedge t} - B_{t_j \wedge t}) | \mathcal{F}_{t_j \vee s}\right) | \mathcal{F}_s\right) \\
&= \sum_{j=0}^{n-1} \mathbb{E}\left(\phi_j \mathbb{E}\left(B_{t_{j+1}\wedge t} - B_{t_j \wedge t} | \mathcal{F}_{t_j \vee s}\right) | \mathcal{F}_s\right) \\
&= \sum_{j=0}^{n-1} \phi_j (B_{t_{j+1}\wedge s} - B_{t_j \wedge s}) = \int_0^s u_v dB_v.
\end{aligned}$$

So, $M_t = \int_0^t u_s dB_s$ is an $\mathcal{F}_t$-martingale if $u \in \mathcal{E}$.

Fix $T > 0$. In the general case, let $u^{(n)}$ be a sequence of simple processes that converges to u in $L^2_T(\mathcal{P})$. Then, for any $t \in [0, T]$,

$$\int_0^t u_s^{(n)} dB_s \overset{L^2(\Omega)}{\longrightarrow} \int_0^t u_s dB_s.$$

Taking into account that the convergence in $L^2(\Omega)$ implies the convergence in $L^2(\Omega)$ of the conditional expectations, we deduce that $\left(\int_0^t u_s dB_s\right)_{t \in [0,T]}$ is a martingale. This holds for any $T > 0$, which implies that the indefinite integral process is a martingale on $\mathbb{R}_+$.

Let us prove that the indefinite integral has a continuous version. Let $u \in L^2_\infty(\mathcal{P})$ and fix $T > 0$. Consider a sequence of simple processes $u^{(n)}$ which converges to u in $L^2_T(\mathcal{P})$. By the continuity of the paths of Brownian motion, the stochastic integral $M_t^{(n)} = \int_0^t u_s^{(n)} dB_s$ has continuous trajectories. Then, taking into account that $M^{(n)}$ is a martingale, Doob's maximal inequality yields for any $\lambda > 0$,

$$\begin{aligned}
P\left(\sup_{0 \le t \le T} |M_t^{(n)} - M_t^{(m)}| > \lambda\right) &\le \frac{1}{\lambda^2} \mathbb{E}\left(|M_T^{(n)} - M_T^{(m)}|^2\right) \\
&= \frac{1}{\lambda^2} \mathbb{E}\left(\int_0^T \left|u_t^{(n)} - u_t^{(m)}\right|^2 dt\right) \overset{n,m\to\infty}{\longrightarrow} 0.
\end{aligned}$$

We can choose an increasing sequence of natural numbers $n_k, k = 1, 2, \ldots$ such that

$$P\left(\sup_{0 \le t \le T} |M_t^{(n_{k+1})} - M_t^{(n_k)}| > 2^{-k}\right) \le 2^{-k}.$$

The events $A_k := \left\{\sup_{0 \le t \le T} |M_t^{(n_{k+1})} - M_t^{(n_k)}| > 2^{-k}\right\}$ satisfy

$$\sum_{k=1}^{\infty} P(A_k) < \infty.$$

Hence, the Borel-Cantelli lemma implies that $P(\limsup_{k\to\infty} A_k) = 0$. Set $N = \limsup_{k\to\infty} A_k$. Then for any $\omega \notin N$, there exists $k_1(\omega)$ such that for all $k \geq k_1(\omega)$,

$$\sup_{0\leq t\leq T} |M_t^{(n_{k+1})}(\omega) - M_t^{(n_k)}(\omega)| \leq 2^{-k}.$$

As a consequence, if $\omega \notin N$, the sequence $M_t^{(n_k)}(\omega)$ is uniformly convergent on $[0,T]$ to a continuous function $J_t(\omega)$. On the other hand, we know that for any $t \in [0,T]$, $M_t^{(n_k)}$ converges in $L^2(\Omega)$ to $\int_0^t u_s dB_s$. So, $J_t(\omega) = \int_0^t u_s dB_s$ almost surely, for all $t \in [0,T]$. Since $T > 0$ is arbitrary, this implies the existence of a continuous version for $(M_t)_{t\geq 0}$. □

4. *Maximal inequalities:* For any $T, \lambda > 0$, and $u \in L^2_\infty(\mathcal{P})$,

$$P\left(\sup_{t\in[0,T]} |M_t| > \lambda\right) \leq \frac{1}{\lambda^2}\mathbb{E}\left(\int_0^T u_t^2 dt\right), \tag{11}$$

and

$$\mathbb{E}\left(\sup_{t\in[0,T]} |M_t|^2\right) \leq 4\mathbb{E}\left(\int_0^T u_t^2 dt\right). \tag{12}$$

These inequalities are a direct consequence of Proposition 4.3 and Doob's maximal inequalities. We remark that if u belongs to $L^2(\mathcal{P})$, then these inequalities also hold replacing T by ∞.

5. *Quadratic variation of the integral process*:

Proposition 4.4. *Let $u \in L^2_\infty(\mathcal{P})$. Consider a subdivision of the interval* $[0,t]$

$$\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}.$$

Then, as $|\pi| \to 0$,

$$S_\pi^2(u) := \sum_{j=0}^{n-1} \left(\int_{t_j}^{t_{j+1}} u_s dB_s\right)^2 \stackrel{L^1(\Omega)}{\longrightarrow} \int_0^t u_s^2 ds. \tag{13}$$

Proof. As a consequence of the quadratic variation property of Brownian motion (see Proposition 3.2), it is easy to see that (13) holds when $u \in \mathcal{E}$. Let $u \in L^2_\infty(\mathcal{P})$. Fix $\epsilon > 0$ and let $v \in \mathcal{E}$ such that

$$\mathbb{E}\left(\int_0^t (u_s - v_s)^2 ds\right) < \epsilon. \tag{14}$$

Then, we write

$$\mathbb{E}\left(\left|S_\pi^2(u) - \int_0^t u_s^2 ds\right|\right) \leq \mathbb{E}\left(\left|S_\pi^2(u) - S_\pi^2(v)\right|\right) + \mathbb{E}\left(\left|S_\pi^2(v) - \int_0^t v_s^2 ds\right|\right)$$
$$+ \mathbb{E}\left(\left|\int_0^t (u_s^2 - v_s^2) ds\right|\right).$$

As (13) holds for $v \in \mathcal{E}$, the second summand converges to 0 as $|\pi| \to 0$. On the other hand, since $u_s^2 - v_s^2 = (u_s + v_s)(u_s - v_s)$, using Cauchy-Schwarz inequality and (14), we obtain

$$\begin{aligned}\mathbb{E}\left(\Big|\int_0^t (u_s^2 - v_s^2)ds\Big|\right) &\leq \sqrt{\epsilon}\left\{\mathbb{E}\left(\int_0^t (u_s + v_s)^2 ds\right)\right\}^{1/2} \\ &\leq \sqrt{\epsilon}\left(\left\{\mathbb{E}\left(\int_0^t u_s^2 ds\right)\right\}^{1/2} + \sqrt{\epsilon}\right).\end{aligned}$$

Similarly,

$$\mathbb{E}\left(\big|S_\pi^2(u) - S_\pi^2(v)\big|\right) \leq \sqrt{\epsilon}\left(\left\{\mathbb{E}\left(\int_0^t u_s^2 ds\right)\right\}^{1/2} + \sqrt{\epsilon}\right).$$

Therefore,

$$\limsup_{|\pi|\to 0} \mathbb{E}\left(\Big|S_\pi^2(u) - \int_0^t u_s^2 ds\Big|\right) \leq 2\sqrt{\epsilon}\left(\left\{\mathbb{E}\left(\int_0^t u_s^2 ds\right)\right\}^{1/2} + \sqrt{\epsilon}\right).$$

Taking into account that $\epsilon > 0$ is arbitrary, we get the desired convergence. □

As a consequence of this proposition and Burkholder-Davis-Gundy inequality, we get the following result.

Theorem 4.5. *Let $u \in L^2_\infty(\mathcal{P})$. Then, for any $p > 0$ and $T > 0$, we have*

$$c_p \mathbb{E}\left(\left|\int_0^T u_s^2 ds\right|^{\frac{p}{2}}\right) \leq \mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t u_s dB_s\right|^p\right) \leq C_p \mathbb{E}\left(\left|\int_0^T u_s^2 ds\right|^{\frac{p}{2}}\right).$$

6. *Stochastic integration up to a stopping time*:

Proposition 4.5. *Suppose that $u \in L^2(\mathcal{P})$ and let τ be a finite stopping time. Then, the process $u\mathbf{1}_{[0,\tau]}$ also belongs to $L^2(\mathcal{P})$, and we have*

$$\int_0^\infty u_t \mathbf{1}_{[0,\tau]}(t) dB_t = \int_0^\tau u_t dB_t.$$

Proof. Suppose first that $u_t = F\mathbf{1}_{(a,b]}(t)$, where $0 \leq a < b$, $F \in L^2(\Omega, \mathcal{F}_a, P)$, and τ takes values in a finite set $\{0 = t_0 < t_1 < \cdots < t_n\}$. On one hand, we have

$$\int_0^\tau u_t dB_t = F(B_{b\wedge\tau} - B_{a\wedge\tau}).$$

On the other hand, the process $\mathbf{1}_{(0,\tau]}$ is simple because

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

and $\mathbf{1}_{\{\tau \geq t_{j+1}\}} = \mathbf{1}_{\{\tau \leq t_j\}^c} \in \mathcal{F}_{t_j}$. Therefore,

$$
\begin{aligned}
\int_0^\infty u_t \mathbf{1}_{(0,\tau]}(t) dB_t &= F \sum_{j=0}^{n-1} \mathbf{1}_{\{\tau \geq t_{j+1}\}} \int_0^\infty \mathbf{1}_{(a,b] \cap (t_j, t_{j+1}]}(t) dB_t \\
&= F \sum_{i=1}^{n} \mathbf{1}_{\{\tau = t_i\}} \int_0^\infty \mathbf{1}_{(a,b] \cap [0, t_i]}(t) dB_t \\
&= F(B_{b \wedge \tau} - B_{a \wedge \tau}).
\end{aligned}
$$

For a general finite stopping time τ, we approximate τ by the sequence of stopping times $\tau_n = \sum_{i=1}^{n2^n} \frac{i}{2^n} 1_{\{\frac{i-1}{2^n} \leq \tau < \frac{i}{2^n}\}}$, that satisfy $\tau_n \downarrow \tau$. Taking the limit as n tends to infinity we deduce the equality in the case of a simple process.

In the case of a general process $u \in L^2(\mathcal{P})$, we approximate u by simple processes $u^{(n)}$ in the norm of $L^2(\mathcal{P})$. The convergence

$$
\int_0^\tau u_t^{(n)} dB_t \stackrel{L^2(\Omega)}{\longrightarrow} \int_0^\tau u_t dB_t
$$

follows from the maximal inequality (12) when $T = \infty$. In fact,

$$
\begin{aligned}
\mathbb{E}\left(\left| \int_0^\tau u_t^{(n)} dB_t - \int_0^\tau u_t dB_t \right|^2 \right) &\leq \mathbb{E}\left(\sup_{t \geq 0} \left| \int_0^t [u_s^{(n)} - u_s] dB_s \right|^2 \right) \\
&\leq 4\mathbb{E}\left(\int_0^\infty |u_s^{(n)} - u_s|^2 ds \right).
\end{aligned}
$$

□

We observe that Proposition 4.5 also holds when $u \in L^2_\infty(\mathcal{P})$ and τ is bounded.

4.3 Integral of general processes

The stochastic integral can be defined for a class of processes larger than $L^2_\infty(\mathcal{P})$. Let $L^2_{loc}(\mathcal{P})$ the set of progressively measurable processes $u = (u_t)_{t \geq 0}$ such that for all $t \geq 0$,

$$
P\left(\int_0^t u_s^2 ds < \infty \right) = 1.
$$

Suppose that $u \in L^2_{loc}(\mathcal{P})$. For each $n \geq 1$, we define the stopping time

$$
T_n = \inf\left\{ t \geq 0 : \int_0^t u_s^2 ds = n \right\}, \tag{15}
$$

and the sequence of processes $u_t^{(n)} = u_t \mathbf{1}_{[0, T_n]}(t)$, which belong to $L^2(\mathcal{P})$.

Proposition 4.6. *Let $u \in L^2_{loc}(\mathcal{P})$. Then, there exists an adapted and continuous process $\left(\int_0^t u_s dB_s\right)_{t\geq 0}$ such that for any $n \geq 1$,*

$$\int_0^t u_s^{(n)} dB_s = \int_0^t u_s dB_s, \quad \text{on} \quad t \leq T_n.$$

Proof. If $n \leq m$, as a consequence of Proposition 4.5, we have that on the set $\{t \leq T_n\}$,

$$\int_0^t u_s^{(n)} dB_s = \int_0^t u_s^{(m)} dB_s. \tag{16}$$

We define

$$\int_0^t u_s dB_s = \int_0^t u_s^{(n)} dB_s, \quad \text{on} \quad t \leq T_n.$$

By (16), this definition does not depend on n and produces an adapted and continuous process because $T_n \uparrow \infty$. □

If $u \in L^2_{loc}(\mathcal{P})$, the process $(M_t)_{t\geq 0}$ defined by $M_t = \int_0^t u_s dB_s$ is a continuous local martingale; that is, there exists a sequence of stopping times $T_n \uparrow \infty$ (we can take the stopping times defined in (15)), such that for each $n \geq 1$, $(M_{t\wedge T_n})_{t\geq 0}$ is a martingale.

Instead of the isometry property, the stochastic integral of processes in $L^2_{loc}(\mathcal{P})$ has the following continuity property in probability.

Proposition 4.7. *Suppose that $u \in L^2_{loc}(\mathcal{P})$. Then, for all $K, \delta, T > 0$, we have*

$$P\left(\left|\int_0^T u_s dB_s\right| \geq K\right) \leq P\left(\int_0^T u_s^2 ds \geq \delta\right) + \frac{\delta}{K^2}.$$

Proof. Consider the stopping time defined by

$$\tau = \inf\left\{t \geq 0 : \int_0^t u_s^2 ds = \delta\right\},$$

with the convention that $\tau = T$ if $\int_0^T u_s^2 ds < \delta$. We have

$$\begin{aligned} P\left(\left|\int_0^T u_s dB_s\right| \geq K\right) &\leq P\left(\int_0^T u_s^2 ds \geq \delta\right) \\ &\quad + P\left(\left|\int_0^T u_s dB_s\right| \geq K, \int_0^T u_s^2 ds \leq \delta\right). \end{aligned}$$

On the other hand,

$$\begin{aligned} P\left(\left|\int_0^T u_s dB_s\right| \geq K, \int_0^T u_s^2 ds \leq \delta\right) &= P\left(\left|\int_0^T u_s dB_s\right| \geq K, \tau = T\right) \\ &\leq P\left(\left|\int_0^\tau u_s dB_s\right| \geq K\right) \\ &\leq \frac{1}{K^2}\mathbb{E}\left(\left|\int_0^\tau u_s dB_s\right|^2\right) = \frac{1}{K^2}\mathbb{E}\left(\int_0^\tau u_s^2 ds\right) \leq \frac{\delta}{K^2}, \end{aligned}$$

which concludes the proof. $\square$

As a consequence of the above proposition, if $u^{(n)}$ is a sequence of processes in $L^2_{loc}(\mathcal{P})$ that converges to $u \in L^2_{loc}(\mathcal{P})$ in probability, that is, for all $\epsilon > 0$ and $T > 0$,

$$\lim_{n\to\infty} P\left(\left|\int_0^T \{u_s^n - u_s\}^2 \, ds\right| > \epsilon\right) = 0,$$

then the sequence $\int_0^T u_s^n \, dB_s$ converges in probability to $\int_0^T u_s \, dB_s$, for all $T > 0$.

Moreover, we can show that for all $u \in L^2_{loc}(\mathcal{P})$ and $t \geq 0$, if $\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ denotes a subdivision of the interval $[0, t]$, we have that as $|\pi| \to 0$,

$$\sum_{j=0}^{n-1} \left(\int_{t_j}^{t_{j+1}} u_s dB_s\right)^2 \xrightarrow{P} \int_0^t u_s^2 ds. \tag{17}$$

4.4 Itô's formula

Itô's stochastic integral does not follow the chain rule of classical calculus. For instance, in Example 4.4, we have shown that

$$\int_0^t B_s dB_s = \frac{1}{2} B_t^2 - \frac{t}{2}, \tag{18}$$

whereas if x_t is a differentiable function such that $x_0 = 0$,

$$\int_0^t x_s dx_s = \int_0^t x_s x_s' ds = \frac{1}{2} x_t^2.$$

Equation (18) can also be written as

$$B_t^2 = \int_0^t 2B_s dB_s + t.$$

That is, the stochastic process B_t^2 can be expressed as the sum of the indefinite stochastic integral $\int_0^t 2B_s dB_s$ plus a differentiable function. More generally, Itô's formula below shows that any process of the form $f(B_t)$, where f is twice continuously differentiable, can be expressed as the sum of an indefinite Itô's integral, plus a process with differentiable trajectories. This leads to the definition of Itô process.

Denote by $L^1_{loc}(\mathcal{P})$ the space of progressively measurable processes $v = (v_t)_{t\geq 0}$ such that for all $t > 0$,

$$P\left(\int_0^t |v_s| \, ds < \infty\right) = 1.$$

Definition 4.6. *A continuous and adapted stochastic process $(X_t)_{t\geq 0}$ is called an Itô process if*

$$X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds, \tag{19}$$

where $u \in L^2_{loc}(\mathcal{P})$, $v \in L^1_{loc}(\mathcal{P})$, and X_0 is an $\mathcal{F}_0$-measurable random variable.

Notice that $\mathcal{F}_0$-measurable random variables are constant a.s.

We say that a function $f\colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is of class $C^{1,2}$ if it is twice differentiable with respect to the variable $x \in \mathbb{R}$ and once differentiable with respect to $t \in \mathbb{R}_+$, with continuous partial derivatives $\frac{\partial f}{\partial x}$, $\frac{\partial^2 f}{\partial x^2}$, and $\frac{\partial f}{\partial t}$.

Theorem 4.7 (Itô's formula)**.** *Let $f\colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a function of class $C^{1,2}$. Suppose that X is an Itô process of the form* (19)*. Then, the process $Y_t = f(t, X_t)$ is again an Itô process with the representation*

$$\begin{aligned} Y_t = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial t}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)u_s dB_s \\ &+ \int_0^t \frac{\partial f}{\partial x}(s, X_s)v_s ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)u_s^2 ds, \end{aligned} \tag{20}$$

which holds a.s., for all $t \geq 0$.

Remark 4.8. *1. Notice that all the integrals in (20) are well-defined, that is, the processes $\frac{\partial f}{\partial t}(s, X_s)$, $\frac{\partial f}{\partial x}(s, X_s)v_s$, and $\frac{\partial^2 f}{\partial x^2}(s, X_s)u_s^2$ belong to $L^1_{loc}(\mathcal{P})$, and $\frac{\partial f}{\partial x}(s, X_s)u_s$ belongs to $L^2_{loc}(\mathcal{P})$.*

2. We set $\langle X\rangle_t = \int_0^t u_s^2 ds$, that is, $\langle X\rangle_t$ is the quadratic variation of the local martingale $\int_0^t u_s dB_s$ in the sense of (17)*. Then (20) can be written as*

$$Y_t = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)d\langle X\rangle_s.$$

3. In differential notation, we write

$$dX_t = u_t dB_t + v_t dt,$$

and Itô's formula can be written as

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \frac{\partial f}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\left(dX_t\right)^2,$$

where $(dX_t)^2$ is computed using the product rule

$\times$	dB_t	dt
dB_t	dt	0
dt	0	0

Notice also that $(dX_t)^2 = u_t^2 dt = d\langle X\rangle_t$.

4. In the particular case $u_t = 1$, $v_t = 0$, and $X_0 = 0$, the process X is the Brownian motion B, and Itô's formula has the following simple version

$$f(t, B_t) = f(0, 0) + \int_0^t \frac{\partial f}{\partial x}(s, B_s)dB_s + \int_0^t \frac{\partial f}{\partial t}(s, B_s)ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s)ds.$$

Proof of Theorem 4.7. We will do the proof only in the case that $v_t = 0$, that is,

$$X_t = X_0 + \int_0^t u_s dB_s,$$

and f does not depend on t. The extension of the proof to the general case is easy and we omit the details. We claim that

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) u_s dB_s + \frac{1}{2}\int_0^t f''(X_s) u_s^2 ds.$$

By a localization argument, we may assume that $f \in C_b^2(\mathbb{R})$, $\int_0^\infty u_s^2 ds \leq N$, and $\sup_{t\geq 0} |X_t| \leq N$, for some integer $N \geq 1$, such that $|X_0| \leq N$. In fact, consider the sequence of stopping times

$$T_N = \inf\left\{ t \geq 0 : \int_0^t u_s^2 ds \geq N, \quad \text{or} \quad |X_t| \geq N \right\}.$$

Let $f_N \colon \mathbb{R} \to \mathbb{R}$ be a function in $C_0^2(\mathbb{R})$ such that $f(x) = f_N(x)$ for $|x| \leq N$. Then, if $u_t^{(N)} = u_t \mathbf{1}_{[0,T_N]}(t)$, and

$$X_t^{(N)} = X_0 + \int_0^t u_s^{(N)} dB_s,$$

by Itô's formula we would get

$$f_N(X_t^{(N)}) = f_N(X_0) + \int_0^t f_N'(X_s^{(N)}) u_s^{(N)} dB_s + \frac{1}{2}\int_0^t f_N''(X_s^{(N)}) (u_s^{(N)})^2 ds.$$

By Proposition 4.6,

$$X_t^{(N)} = X_0 + \int_0^{T_N \wedge t} u_s dB_s = X_{T_N \wedge t},$$

and

$$f(X_{T_N\wedge t}) = f(X_0) + \int_0^{T_N \wedge t} f'(X_s) u_s dB_s + \frac{1}{2}\int_0^{T_N\wedge t} f''(X_s) u_s^2 ds.$$

Then, we let $N \to \infty$ to get the result.

Consider the uniform partition $0 = t_0 < t_1 < \cdots < t_n = t$, where $t_i = \frac{it}{n}$. We can write, using Taylor's formula

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{i=0}^{n-1} \left[f(X_{t_{i+1}}) - f(X_{t_i}) \right] \\ &= \sum_{i=0}^{n-1} f'(X_{t_i})(X_{t_{i+1}} - X_{t_i}) + \frac{1}{2}\sum_{i=0}^{n-1} f''(\widetilde{X_i})(X_{t_{i+1}} - X_{t_i})^2, \end{aligned}$$

where $\widetilde{X_i}$ is a random point between X_{t_i} and $X_{t_{i+1}}$.

It is an easy exercise to show that

$$\sum_{i=0}^{n-1} f'(X_{t_i})(X_{t_{i+1}} - X_{t_i}) \overset{L^2(\Omega)}{\longrightarrow} \int_0^t f'(X_s) u_s dB_s, \tag{21}$$

as n tends to infinity.

For the second term, we write

$$\begin{aligned}
&\int_0^t f''(X_s)u_s^2 ds - \sum_{i=0}^{n-1} f''(\widetilde{X_i})(X_{t_{i+1}} - X_{t_i})^2 \\
&\qquad = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \left[f''(X_s) - f''(X_{t_i})\right] u_s^2 ds \\
&\qquad\qquad + \sum_{i=0}^{n-1} f''(X_{t_i}) \left(\int_{t_i}^{t_{i+1}} u_s^2 ds - \left(\int_{t_i}^{t_{i+1}} u_s dB_s \right)^2 \right) \\
&\qquad\qquad + \sum_{i=0}^{n-1} [f''(X_{t_i}) - f''(\widetilde{X}_i)] \left(\int_{t_i}^{t_{i+1}} u_s dB_s \right)^2 \\
&\qquad =: A_1^n + A_2^n + A_3^n.
\end{aligned}$$

Then, it suffices to show that each term A_i^n converges to zero in probability.

We have

$$|A_1^n| \leq \sup_{|s-r|\leq t/n} |f''(X_s) - f''(X_r)| \int_0^t u_s^2 ds,$$

and

$$|A_3^n| \leq \sup_{0\leq i\leq n-1} |f''(X_{t_i}) - f''(\widetilde{X}_i)| \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} u_s dB_s \right)^2.$$

Taking into account that f'' and X are continuous together with Proposition 4.4, we obtain that both expressions converge to zero in probability as n tends to infinity.

Using that the sequence $\xi_i = \int_{t_i}^{t_{i+1}} u_s^2 ds - \left(\int_{t_i}^{t_{i+1}} u_s dB_s \right)^2$ is bounded, and satisfies

$$\mathbb{E}[\xi_i | \mathcal{F}_{t_i}] = 0,$$

we obtain

$$\begin{aligned}
\mathbb{E}[(A_n^2)^2] &= \sum_{i=0}^{n-1} \mathbb{E}[f''(X_{t_i})^2 \xi_i^2] \le \|f''\|_\infty^2 \sum_{i=0}^{n-1} \mathbb{E}[\xi_i^2] \\
&\le 2\|f''\|_\infty^2 \sum_{i=0}^{n-1} \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} u_s^2 ds\right)^2 + \left(\int_{t_i}^{t_{i+1}} u_s dB_s\right)^4\right] \\
&\le 2\|f''\|_\infty^2 \mathbb{E}\left[N \sup_i \int_{t_i}^{t_{i+1}} u_s^2 ds\right. \\
&\qquad \left. + \sup_i |X_{t_{i+1}} - X_{t_i}|^2 \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} u_s dB_s\right)^2\right],
\end{aligned}$$

which converges to zero as $n \to \infty$. This completes the proof. □

Example 4.9. *1. If $f(x) = x^2$ and $X_t = B_t$, we obtain*

$$B_t^2 = 2\int_0^t B_s dB_s + t,$$

because $f'(x) = 2x$ and $f''(x) = 2$.

2. If $f(x) = x^3$ and $X_t = B_t$, we obtain

$$B_t^3 = 3\int_0^t B_s^2 dB_s + 3\int_0^t B_s ds,$$

because $f'(x) = 3x^2$ and $f''(x) = 6x$. More generally, if $n \ge 2$ is a natural number,

$$B_t^n = n\int_0^t B_s^{n-1} dB_s + \frac{n(n-1)}{2}\int_0^t B_s^{n-2} ds.$$

3. If $f(t,x) = e^{ax - \frac{a^2}{2}t}$, $X_t = B_t$, and $Y_t = e^{aB_t - \frac{a^2}{2}t}$, we obtain

$$Y_t = 1 + a\int_0^t Y_s dB_s,$$

because

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = 0. \tag{22}$$

In particular, the solution to the stochastic differential equation

$$dY_t = aY_t dB_t, \quad Y_0 = 1,$$

is not $Y_t = e^{aB_t}$, but $Y_t = e^{aB_t - \frac{a^2}{2}t}$.

4. If a function f of class $C^{1,2}$ satisfies equality (22)*, then,*

$$f(t, B_t) = f(0,0) + \int_0^t \frac{\partial f}{\partial x}(s, B_s) dB_s.$$

This implies that $f(t, B_t)$ is a continuous local martingale. The process is a square integrable martingale if

$$\mathbb{E}\left[\int_0^t \left(\frac{\partial f}{\partial x}(s, B_s)\right)^2 ds\right] < \infty,$$

for all $t \geq 0$.

4.5 Multidimensional Itô's formula

Itô's formula can be extended to a multidimensional setting as follows. Suppose that $B = (B_t)_{t\geq 0}$, $B_t = (B_t^1, B_t^2, \ldots, B_t^m)$, is an m-dimensional Brownian motion.

Definition 4.10. *An n-dimensional continuous and adapted process $(X_t)_{t\geq 0}$ is called a multidimensional Itô process if*

$$X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds, \tag{23}$$

where $v = (v_t)_{t\geq 0}$ is an n-dimensional process and $u = (u_t)_{t\geq 0}$ is a process with values in the set of $n \times m$ matrices, and we assume that the components of u belong to $L^2_{loc}(\mathcal{P})$ and those of v belong to $L^1_{loc}(\mathcal{P})$.

We say that a function $f\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is of class $C^{1,2}$ if the partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$, and $\frac{\partial f}{\partial x_i \partial x_j}$, $1 \leq i, j \leq n$, exist and are continuous.

Theorem 4.11 (Multidimensional Itô's formula)**.** *Suppose that X is a multidimensional Itô process of the form* (23)*. Let $f\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ be a function of class $C^{1,2}$. Then, the process $Y_t = f(t, X_t)$ is again a multidimensional Itô process with the representation*

$$\begin{aligned} Y_t = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s, X_s) dX_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X_s)(u_s u_s^T)_{ij} ds \end{aligned}$$

almost surely, for all $t \geq 0$.

Proof. The proof follows along the same lines as the one-dimensional Itô's formula (Theorem 4.7). □

Remark 4.12. *1. Set*

$$\langle X^i, X^j\rangle_t = \int_0^t (u_s u_s^{\mathrm{T}})_{ij} ds = \int_0^t \sum_{k=1}^m (u_s)_{ik}(u_s)_{jk} ds.$$

That is, $\langle X^i, X^j\rangle_t$ is the quadratic covariation of the local martingales $\sum_{k=1}^m \int_0^t (u_s)_{ik} dB_s^k$ and $\sum_{k=1}^m \int_0^t (u_s)_{jk} dB_s^k$. Then, Itô's formula can be written in the form

$$\begin{aligned} Y_t = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial t}(s, X_s) ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s, X_s) dX_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X_s) d\langle X^i, X^j\rangle_s. \end{aligned}$$

2. In differential notation, multidimensional Itô's formula can be written as

$$dY_t = \frac{\partial f}{\partial t}(t, X_t) dt + \nabla f(t, X_t) dX_t + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X_t) dX_t^i dX_t^j.$$

The product of differentials $dX_t^i dX_t^j$ is computed by means of the product rules: $dB_t^i dt = 0$, $(dt)^2 = 0$ and

$$dB_t^i dB_t^j = \begin{cases} 0 & \text{if } i \neq j, \\ dt & \text{if } i = j. \end{cases}$$

3. As an application of the multidimensional Itô's formula we can deduce the following integration-by-parts formula. Suppose that $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are one-dimensional Itô processes. Then,

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t d\langle X, Y\rangle_s. \tag{24}$$

4.6 Stratonovich integral

The Itô's integral has been defined as the limit in $L^2(\Omega)$ or in probability of forward Riemann sums. If we consider symmetric Riemann sums defined by taking the value of the integrand in the middle point of each interval, we obtain a different type of integral, called Stratonovich integral. In the next proposition we define the Stratonovich integral of an Itô process with respect to another one.

Proposition 4.8. *Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be two Itô processes of the form*

$$X_t = X_0 + \int_0^t u_s^X dB_s + \int_0^t v_s^X ds,$$

and

$$Y_t = Y_0 + \int_0^t u_s^Y dB_s + \int_0^t v_s^Y ds,$$

where $u^X, u^Y \in L^2_{loc}(\mathcal{P})$ *and* $v^X, v^Y \in L^1_{loc}(\mathcal{P})$. *Let* $\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ *be a partition of the interval* $[0,t]$. *Then, as* $|\pi| \to 0$,

$$\sum_{j=0}^{n-1} \frac{1}{2}(Y_{t_j} + Y_{t_{j+1}})(X_{t_{j+1}} - X_{t_j}) \xrightarrow{P} \int_0^t Y_s dX_s + \frac{1}{2}\langle X, Y\rangle_t.$$

This limit is called the Stratonovich integral of Y *with respect to* X *and is denoted by* $\int_0^t Y_s \circ dX_s$.

Proof. The result follows easily from the decomposition

$$\frac{1}{2}(Y_{t_j} + Y_{t_{j+1}})(X_{t_{j+1}} - X_{t_j}) = Y_{t_j}(X_{t_{j+1}} - X_{t_j}) + \frac{1}{2}(Y_{t_{j+1}} - Y_{t_j})(X_{t_{j+1}} - X_{t_j}),$$

and the following property: as $|\pi| \to 0$,

$$\sum_{j=0}^{n-1}(Y_{t_{j+1}} - Y_{t_j})(X_{t_{j+1}} - X_{t_j}) \xrightarrow{P} \int_0^t u_s^X u_s^Y ds.$$

□

The Stratonovich integral follows the rules of the classical calculus. That is, if $f \in C^3(\mathbb{R})$,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \circ dX_s.$$

Indeed, by Itô's formula (Theorem 4.7) and Proposition 4.8,

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2}\int_0^t f''(X_s) d\langle X\rangle_s \\ &= f(X_0) + \int_0^t f'(X_s) \circ dX_s - \frac{1}{2}\langle f'(X), X\rangle_t + \frac{1}{2}\int_0^t f''(X_s) d\langle X\rangle_s. \end{aligned}$$

On the other hand, again by Itô's formula, $\langle f'(X), X\rangle_t$ is the quadratic covariation between the martingales $\int_0^t f''(X_s) u_s dB_s$ and $\int_0^t u_s dB_s$, which coincides with $\int_0^t f''(X_s) d\langle X\rangle_s$.

4.7 Integral representation theorem

Consider a process u in the space $L^2_\infty(\mathcal{P})$. We know that the indefinite stochastic integral

$$X_t = \int_0^t u_s dB_s$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. The aim of this subsection is to show that any square integrable martingale is of this form. We start with the integral representation of square integrable random variables.

Theorem 4.13 (Integral representation theorem)**.** *Fix $T > 0$ and let $F \in L^2(\Omega, \mathcal{F}_T, P)$. Then, there exists a unique process u in the space $L^2_T(\mathcal{P})$ such that*

$$F = \mathbb{E}(F) + \int_0^T u_s dB_s.$$

Proof. Suppose first that

$$F = \exp\left(\int_0^T h_s dB_s - \frac{1}{2}\int_0^T h_s^2 ds\right), \tag{25}$$

where $h \in L^2([0,T])$. Define

$$Y_t = \exp\left(\int_0^t h_s dB_s - \frac{1}{2}\int_0^t h_s^2 ds\right), \quad 0 \le t \le T.$$

By Itô's formula applied to the function $f(x) = e^x$ and the process $X_t = \int_0^t h_s dB_s - \frac{1}{2}\int_0^t h_s^2 ds$, we obtain for all $t \in [0,T]$,

$$Y_t = 1 + \int_0^t Y_s h_s dB_s.$$

Hence,

$$F = 1 + \int_0^T Y_s h_s dB_s,$$

and we get the desired representation because $\mathbb{E}(F) = 1$ and $Yh \in L^2_\infty(\mathcal{P})$. By linearity, the representation holds for linear combinations of exponentials of the form (25).

In the general case, any random variable $F \in L^2(\Omega, \mathcal{F}_T, P)$ can be approximated in $L^2(\Omega)$ by a sequence F_n of linear combinations of exponentials of the form (25). Then, we have

$$F_n = \mathbb{E}(F_n) + \int_0^T u_s^{(n)} dB_s.$$

By the isometry of the stochastic integral,

$$\begin{aligned} \mathbb{E}\left[(F_n - F_m)^2\right] &\ge \operatorname{Var}(F_n - F_m) = \mathbb{E}\left[\left(\int_0^T \left(u_s^{(n)} - u_s^{(m)}\right) dB_s\right)^2\right] \\ &= \mathbb{E}\left[\int_0^T \left(u_s^{(n)} - u_s^{(m)}\right)^2 ds\right]. \end{aligned}$$

Hence, $u^{(n)}$ is a Cauchy sequence in $L^2_T(\mathcal{P})$ and it converges to a process u in $L^2_T(\mathcal{P})$. Applying again the isometry property, and taking into account that $\mathbb{E}(F_n)$ converges to $\mathbb{E}(F)$, we obtain

$$\begin{aligned} F &= \lim_{n\to\infty} F_n = \lim_{n\to\infty}\left(\mathbb{E}(F_n) + \int_0^T u_s^{(n)} dB_s\right) \\ &= \mathbb{E}(F) + \int_0^T u_s dB_s. \end{aligned}$$

Finally, uniqueness follows also from the isometry property. Indeed, suppose that $u^{(1)}$ and $u^{(2)}$ are processes in $L_T^2(\mathcal{P})$ such that

$$F = \mathbb{E}(F) + \int_0^T u_s^{(1)} dB_s = \mathbb{E}(F) + \int_0^T u_s^{(2)} dB_s.$$

Then,

$$0 = \mathbb{E}\left[\left(\int_0^T \left(u_s^{(1)} - u_s^{(2)}\right) dB_s\right)^2\right] = \mathbb{E}\left[\int_0^T \left(u_s^{(1)} - u_s^{(2)}\right)^2 ds\right],$$

and, hence, $u_s^{(1)}(\omega) = u_s^{(2)}(\omega)$ for almost all $(\omega, s) \in \Omega \times [0, T]$. □

Corollary 4.14 (Martingale representation theorem)**.** *Suppose that $M = (M_t)_{t \geq 0}$ is a square integrable martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$. Then, there exists a unique process $u \in L_\infty^2(\mathcal{P})$ such that*

$$M_t = \mathbb{E}(M_0) + \int_0^t u_s dB_s.$$

In particular, M has a continuous version.

Example 4.15. *We want to find the integral representation of $F = B_T^3$. By Itô's formula,*

$$B_T^3 = \int_0^T 3B_t^2 dB_t + 3\int_0^T B_t dt,$$

and using the integrating by parts formula (24) yields

$$\int_0^T B_t dt = TB_T - \int_0^T t dB_t = \int_0^T (T-t) dB_t.$$

So, we obtain the representation

$$B_T^3 = \int_0^T 3\left[B_t^2 + (T-t)\right] dB_t.$$

Exercises

1. Use Itô's formula in order to compute $\mathbb{E}(B_t^4)$. Do it again for $\mathbb{E}(B_t^6)$. Write a recurrence formula in order to compute $\mathbb{E}(B_t^n)$ in terms of $\mathbb{E}(B_t^{n-2})$.

2. Show that for any $x \in \mathbb{R}$, the process

$$X_t = \int_0^t \operatorname{sign}(B_s - x) dB_s, \quad t \geq 0,$$

   is a Brownian motion.

   Hint: Apply Itô's formula to the process $e^{i\lambda(X_t - X_s)}$.

3. Let B be a Brownian motion. Using Itô's formula check whether the following processes are martingales:

   (a) $X_t^{(1)} = B_t^3 - 3tB_t$.

   (b) $X_t^{(2)} = t^2 B_t - 2\int_0^t sB_s ds$.

   (c) $X_t^{(3)} = e^{t/2}\cos B_t$.

   (d) $X_t^{(4)} = e^{t/2}\sin B_t$.

   (e) $X_t^{(5)} = (B_t + t)\exp(-B_t - \frac{1}{2}t)$.

4. Let B_1 and B_2 two independent Brownian motions. Show that $B_1(t)B_2(t)$ is a martingale.

5. Let B be an m-dimensional Brownian motion. Consider a partition $\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ of the interval $[0, t]$. Show that if $i \neq j$, as $|\pi| \to 0$,

$$\sum_{k=0}^{n-1} (B^i_{t_{k+1}} - B^i_{t_k})(B^j_{t_{k+1}} - B^j_{t_k}) \xrightarrow{L^2(\Omega)} 0.$$

6. Find the stochastic integral representation on the time interval $[0, T]$ of the following random variables:

$$B_T,\ B_T^2,\ e^{B_T},\ \int_0^T B_t dt,\ B_T^3,\ \sin B_T,\ \int_0^T tB_t^2 dt.$$

7. Let $p(t, x) = 1/\sqrt{1-t}\exp(-x^2/2(1-t))$, for $0 \le t < 1$, $x \in \mathbb{R}$, and $p(1, x) = 0$. Define $M_t = p(t, B_t)$, where $(B_t)_{t\in[0,1]}$ is a Brownian motion.

   (a) Show that for each $0 \le t < 1$,

$$M_t = M_0 + \int_0^t \frac{\partial p}{\partial x}(s, B_s) dB_s.$$

   (b) Set $H_t = \frac{\partial p}{\partial x}(t, B_t)$. Show that $\int_0^1 H_t^2 dt < \infty$ almost surely, but

$$\mathbb{E}\left(\int_0^1 H_t^2 dt\right) = \infty.$$

5 Derivative and divergence operators

In this section we introduce the derivative and divergence operators in the context of an isonormal Gaussian process on a Hilbert space H.

5.1 Derivative operator

Consider an isonormal Gaussian process $B = \{B(h), h \in H\}$ on a separable Hilbert space H. We will introduce the derivative of a random variable $F\colon \Omega \to \mathbb{R}$ as a random element that takes values in H, that is, $DF \in L^2(\Omega; H)$.

The basic example will be $H = L^2(\mathbb{R}_+)$. In this case $B_t = B(\mathbf{1}_{[0,t]})$, $t \geq 0$, is a Brownian motion and $B(h) = \int_0^\infty h(t)dB_t$ is the Wiener integral of h. The derivative of F, $(D_tF)_{t\geq 0}$ is a stochastic process.

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

$$F = f(B(h_1), \dots, B(h_n)), \tag{26}$$

where $f \in C_p^\infty(\mathbb{R}^n)$ (f and all its partial derivatives have polynomial growth order) and $h_i \in H$.

Definition 5.1. *If $F \in \mathcal{S}$ is a smooth and cylindrical random variable of the form* (26), *the derivative DF is the H-valued random variable defined by*

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

For instance, $D(B(h)) = h$.

Remark 5.2. *In the Brownian motion case $(B_t)_{t\in[0,T]}$, defined on the Wiener space $\Omega = C([0,T])$, 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 $h \in H = L^2([0,T])$. Then, for any $h \in H$, $\langle DF, h\rangle_H$ 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 $D_t F = \mathbf{1}_{[0,t_1]}(t)$.

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 H-valued random elements of the form

$$u_t = \sum_{j=1}^{n} F_j h_j, \tag{27}$$

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

Definition 5.3. *We define the divergence of an element u of the form (27) 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 5.1. *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 $F = f(B(h_1) \dots, B(h_n))$ and

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

where $h_1, \dots, h_n$ are orthonormal elements in H. In this case, the duality relationship reduces to the finite-dimensional case proved in Proposition 2.4. $\square$

We will make use of the notation $D_h F = \langle DF, h \rangle_H$ 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 5.2. *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*

$$\mathbb{E}(\delta(u)\delta(v)) = \mathbb{E}(\langle u, v \rangle_H) + \mathbb{E}\left(\sum_{i,j=1}^{\infty} D_{e_i} \langle u, e_j \rangle_H D_{e_j} \langle v, e_i \rangle_H \right), \tag{28}$$

$$D_h(\delta(u)) = \delta(D_h u) + \langle h, u \rangle_H, \tag{29}$$

$$\delta(Fu) = F\delta(u) - \langle DF, u \rangle_H. \tag{30}$$

In the Brownian motion case, Property (28) 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 5.2. We first show property (29). 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

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

To show property (28), using the duality formula (Proposition 5.1) and property (29), 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 (30) we choose a smooth random variable $G \in \mathcal{S}$ and write, using the duality relationship (Proposition 5.1),

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

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

5.2 Sobolev spaces

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

Proposition 5.3. *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 5.1), 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 $p \geq 1$. $\square$

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

$$\|F\|_{1,p} = \left(\mathbb{E}(|F|^p) + \mathbb{E}\left(\|DF\|_H^p\right)\right)^{1/p}.$$

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

$$\langle F, G \rangle_{1,2} = \mathbb{E}(FG) + \mathbb{E}\left(\langle DF, DG \rangle_H\right).$$

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 5.4. *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 $|\varphi(x)| \leq C'(1 + |x|^{\alpha+1})$, 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. $\square$

We next define the domain of the divergence operator.

Definition 5.4. *The domain of the divergence operator* $\operatorname{Dom}\delta$ *in $L^2(\Omega)$ is the set of random elements $u \in L^2(\Omega; H)$ 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 5.2 can be extended to random variables in suitable Sobolev spaces. Property (28) 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

$$\mathbb{E}(\delta(u)^2) \leq \mathbb{E}\left(\|u\|_H^2\right) + \mathbb{E}\left(\|Du\|_{H\otimes H}^2\right) = \|u\|_{1,2,H}^2.$$

Property (29) holds if $u \in \mathbb{D}^{1,2}(H)$ and $D_h u \in \operatorname{Dom}\delta$. Finally, property (30) 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 $H^{\otimes k}$-valued random element obtained by iteration:

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

For any $p \geq 1$, 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

$$\|F\|_{k,p} = \left(\mathbb{E}(|F|^p) + \mathbb{E}\left(\sum_{j=1}^{k} \|D^j F\|_{H^{\otimes j}}^p\right)\right)^{1/p}.$$

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

These spaces satisfy that for any $q \geq p \geq 2$, and $\ell \geq k$, $\mathbb{D}^{\ell,q} \subset \mathbb{D}^{k,p}$. We also have the following Hölder's inequality for the $\|\cdot\|_{k,p}$-seminorms.

Proposition 5.5. *Let $p, q, r \geq 2$ such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Let $F \in \mathbb{D}^{k,p}$ and $G \in \mathbb{D}^{k,q}$. Then, FG belongs to $\mathbb{D}^{k,r}$, and*

$$\|FG\|_{k,r} \leq c_{p,q,r} \|F\|_{k,p} \|G\|_{k,q}.$$

Proof. The result follows from the Leibnitz rule and Hölder's inequality. $\square$

Remark 5.5. *In the Brownian motion case, $H = L^2(\mathbb{R}_+)$ and we identify the Hilbert space $L^2(\Omega; H)$ with $L^2(\Omega \times \mathbb{R}_+)$. In this form an element $u \in L^2(\Omega; H) = L^2(\Omega \times \mathbb{R}_+)$ is a stochastic process $u = (u_t)_{t\geq 0}$. Moreover $L^2(\Omega; H^{\otimes k})$ can be indetified with $L^2(\Omega \times \mathbb{R}_+^k)$, and $D^k F$ is a k-parameter process denoted by*

$$D^k F = (D^k_{t_1,\dots,t_k} F)_{t_1,\dots,t_k \geq 0}.$$

5.3 The divergence as a stochastic integral

Consider the case of the Brownian motion $B = (B_t)_{t\geq 0}$. The derivative operator is local 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 5.6. *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 $(t, \omega) \in \Omega \times [a,b]^c$.*

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. Recall that $(\mathcal{F}_t)_{t \geq 0}$ is the filtration generated by the Brownian motion.

Theorem 5.7. *Any progressively measurable process u in $L^2(\mathcal{P}) \subset L^2(\Omega \times \mathbb{R}_+)$ belongs to* $\operatorname{Dom} \delta$ *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 (30),

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

taking into account that $D_t \phi_j = 0$ if $t > t_j$ by Lemma 5.6. 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 (see [18]).

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 (28) 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 (29), we obtain

$$D_t \left(\int_0^\infty u_s dB_s \right) = u_t + \int_t^\infty D_t u_s dB_s, \tag{31}$$

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

Exercises

1. Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be a function such that

$$|\varphi(x) - \varphi(y)| \leq K|x - y|,$$

for all $x, y \in \mathbb{R}$. Let F be a random variable in the space $\mathbb{D}^{1,2}$. Show that $\varphi(F)$ belongs to $\mathbb{D}^{1,2}$, and there exists a random variable G bounded by K such that

$$D(\varphi(F)) = GDF.$$

Moreover, show that when the law of the random variable F is absolutely continuous with respect to the Lebesgue measure, then $G = \varphi'(F)$.

Hint: Approximate φ by the sequence $\varphi_n = \varphi * \alpha_n$, where α_n is an approximation of the identity, apply the chain rule in Proposition 5.4 to the random variable F and the function φ_n, and use Corollary 6.4 to conclude.

2. Let $X = (X_t)_{t\in[0,1]}$ be a continuous centered Gaussian process. Assume that a.s. X attains its maximum on a unique random point T at $[0, 1]$. Show that the random variable $M = \sup_{t\in[0,1]} X_t$ belongs to the space $\mathbb{D}^{1,2}$, and $D_t M = \mathbf{1}_{[0,T]}(t)$.

Hint: Approximate the supremum of X by the maximum on a finite set, and use the chain rule for Lipschitz functions proved in the previous exercise, and Corollary 6.4.

3. Show that, in the Brownian motion case, for any $u, v \in \mathbb{D}^{1,2}(H)$,

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

4. Show that for any $u \in \mathbb{D}^{1,2}(H)$ and $h \in H$ such that $D_h u \in \operatorname{Dom}\delta$,

$$D_h(\delta(u)) = \delta(D_h u) + \langle h, u\rangle_H.$$

5. Show that for any $F \in \mathbb{D}^{1,2}$ and $u \in \operatorname{Dom}\delta$ such that $Fu \in L^2(\Omega; H)$,

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

provided that the right-hand side is square integrable.

6. Show that for any $p \geq 1$, the operator D^k is closable from $L^p(\Omega)$ into $L^p(\Omega; H^{\otimes k})$.

7. Prove the following Leibniz rule for the iterated derivative: for any random variables $F, G \in \mathcal{S}$ and any $k \geq 2$,

$$D^k(FG) = \sum_{i=0}^{k} \binom{k}{i} D^i F D^{k-i} G.$$

6 Multiple stochastic integrals. Wiener chaos

In this section we present the Wiener chaos expansion, which provides an orthogonal decomposition of square integrable random variables in the Wiener space in terms of multiple stochastic integrals. We then compute the derivative and the divergence operators on the Wiener chaos expansion.

6.1 Multiple stochastic integrals

Suppose 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 $f\colon \mathbb{R}^n_+ \to \mathbb{R}$. 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 6.1. *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 $g \in L^2(\mathbb{R}^m_+)$,

$$\mathbb{E}(I_n(f) I_m(g)) = \begin{cases} 0 & \text{if } n \neq m, \\ n! \langle \tilde{f}, \tilde{g} \rangle_{L^2(\mathbb{R}^n_+)} & \text{if } n = m. \end{cases} \tag{32}$$

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:

$$I_n(f)I_m(g) = \sum_{r=0}^{n \wedge m} r! \binom{n}{r} \binom{m}{r} I_{n+m-2r}(f \otimes_r g). \tag{33}$$

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

Proposition 6.1. *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 $g^{\otimes n}(t_1, \ldots, t_n) = g(t_1) \cdots g(t_n)$.

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 $1, \ldots, n$. Using the product rule (33), the induction hypothesis, and the recursive relation for the Hermite polynomials, we get

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

which concludes the proof. □

The next result is the *Wiener chaos expansion.*

Theorem 6.2. *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 6.2 can be reformulated by saying that we have the orthogonal decomposition

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

Proof of Theorem 6.2. 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 $G = 0$. □

6.2 Derivative operator on the Wiener chaos

Let us compute the derivative of a multiple stochastic integral.

Proposition 6.2. *Let $f \in L^2_s(\mathbb{R}^n_+)$. 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 6.1 and the properties of Hermite polynomials, we have

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

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

Moreover, applying (32), we have

$$\begin{aligned} \mathbb{E}\left(\int_{\mathbb{R}_+} (D_t I_n(f))^2 dt \right) &= n^2 \int_{\mathbb{R}_+} \mathbb{E}(I_{n-1}(f(\cdot, t))^2) dt \\ &= n^2 (n-1)! \int_{\mathbb{R}_+} \|f(\cdot, t)\|^2_{L^2(\mathbb{R}^{n-1}_+)} dt \\ &= n n! \|f\|^2_{L^2(\mathbb{R}^n_+)} \\ &= n \mathbb{E}(I_n(f)^2). \end{aligned} \tag{34}$$

As a consequence of Proposition 6.2 and (34), we deduce the following result.

Proposition 6.3. *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*

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

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

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

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:

$$f_n = \frac{1}{n!}\mathbb{E}(D^n F). \tag{35}$$

Example 6.3. *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 6.3 implies the following characterization of the space $\mathbb{D}^{1,2}$.

Proposition 6.4. *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:*

$$\mathbb{E}(\langle u, h\rangle_H G) = \mathbb{E}(F\delta(Gh)). \tag{36}$$

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 5.1) and Proposition 6.2, 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 (36), 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 6.3 allows us to conclude the proof. □

Corollary 6.4. *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 6.4, it suffices to show that, for all $G \in \mathcal{S}$ and $h \in H$,

$$\mathbb{E}(\langle \alpha, h\rangle_H G) = \mathbb{E}(F\delta(Gh)). \tag{37}$$

By the duality formula (Proposition 5.1), 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 (37), 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 6.5. *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 5.4, 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 6.3, we deduce that $\mathbf{1}_A = P(A)$, which completes the proof. □

6.3 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 6.6. *The process u belongs to the domain of δ if and only if the series*

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \tag{38}$$

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

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

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 (38) converges in $L^2(\Omega)$ and its sum is equal to $\delta(u)$. The converse can be proved by similar arguments. □

6.4 Generalized multiple stochastic integrals

Let $B = \{B(h), h \in H\}$ be an isomormal Gaussian process on a Hilbert space H defined on a probability space $L^2(\Omega, \mathcal{F}, P)$. Assume that $\mathcal{F}$ is generated by B. For each $n \geq 1$ we will denote by $\mathcal{H}_n$ the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_n(B(h)), h \in H, \|h\|_H = 1\}$. $\mathcal{H}_0$ will be the set of constants. For $n = 1$, $\mathcal{H}_1$ coincide with the set B. For $n \neq m$ the spaces $\mathcal{H}_n$ and $\mathcal{H}_m$ are orthogonal. The space $\mathcal{H}_n$ is called the Wiener chaos of order n. In view of Proposition 6.1, this definition generalizes the Wiener chaos defined previously for the Brownian motion.

As in the case of Brownian motion, one can show the following orthogonal decomposition:

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

Suppose that $\{e_i, i \geq 1\}$ is a complete orthonormal system in H. We will denote by Λ the set of all sequences $a = (a_1, a_2, \dots)$, $a_i \in \mathbb{N}$ such that all the terms, except a finite number of them, vanish. For $a \in \Lambda$ we set $a! = \prod_{i=1}^{\infty} a_i!$ and $|a| = \sum_{i=1}^{\infty} a_i$. One can show the following facts:

(i) For any $n \geq 1$, the random variables

$$\left\{ \frac{1}{\sqrt{a!}} \prod_{i=1}^{\infty} H_{a_i}(B(e_i)), a \in \Lambda, |a| = n \right\}$$

form a complete orthonormal system in $\mathcal{H}_n$.

(ii) The mapping

$$I_n\left(\mathrm{symm}(\otimes_{i=1}^{\infty} e_i^{\otimes a_i})\right) = \prod_{i=1}^{\infty} H_{a_i}(B(e_i))$$

provides an isometry between the symmetric tensor product $H^{\hat{\otimes} n}$ equipped with the norm $\sqrt{n!}\|\cdot\|_{H^{\otimes n}}$ and the nth Wiener chaos $\mathcal{H}_n$. The mapping I_n is called the generalized stochastic integral of order n.

Exercises

1. Show that for all $n, m \geq 1$, $f \in L^2(\mathbb{R}_+^n)$, and $g \in L^2(\mathbb{R}_+^m)$,

$$\mathbb{E}\left(I_n(f)I_m(g)\right) = n!\langle \tilde{f}, \tilde{g}\rangle_{L^2(\mathbb{R}_+^n)} \mathbf{1}_{\{n=m\}}.$$

2. Let $B = \{B(h), h \in H\}$ be an isomormal Gaussian process on a Hilbert space H. Show that the exponentials $\{\exp(B(h)), h \in H\}$ form a total set in $L^2(\Omega, \mathcal{F}, P)$, assuming that $\mathcal{F}$ is generated by B.

3. For every n we define the extended Hermite polynomial $H_n(x, \lambda)$ by

$$H_n(x,\lambda) = \lambda^{\frac{n}{2}} H_n\left(\frac{x}{\sqrt{\lambda}}\right), \qquad \text{for } x \in \mathbb{R} \text{ and } \lambda > 0.$$

Check that

$$\exp\left(tx - \frac{t^2\lambda}{2}\right) = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x, \lambda).$$

Show that for any $g \in L^2(\mathbb{R}_+)$,

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

4. Show that the process $(H_n(B_t, t))_{t\geq 0}$ is a martingale.

5. Show that if $F \in \mathbb{D}^{\infty,2}$, then the kernels in the Wiener chaos expansion of F satisfy the following formula

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

6. Let $F = \exp\left(B(h) - \frac{1}{2}\int_{\mathbb{R}_+} h^2(s)ds\right)$, $h \in H$. Compute the iterated derivatives of F and the kernels of the Wiener chaos expansion.

7. Compute the expansion on the Wiener chaos expansion of the following of the random variables:

$$e^{B_T - \frac{T}{2}}, \quad B_T^5, \quad \int_0^1 (t^3 B_t^3 + 2tB_t^2)dB_t, \quad \int_0^1 t e^{B_t}\, dB_t.$$

8. Let $F \in \mathbb{D}^{1,2}$ be a random variable such that $\mathbb{E}(F^{-2}) < \infty$. Show that $P(F > 0)$ is zero or one.

   Hint: Using the duality relation, compute $\mathbb{E}(\operatorname{sign}(F)\delta(u))$, where u is an arbitrary bounded process in the domain of δ, using an approximation of the sign function.

7 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. Finally we show Meyer's inequality.

7.1 Ornstein-Uhlenbeck semigroup

Let $B = \{B(h), h \in H\}$ be an isomormal Gaussian process on a Hilbert space H defined on a probability space $L^2(\Omega, \mathcal{F}, P)$. Assume that $\mathcal{F}$ is generated by B. Recall the Wiener chaos expansion

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

We will denote by J_n the projection on each Wiener chaos $\mathcal{H}_n$. In the Brownian motion case, $J_n F = I_n(f_n)$, where $F = \mathbb{E}(F) + \sum_{n=1}^{\infty} I_n(f_n)$, $f_n \in L_s^2(\mathbb{R}_+^n)$.

Definition 7.1. *The Ornstein–Uhlenbeck semigroup is the one-parameter semigroup $(T_t)_{t\geq 0}$ of operators on $L^2(\Omega)$ defined by*

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

7.2 Mehler's formula

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

Proposition 7.1. *Let B' be an independent copy of B. Then, for any $t \geq 0$ and $F \in L^2(\Omega)$, we have*

$$T_t(F) = \mathbb{E}'(F(e^{-t}B + \sqrt{1 - e^{-2t}}B')), \tag{39}$$

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

Proof. Both T_t in Definition 7.1 and the right-hand side of (39) 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 (39) for random variables of the form $F = \exp\big(\lambda B(h) - \frac{1}{2}\lambda^2\big)$, where $B(h)$, $h \in H$, is an element of norm one, and $\lambda \in \mathbb{R}$.

We have, using formula (1),

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

because

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

is the Wiener chaos expansion of F. This completes the proof. $\square$

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

$$\mathbb{E}(GT_t(F))=\mathbb{E}(FT_t(G))=\sum_{n=0}^{\infty}e^{-nt}\mathbb{E}(J_nFJ_nG).$$

7.3 Hypercontractivity and applications

The Ornstein–Uhlenbeck semigroup has the following hypercontractivity property

Theorem 7.2. *Let $F\in L^p(\Omega)$, $p>1$, and $q(t)=e^{2t}(p-1)+1>p$, $t>0$. Then*

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

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 $F\in\mathcal{H}_n$,

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

which implies that

$$\|F\|_q\le\left(\frac{q-1}{p-1}\right)^{n/2}\|F\|_p. \tag{40}$$

Moreover, for any $n\ge1$ and $1<p<\infty$, the orthogonal projection onto the nth Wiener chaos J_n is bounded in $L^p(\Omega)$, and

$$\|J_nF\|_p\le\begin{cases}(p-1)^{n/2}\|F\|_p & \text{if } p>2,\\ (p-1)^{-n/2}\|F\|_p & \text{if } p<2.\end{cases} \tag{41}$$

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

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

If $p < 2$, we have

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

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

As an application we can establish the following lemma.

Lemma 7.3. *Fix an integer $k \ge 1$ and a real number $p > 1$. Then, there exists a constant $c_{p,k}$ such that, for any random variable $F \in \mathbb{D}^{k,2}$,*

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

Proof. Suppose that $F = \sum_{n=0}^{\infty} I_n(f_n)$, where $f_n \in H^{\hat{\otimes} n}$. Then, by Stroock's formula (35), $\mathbb{E}(D^k F) = k! f_k$. Therefore,

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

From (40) we obtain

$$\|J_k F\|_2 \le \big((p-1) \wedge 1\big)^{-k/2}\|J_k F\|_p.$$

Finally, applying (41) we get

$$\|J_k F\|_p \le (p-1)^{\operatorname{sign}(p-2)k/2}\|F\|_p,$$

which concludes the proof. □

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

Proposition 7.2. *Let $F \in L^p(\Omega)$ for some $p > 1$. Then, for any $t > 0$, we have that $T_t F \in \mathbb{D}^{1,p}$ and there exists a constant c_p such that*

$$\|DT_t F\|_{L^p(\Omega;H)} \le c_p t^{-1/2}\|F\|_p. \tag{42}$$

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 $T_t F_n$ converges to $T_t F$ in $L^p(\Omega)$. We have $T_t F_n \in \mathbb{D}^{1,p}$, and using Mehler's formula (39), 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 7.3 implies that

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

Hence, $DT_t F_n$ is a Cauchy sequence in $L^p(\Omega; H)$. Therefore, $T_t F \in \mathbb{D}^{1,p}$ and $DT_t F$ is the limit in $L^p(\Omega; H)$ of $DT_t F_n$. The estimate (42) follows by the same arguments. □

With the above ingredients, we can show an extension of Corollary 6.4 to any $p > 1$.

Proposition 7.3. *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 7.2, for any $t > 0$, we have that $T_t F$ belongs to $\mathbb{D}^{1,p}$ and $DT_t F_{n(k)}$ converges to $DT_t F$ in $L^p(\Omega; H)$. Then, for any $\beta \in L^q(\Omega; H)$, we can write

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

Therefore, $DT_t F = e^{-t}T_t\alpha$. This implies that $DT_t F$ 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 $DF = \alpha$. □

7.4 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 $F \in L^2(\Omega)$, belongs to the domain of L if and only if

$$\sum_{n=1}^{\infty} n^2 \mathbb{E}((J_n F)^2) < \infty;$$

and, in this case, $LF = \sum_{n=1}^{\infty} -nJ_n F$. 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

$$LF = -\sum_{n=1}^{\infty} \frac{1}{n} J_n F.$$

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 7.4. *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. Suppose first that $F \in \mathbb{D}^{1,2}$ and $DF \in \operatorname{Dom}\delta$. Then, for any random variable $G = \mathcal{H}_m$, we have, using the duality relationship (Proposition 5.1),

$$\mathbb{E}(G\delta DF) = \mathbb{E}(\langle DG, DF\rangle_H) = m\mathbb{E}(J_m FG) = \mathbb{E}(GmJ_m F).$$

Therefore, the projection of δDF onto the mth Wiener chaos is equal to mJ_mF. This implies that the series $\sum_{n=1}^{\infty} nJ_nF$ 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 $G \in \mathbb{D}^{1,2}$ we have

$$\mathbb{E}(\langle DG, DF\rangle_H) = \sum_{n=1}^{\infty} n\mathbb{E}(J_nFJ_nG) = -\mathbb{E}(GLF).$$

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 7.5. *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 5.4), $\varphi(F)$ 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 7.4, $\varphi(F)$ 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.

7.5 Meyer's inequalities

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

$$\mathbb{E}(|\delta(u)|^p) \leq c_p \left(\mathbb{E}(\|Du\|^p_{L^2(\mathbb{R}^2_+)}) + \|\mathbb{E}(u)\|^p_H \right). \tag{43}$$

Proof. We can assume that $\mathbb{E}(u) = 0$. Indeed, if $\tilde{u} = u - \mathbb{E}(u)$, then

$$\|\delta(u)\|_p \leq \|\delta(\tilde{u})\|_p + \|\delta(\mathbb{E}(u))\|_p,$$

and $\delta(\mathbb{E}(u)) = B(\mathbb{E}(u))$ being a Gaussian random variable,

$$\|\delta(\mathbb{E}(u))\|_p = c_p \|\delta(\mathbb{E}(u))\|_2 = c_p \|\mathbb{E}(u)\|_H.$$

Let $G \in \mathcal{S}$ be a smooth and cylindrical random variable such that $\|G\|_q \leq 1$, where $\frac{1}{p} + \frac{1}{q} = 1$. Then,

$$\begin{aligned}
\mathbb{E}\left(\delta(u)G\right) &= \mathbb{E}\left(\langle u, DG\rangle_H\right) = -\int_0^\infty \mathbb{E}\left(\langle LT_t u, DG\rangle_H\right) dt \\
&= \int_0^\infty \mathbb{E}\left(\langle \delta D T_t u, DG\rangle_H\right) dt = \int_0^\infty \mathbb{E}\left(\langle DT_t u, D^2G\rangle_{H^{\otimes 2}}\right) dt \\
&= \int_0^\infty e^{-t}\mathbb{E}\left(\langle T_t Du, D^2G\rangle_{H^{\otimes 2}}\right) dt \\
&= \int_0^\infty e^{-t}\mathbb{E}\left(\langle Du, T_t D^2G\rangle_{H^{\otimes 2}}\right) dt \\
&\leq \|Du\|_{L^p(\Omega;H^{\otimes 2})} \left\| \int_0^\infty e^{-t} T_t D^2 G dt \right\|_{L^q(\Omega;H^{\otimes 2})}.
\end{aligned}$$

By Mehler's formula (39),

$$\begin{aligned}
T_t D^2 G &= \tilde{\mathbb{E}}\left((D^2G)(e^{-t}B + \sqrt{1-e^{-2t}}\tilde{B}) \right) \\
&= \frac{1}{1-e^{-2t}} \tilde{\mathbb{E}}\left(\tilde{D}^2(G(e^{-t}B + \sqrt{1-e^{-2t}}\tilde{B})) \right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int_0^\infty e^{-t} T_t D^2 G dt &= \int_0^\infty \frac{e^{-t}}{1-e^{-2t}} \tilde{\mathbb{E}}[\tilde{D}^2(G(e^{-t}B + \sqrt{1-e^{-2t}}\tilde{B}))] dt \\
&= \int_0^1 \frac{1}{1-x^2} \tilde{\mathbb{E}}[\tilde{D}^2(G(xB + \sqrt{1-x^2}\tilde{B}))] dx \\
&= \int_0^{\frac{\pi}{2}} \frac{1}{\sin\theta} \tilde{\mathbb{E}}[\tilde{D}^2(G(R_\theta(B, \tilde{B}))] d\theta,
\end{aligned}$$

where $R_\theta(B,\tilde{B}) = (\cos\theta)B + (\sin\theta)\tilde{B}$.

Notice that $\tilde{D}^2(G(R_\theta(B,\tilde{B}))$ vanishes for $\theta = 0$, so the above integral is well defined. Also $\tilde{D}^2 G = 0$. Therefore, we can write

$$\int_0^\infty e^{-t} T_t D^2 G dt = \tilde{\mathbb{E}}\left[\tilde{D}^2\left(\int_0^{\frac{\pi}{2}} \frac{1}{\sin\theta}(G(R_\theta(B,\tilde{B})) - G)d\theta\right)\right].$$

From Lemma 7.3, we can write

$$\left\|\int_0^\infty e^{-t} T_t D^2 G dt\right\|_{H^{\otimes 2}}^q \leq c_{q,2}^q \tilde{\mathbb{E}}\left|\int_0^{\frac{\pi}{2}} \frac{1}{\sin\theta}(G(R_\theta(B,\tilde{B})) - G)d\theta\right|^q,$$

which implies

$$\mathbb{E}\left\|\int_0^\infty e^{-t} T_t D^2 G dt\right\|_{H^{\otimes 2}}^q \leq c_{q,2}^q \mathbb{E}\tilde{\mathbb{E}}\left|\int_0^{\frac{\pi}{2}} \frac{1}{\sin\theta}(G(R_\theta(B,\tilde{B})) - G)d\theta\right|^q.$$

Since the function $\frac{1}{\sin\theta} - \frac{1}{\theta}$ is integrable in $[0,\pi/2]$, we have

$$\mathbb{E}\tilde{\mathbb{E}}\left|\int_0^{\frac{\pi}{2}} \left(\frac{1}{\sin\theta} - \frac{1}{\theta}\right)(G(R_\theta(B,\tilde{B})) - G)d\theta\right|^q \leq c_q \mathbb{E}(|G|^q).$$

Therefore, it suffices to study the term

$$\mathbb{E}\tilde{\mathbb{E}}\left|\int_0^{\frac{\pi}{2}} \frac{1}{\theta}(G(R_\theta(B,\tilde{B})) - G)d\theta\right|^q.$$

Consider the transformation

$$T_\xi(B,\tilde{B}) = \begin{pmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{pmatrix}\begin{pmatrix} B \\ \tilde{B} \end{pmatrix},$$

which preserves the law of $(B,\tilde{B})$. Moreover, $R_\theta(T_\xi(B,\tilde{B})) = R_{\theta+\xi}(B,\tilde{B})$. Therefore,

$$\begin{aligned}
&\mathbb{E}\tilde{\mathbb{E}}\left|\int_0^{\frac{\pi}{2}} \frac{1}{\theta}(G(R_\theta(B,\tilde{B})) - G)d\theta\right|^q \\
&\quad = \frac{2}{\pi}\mathbb{E}\tilde{\mathbb{E}}\int_0^{\frac{\pi}{2}}\left|\int_0^{\frac{\pi}{2}} \frac{1}{\theta}(G(R_{\theta+\xi}(B,\tilde{B})) - G(R_\xi(B,\tilde{B}))d\theta\right|^q d\xi,
\end{aligned}$$

which is bounded by a constant times $\mathbb{E}(|G|^q)$ due to the boundedness in $L^q(\mathbb{R})$ of the Hilbert transform. $\square$

As a consequence of Theorem 7.4, the divergence operator is continuous from $\mathbb{D}^{1,p}(H)$ to $L^p(\Omega)$, and so we have *Meyer's inequality*:

$$\mathbb{E}(|\delta(u)|^p) \le c_p \left(\mathbb{E}(\|Du\|^p_{L^2(\mathbb{R}^2_+)}) + \mathbb{E}(\|u\|^p_H) \right) = c_p \|u\|^p_{1,p,H}. \tag{44}$$

This result can be extended as follows.

Theorem 7.5. *For any $p > 1$, $k \ge 1$, and $u \in \mathbb{D}^{k,p}(H)$,*

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

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

7.6 Integration-by-parts formula

The following integration-by-parts formula will play an important role in the applications of Malliavin calculus to normal approximations.

Proposition 7.6. *Let $F \in \mathbb{D}^{1,2}$ with $\mathbb{E}(F) = 0$. Let f be a continuously differentiable function with bounded derivative. Then,*

$$\mathbb{E}\left(f(F)F\right) = \mathbb{E}\left(f'(F)\langle DF, -DL^{-1}F\rangle_H\right). \tag{45}$$

Proof. By Proposition 7.4,

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

Then, using the duality relationship (Proposition 5.1) and the chain rule (Proposition 5.4), we can write

$$\begin{aligned} \mathbb{E}\left(f(F)F\right) &= -\mathbb{E}\left(f(F)\delta(DL^{-1}F)\right) \\ &= -\mathbb{E}\left(\langle D(f(F)), DL^{-1}F\rangle_H\right) \\ &= \mathbb{E}\left(f'(F)\langle DF, -DL^{-1}F\rangle_H\right). \end{aligned}$$

This completes the proof. □

Exercises

1. Let $F = \exp\left(B(h) - \frac{1}{2}\int_{\mathbb{R}_+} h^2(s)ds\right)$, $h \in H$. Show that

$$LF = -\left(B(h) - \int_{\mathbb{R}_+} h^2(s)ds\right) F.$$

2. Show that the operator $(I - L)^{-\alpha}$ is a contraction in $L^p(\Omega)$, for any $p \ge 1$, where $\alpha > 0$.

   Hint: Use the equation $(1+n)^{-\alpha} = \frac{1}{\Gamma(\alpha)}\int_0^\infty e^{-(n+1)\alpha}t^{\alpha-1}dt$.

3. Let $F \in L^2(\Omega)$ with $\mathbb{E}(F) = 0$. Show that

$$L^{-1}F = -\int_0^\infty T_t(F)dt.$$

4. Show that if $F \in L^p(\Omega)$ for some $p > 1$, then $L^{-1}F \in \mathbb{D}^{1,p}$ and

$$\|DL^{-1}F\|_{L^p(\Omega;H)} \leq c_p \|F\|_p.$$

8 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.

8.1 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 8.1 (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 5.1), 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. $\square$

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

Example 8.2. *Suppose that* $F = B_t^3$*. Then* $D_sF = 3B_t^2\mathbf{1}_{[0,t]}(s)$ *and*

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

Therefore

$$B_t^3 = 3\int_0^t (t - s + B_s^2)dB_s. \tag{46}$$

This formula should be compared with Itô's formula,

$$B_t^3 = 3\int_0^t B_s^2 dB_s + 3\int_0^t B_s ds. \tag{47}$$

Notice that equation (46) contains only a stochastic integral but it is not a martingale, because the integrand depends on t*, whereas (47) contains two terms and one is a martingale. Moreover, the integrand in (46) is unique.*

Example 8.3. *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 $\epsilon \to 0$*,*

$$F_\varepsilon = \int_0^t p_\varepsilon(B_s - x)ds \overset{L^2(\Omega)}{\longrightarrow} L_t^x. \tag{48}$$

Applying the derivative operator yields

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

Thus

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

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

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

8.2 Second integral representation

Let $B = \{B(h), h \in H\}$ be an isomormal Gaussian process on a Hilbert space H defined on a probability space $L^2(\Omega, \mathcal{F}, P)$. Assume that $\mathcal{F}$ is generated by B. Recall that L is the generator of the Ornstein–Uhlenbeck semigroup.

Proposition 8.1. *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, in the Brownian motion case, $u \in L^2(\Omega; H)$ 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 7.4,

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

In the Brownian motion case, 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 $t, t_1, \ldots, t_q$ 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 8.1 holds in the context of a general Gaussian isonormal process.

Exercises

1. Using the Clark-Ocone formula find the stochastic integral representation of the following random on the time interval $[0, T]$:

$$B_T,\ B_T^2,\ e^{B_T},\ \int_0^T B_t dt,\ B_T^3,\ \sin B_T,\ \int_0^T tB_t^2 dt.$$

2. Let $F = I_n(f)$, $f \in L^2_s([0,T]^n)$, and for any $A \subset \mathcal{B}([0,T])$ we consider the σ-field $\mathcal{F}_A = \sigma\{B(\mathbf{1}_D), D \subset A\}$. Show that

$$\mathbb{E}\left(F|\mathcal{F}_A\right) = I_n(f_n \mathbf{1}_A^{\otimes n}).$$

3. Let $F = I_n(f)$, $f \in L^2_s([0,T]^n)$. Use Clark-Ocone formula to show that $F = \delta(u)$, where $u_t = nI_{n-1}(f(\cdot,t)\mathbf{1}_{[0,t]}^{\otimes n})$. Deduce that $u_t = \mathbb{E}\left(D_tF|\mathcal{F}_t\right)$. Apply this approach to prove the Clark-Ocone formula via the Wiener chaos expansion.

4. Let $F \in L^2(\Omega, \mathcal{F}_{[a,b]}, P)$. Show that F admits the following representation

$$F = \mathbb{E}[F] + \int_a^b \mathbb{E}\left(D_tF|\mathcal{F}_{[a,t]}\right) dB_t.$$

5. Using Clark-Ocone formula show the Poincaré inequality

$$\mathrm{Var}(F) \leq \|DF\|^2_{L^2(\Omega\times[0,T])},$$

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

9 Existence and regularity of densities. Density formulas

In this Section we apply Malliavin calculus to derive explicit formulas for the densities of random variables which are functionals of an isonormal Gaussian process and we establish criteria for their regularity.

9.1 Analysis of densities in the one-dimensional case

Let $B = \{B(h), h \in H\}$ be an isomormal Gaussian process on a Hilbert space H defined on a probability space $L^2(\Omega, \mathcal{F}, P)$. Assume 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 9.1. *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 5.4, $\varphi(F) \in \mathbb{D}^{1,2}$ but, almost surely, $\varphi(F) = \mathbf{1}_{\{F \geq a+\epsilon\}}$. Therefore, by Proposition 6.5, we must have $P(F \geq a + \epsilon) = 0$ or $P(F \geq a + \epsilon) = 1$, which leads to a contradiction. $\square$

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

Proposition 9.2. *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 5.4), $\psi_n(F) \in \mathbb{D}^{1,2}$ 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 9.3. *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*

$$p(x) = \mathbb{E}\left(\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right). \tag{49}$$

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 5.4), $\varphi(F)$ 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 5.1), we obtain

$$\mathbb{E}(\psi(F)) = \mathbb{E}\left(\left\langle D(\varphi(F)), \frac{DF}{\|DF\|_H^2}\right\rangle_H\right) = \mathbb{E}\left(\varphi(F)\,\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right). \tag{50}$$

By an approximation argument, equation (50) 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 9.1. *Equation (49) 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 9.2. *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 (49) 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 (49) we can derive density estimates. Notice first that (49) 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

$$p(x) \leq (P(|F| > |x|))^{1/q}\left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_p, \tag{51}$$

for all $x \in \mathbb{R}$. Applying (51) and Meyer's inequality (44), we can deduce the following result.

Proposition 9.4. *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:*

$$\begin{aligned} p(x) \leq \quad & c_{q,\alpha,\beta}\,(P(|F| > |x|))^{1/q} \\ & \times \left(\mathbb{E}(\|DF\|_H^{-1}) + \left\|D^2F\right\|_{L^\alpha(\Omega;L^2(\mathbb{R}_+^2))}\left\|\|DF\|_H^{-2}\right\|_\beta\right). \end{aligned} \tag{52}$$

9.2 Existence and smoothness of densities for random vectors

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

$$\gamma_F = (\langle DF^i, DF^j\rangle_H)_{1\leq i,j\leq m}. \tag{53}$$

In the one-dimensional case, $\gamma_F = \|DF\|_H^2$. The following theorem is a multidimensional version of Proposition 9.2.

Theorem 9.3. *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. One can also show that 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 9.4. *We say that a random vector $F = (F^1, \ldots, F^m)$ is nondegenerate if $F^i \in \mathbb{D}^{1,2}$ for $i = 1, \ldots, m$ and*

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

for all $p \geq 2$.

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

$$D\left(\gamma^{-1}\right)^{ij} = -\sum_{k,\ell=1}^{m} \left(\gamma^{-1}\right)^{ik} \left(\gamma^{-1}\right)^{\ell j} D\gamma^{k\ell}. \tag{54}$$

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 7.3 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 (54) on applying the derivative operator and letting ϵ tend to zero. $\square$

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 9.5. *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 $H_\alpha(F, G) \in \cap_{p \geq 2} L^p(\Omega)$ such that*

$$\mathbb{E}(\partial_\alpha \varphi(F) G) = \mathbb{E}(\varphi(F) H_\alpha(F, G)), \tag{55}$$

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 5.4), 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 5.1) 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 7.4) and Lemma 9.5 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

$$\|H_\alpha(F,G)\|_p \leq c_{p,q} \left\|\det\gamma_F^{-1}\right\|_\beta^m \|DF\|_{k,\gamma}^n \|G\|_{k,q}. \tag{56}$$

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 (49).

Proposition 9.6. *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*

$$p(x) = \mathbb{E}(\mathbf{1}_{\{F>x\}} H_\alpha(F,1)), \tag{57}$$

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

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

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

Then, equality (55) 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

$$\mathbb{E}(\partial_\alpha \varphi(F)) = \int_{\mathbb{R}^m} \partial_\alpha \varphi(x) \mathbb{E}(\mathbf{1}_{\{F>x\}} H_\alpha(F,1)) dx. \tag{58}$$

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 (58) 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 9.6. *Let $F = (F^1, \dots, F^m)$ be a nondegenerate random vector such that $F^i \in \mathbb{D}^\infty$ for all $i = 1, \dots, m$. 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. □

9.3 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 (57), 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 $1 \leq i \leq m$,

$$\partial_i Q_m(x) = -c_m \frac{x_i}{|x|^m}, \tag{59}$$

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

$$\varphi(x) = \nabla\varphi * \nabla Q_m(x) = \sum_{i=1}^m \int_{\mathbb{R}^m} \partial\varphi(x-y)\partial_i Q_m(y)dy. \tag{60}$$

Indeed,

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

Theorem 9.7. *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 (60), 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 (59) we obtain

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

As a consequence, Fubini's theorem and (55) 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. $\square$

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

Lemma 9.8. *Under the assumptions of Theorem 9.7, 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 (59))

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

yields

$$p(x) \le m c_m A \left(\mathbb{E} \left(|F-x|^{(1-m)q} \right) \right)^{1/q}, \tag{61}$$

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 $\epsilon > 0$,

$$\begin{aligned} \mathbb{E}(|F-x|^{(1-m)q}) &\le \epsilon^{(1-m)q} + \int_{|z-x| \le \epsilon} |z-x|^{(1-m)q} p(x) dx \\ &\le \epsilon^{(1-m)q} + C_{m,p} \epsilon^{(p-m)/(p-1)} M. \end{aligned} \tag{62}$$

Therefore, substituting (62) into (61), we get

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

Now we minimize with respect to ϵ and obtain $M \le A C_{m,p} M^{1-1/m}$, 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. □

Exercises

1. Let F be a random variable with the law $N(0, \sigma^2)$. Show that the density $p(x)$ of F satisfies
$$p(x) = \sigma^{-2} \mathbb{E} \left(\mathbf{1}_{\{F > x\}} F \right).$$

2. Derive (52) applying (51) and Meyer's inequality (43).

3. Set $M_t = \int_0^t u_s dB_s$, where $u = (u_t)_{t\in(0,T]}$ is a process in $L^2_T(\mathcal{P})$ such that for all $t \in [0,T]$, $|u_t| \geq \rho > 0$ for some constant ρ, $u_t \in \mathbb{D}^{2,2}$, and

$$\lambda := \sup_{s,t\in[0,T]} \mathbb{E}(|D_s u_t|^p) + \sup_{r,s\in[0,T]} \mathbb{E}\left(\left\{\int_0^T |D^2_{r,s} u_t|^p dt\right\}^{p/2}\right) < \infty,$$

for some $p > 3$. Applying Proposition 9.4 show that the density of M_t, denoted by $p_t(x)$ satisfies

$$p_t(x) \leq \frac{c}{\sqrt{t}} P(|M_t| > |x|)^{\frac{1}{q}},$$

for all $t \in [0,T]$, where $q > \frac{p}{p-3}$ and the constant c depends on λ, ρ and p.

Hint: Use a lower bound of the form

$$\|DM_t\|^2_H = \int_0^t \left(u_s + \int_s^t D_s u_r dB_r\right)^2 ds \geq \int_{t(1-h)}^t \left(u_s + \int_s^t D_s u_r dB_r\right)^2 ds$$

and choose h small enough.

4. Show inequality (56).

10 Malliavin differentiability of diffusion processes. Proof of Hörmander's theorem

Suppose that $B = (B_t)_{t\geq 0}$, with $B_t = (B_t^1, \dots, B_t^d)$, is a d-dimensional Brownian motion. Consider the m-dimensional stochastic differential equation

$$dX_t = \sum_{j=1}^{d} \sigma_j(X_t) dB_t^j + b(X_t)dt, \tag{63}$$

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

By definition, a solution to equation (63) is an adapted process $X = (X_t)_{t\geq 0}$ such that, for any $T > 0$ and $p \geq 2$,

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

and X satisfies the integral equation

$$X_t = x_0 + \sum_{j=1}^{d} \int_0^t \sigma_j(X_s) dB_s^j + \int_0^t b(X_s)ds. \tag{64}$$

The following result is well known.

Theorem 10.1. *Suppose that the coefficients $\sigma_j, b \colon \mathbb{R}^m \to \mathbb{R}^m$, $1 \leq j \leq d$, satisfy the Lipschitz condition: for all $x, y \in \mathbb{R}^m$,*

$$\max_j \left(|\sigma_j(x) - \sigma_j(y)|, |b(x) - b(y)|\right) \leq K|x - y|. \tag{65}$$

Then there exists a unique solution X to Equation (64).

When the coefficients in equation (63) are continuously differentiable, the components of the solution are differentiable in the Malliavin calculus sense.

Proposition 10.1. *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 $j = 1, \dots, d$,*

$$\begin{aligned} D_r^j X_t = &\sigma_j(X_r) + \sum_{k=1}^{m} \sum_{\ell=1}^{d} \int_r^t \partial_k \sigma_\ell(X_s) D_r^j X_s^k dB_s^\ell \\ &+ \sum_{k=1}^{m} \int_r^t \partial_k b(X_s) D_r^j X_s^k ds. \end{aligned} \tag{66}$$

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 $i = 1, \ldots, m, t \geq 0$. Moreover, for all $p > 1$ and $t \geq 0$,

$$\psi_n(t) := \sup_{0 \leq r \leq t} \mathbb{E}\left(\sup_{s \in [r,t]} |D_r X_s^{(n)}|^p \right) < \infty \tag{67}$$

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

$$\psi_{n+1}(t) \leq c_1 + c_2 \int_0^t \psi_n(s) ds, \tag{68}$$

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 (31) of the divergence operator and the chain rule (Proposition 5.4), for any $r \leq t$, $i = 1, \ldots, m$, and $\ell = 1, \ldots, d$, 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 (67) 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,

$$\mathbb{E}\left(\sup_{r \leq s \leq t} |D_r X_s^{(n+1)}|^p \right) \leq c_p \left(\gamma_p + T^{(p-1)/2} K^p \int_r^t \mathbb{E}\left(|D_r^j X_s^{(n)}|^p \right) ds \right), \tag{69}$$

where

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

So (67) and (68) hold for $n+1$ and the claim is proved.

We know that

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

as n tends to infinity. By Gronwall's lemma applied to (68) we deduce that the derivatives of the sequence $X_t^{(n),i}$ are bounded in $L^p(\Omega; H)$ uniformly in n for all $p \geq 2$. This implies that the random variables X_t^i belong to $\mathbb{D}^{1,\infty}$. Finally, applying the operator D to equation (64) we deduce the linear stochastic differential equation (66) for the derivative of X_t^i.

This completes the proof of the proposition. $\square$

Example 10.2. *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 $(D_r X_t)_{r \leq t}$ satisfies the following linear equation:

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

Therefore, by Itô's formula,

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

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

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

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 (64) 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 10.2. *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 10.3. *The $m \times d$ matrix $(D_rX_t)_j^i = D_r^jX_t^i$ can be expressed as*

$$D_rX_t = Y_tY_r^{-1}\sigma(X_r), \tag{70}$$

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 (70) leads to

$$Q_t = Y_tC_tY_t^T, \tag{71}$$

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*.

10.1 Absolute continuity under ellipticity conditions

Consider the following ellipticity condition:

(E) The vectors $\sigma_1(x_0), \ldots, \sigma_d(x_0)$ span $\mathbb{R}^m$.

Notice that condition **(E)** is equivalent to saying that the matrix $\sigma\sigma^T(x_0)$ is invertible. We need the following technical lemma.

Lemma 10.4. *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 $p \geq 2$ there exists $\epsilon_0(p)$ such that, for all $\epsilon \leq \epsilon_0(p)$,*

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

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

Proof. Let $\lambda = \inf_{|v|=1} v^T C v$ be the smallest eigenvalue of C. We know that $\lambda^m \leq \det C$. Thus, it suffices to show that $\mathbb{E}(\lambda^{-p}) < \infty$ for all $p \geq 2$. 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

$$\begin{aligned} P(\lambda < \epsilon) &= P\left(\inf_{|v|=1} v^T C v < \epsilon\right) \\ &\leq P\left(\inf_{|v|=1} v^T C v < \epsilon, |C| \leq \frac{1}{\epsilon}\right) + P\left(|C| > \frac{1}{\epsilon}\right). \end{aligned} \tag{72}$$

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, (72) 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. $\square$

Theorem 10.5. *Let $(X_t)_{t\geq 0}$ be a diffusion process with $C^{1+\alpha}$ and Lipschitz coefficients. Then, for any $t > 0$, the random vector X_t is non-degenerate.*

Proof. We know that $X_t^i \in \mathbb{D}^{1,\infty}$ for each $i = 1, \ldots, m$. Then, it suffices to show that for all $t > 0$ and all $p \geq 2$, $\mathbb{E}((\det Q_t)^{-p}) < \infty$, where Q_t is the Malliavin matrix of X_t. The proof of this will be done in several steps:

Step 1: 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 $p \geq 2$.

Step 2: Using Lemma 10.4, the problem reduces to showing that, for all $p \geq 2$, we have

$$. \sup_{|v|=1} P\left(v^T C_t v \leq \epsilon\right) \leq \epsilon^p, \tag{73}$$

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

$$v^T C_t v = \sum_{j=1}^{d} \int_0^t \langle v, Y_s^{-1}\sigma_j(X_s)\rangle^2 ds. \tag{74}$$

Step 3: In order to show the estimate (100), we set

$$Z_s = \sum_{j=1}^{d} \langle v, Y_s^{-1}\sigma_j(X_s)\rangle^2.$$

The ellipticity assumption implies that

$$Z_0 = \sum_{j=1}^{d} \langle v, Y_s^{-1}\sigma_j(x_0)\rangle^2 \geq \inf_{|v|=1} \sum_{j=1}^{d} \langle v, Y_s^{-1}\sigma_j(x_0)\rangle^2 =: \gamma > 0.$$

Suppose that $\epsilon < \frac{t\gamma}{2}$. Then, we have

$$\begin{aligned} P\left(v^T C_t v \leq \epsilon\right) &= P\left(\sum_{j=1}^{d} \int_0^t \langle v, Y_s^{-1}\sigma_j(X_s)\rangle^2 ds \leq \epsilon\right) \\ &\leq P\left(\int_0^{2\epsilon/\gamma} Z_s ds \leq \epsilon\right) \\ &\leq P\left(\sup_{0\leq s\leq 2\epsilon/\gamma} |Z_s - Z_0| \leq \frac{Z_0}{2}\right). \end{aligned}$$

. Then, for any $p \geq 2$, we obtain

$$
\begin{aligned}
P\left(v^T C_t v \leq \epsilon\right) &\leq \left(\frac{2}{Z_0}\right)^p E\left[\sup_{0\leq s\leq 2\epsilon/\gamma} |Z_s - Z_0|^p\right] \\
&\leq \left(\frac{2}{\gamma}\right)^{p(1+\alpha)} \epsilon^{p\alpha},
\end{aligned}
$$

for any $\alpha < 1/2$. This completes the proof. □

Corollary 10.6. *Under the assumptions of Theorem 10.5, for any $t > 0$, X_t has a density.*

Example 10.7. *Assume that $\sigma(x_0) \neq 0$ in Example 10.2. Then, for any $t > 0$, the law of X_t is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}$.*

10.2 Regularity of the density under Hörmander's conditions

We need the following regularity result, whose proof is similar to that of Proposition 10.1 and is thus omitted.

Proposition 10.3. *Suppose that the coefficients σ_j, $1 \leq j \leq m$, and b of equation (63) 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 (64) 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

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

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 $n \geq 1$.

Theorem 10.8. *Suppose that the coefficients σ_j, $1 \leq j \leq m$, and b of equation* (63) *are infinitely differentiable with bounded derivatives of all orders.. 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

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

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).

The next lemma was proved by Norris, following the ideas of Stroock, and is the basic ingredient in the proof of Theorem 10.8.

Lemma 10.9 (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 $c = \mathbb{E}\left(\sup_{0\leq t\leq T}(|\beta_t| + |\gamma_t| + |a_t| + |u_t|)^p\right) < \infty$ *for some* $p \geq 2$.

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

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

Proof of Theorem 10.8. The proof will be carried out in several steps:

Step 1 Fix $t > 0$. As in the proof of Theorem 10.5, the problem reduces to showing that, for all $p \geq 2$, we have

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

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

$$v^T C_t v = \sum_{j=1}^{d} \int_0^t \langle v, Y_s^{-1} \sigma_j(X_s) \rangle^2 ds. \tag{75}$$

Step 2 Fix a smooth function V and use Itô's formula to compute the differential of $Y_t^{-1} V(X_t)$:

$$\begin{aligned} d\left(Y_t^{-1} V(X_t)\right) &= Y_t^{-1} \sum_{k=1}^{d} [\sigma_k, V](X_t) dB_t^k \\ &\quad + Y_t^{-1} \left([\sigma_0, V] + \tfrac{1}{2} \sum_{k=1}^{d} [\sigma_k, [\sigma_k, V]] \right)(X_t) dt. \end{aligned} \tag{76}$$

Step 3 We introduce the following sets of vector fields:

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

and

$$\begin{aligned} \Sigma'_0 &= \Sigma_0, \\ \Sigma'_n &= \Big\{ [\sigma_k, V], k = 1, \ldots, d, V \in \Sigma'_{n-1}; \\ &\qquad [\sigma_0, V] + \tfrac{1}{2} \sum_{j=1}^{d} [\sigma_j, [\sigma_j, V]], V \in \Sigma'_{n-1} \Big\} \quad \text{if } n \geq 1, \\ \Sigma' &= \cup_{n=0}^{\infty} \Sigma'_n. \end{aligned}$$

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,

$$\inf_{|v|=1} \sum_{j=0}^{j_0} \sum_{V \in \Sigma'_j} \langle v, V(x_0) \rangle^2 = R > 0. \tag{77}$$

Step 5 For any $j = 0, 1, \ldots, j_0$ 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 5 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 (77), we proceed as in the proof of Theorem 10.5, using the process

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

and we obtain that for any $p \geq 1$, there exists ϵ_0 such that

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

for any $\epsilon < \epsilon_0$.

Step 6 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 (76) 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 10.9) 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. □

Exercises

1. Consider the stochastic differential equation

$$\begin{aligned} dX_t^1 &= dB_t^1 + \sin X_t^2 dB_t^2, \\ dX_t^2 &= 2X_t^1 dB_t^1 + X_t^1 dB_t^2 \end{aligned}$$

with initial condition $x_0 = 0$. Show that the vector fields σ_1 and $[\sigma_1, \sigma_2]$ at $x = 0$ span $\mathbb{R}^2$ and Hörmander's condition holds. Deduce that X_t has a $C^\infty(\mathbb{R}^2)$ density for any $t > 0$.

11 Malliavin-Stein's approach for normal approximations

In this section we will combine the techniques of Malliavin calculus with Stein's method for normal approximations in order to establish quantitative versions of central limit theorems.

11.1 Stein's method for normal approximation

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

Lemma 11.1. *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*

$$\mathbb{E}(f'(X) - f(X)X) = 0. \tag{78}$$

Proof. Suppose first that X has the standard normal distribution $N(0,1)$. Then, equality (78) 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

$$f_h'(x) - xf_h(x) = h(x) - \mathbb{E}(h(X)), \quad x \in \mathbb{R}. \tag{79}$$

Definition 11.2. *A solution to equation (79) is an absolutely continuous function f_h such that there exists a version of the derivative f_h' satisfying (79) for every $x \in \mathbb{R}$.*

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

Proposition 11.1. *The function*

$$f_h(x) = e^{x^2/2}\int_{-\infty}^{x} (h(y) - \mathbb{E}(h(X)))e^{-y^2/2}dy \tag{80}$$

is the unique solution of Stein's equation (79) *satisfying*

$$\lim_{x\to\pm\infty} e^{-x^2/2} f_h(x) = 0. \tag{81}$$

Proof. Equation (79) 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 (79) 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 (81) 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

$$\int_{-\infty}^{x}(h(y) - \mathbb{E}(h(X)))e^{-y^2/2}dy = -\int_{x}^{\infty}(h(y) - \mathbb{E}(h(X)))e^{-y^2/2}dy. \tag{82}$$

Proposition 11.2. *Let $h\colon \mathbb{R} \to [0,1]$ be a measurable function. Then the solution to Stein's equation f_h given by* (80) *satisfies*

$$\|f_h\|_\infty \leq \sqrt{\frac{\pi}{2}} \quad \text{and} \quad \|f_h'\|_\infty \leq 2. \tag{83}$$

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

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

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 (82), 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

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

This completes the proof. □

11.2 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 11.3. *We say that* $F_n \xrightarrow{\mathcal{L}} F$ *if* $\mathbb{E}[g(F_n)] \to \mathbb{E}[g(F)]$ *for any* $g : \mathbb{R} \to \mathbb{R}$ *continuous and bounded.*

We know that $F_n \xrightarrow{\mathcal{L}} F$ if and only if $P(F_n \leq z) \to P(F \leq z)$ 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

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

We denote by $P_F = P \circ F^{-1}$ the probability distribution of F. Then, $d_{TV}(P_{F_n}, P_F) \to 0$ 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 11.3. *Let* ν *be a probability on* $\mathbb{R}$*. Then,*

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

where

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

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,

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

Integrating with respect to ν yields

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

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. $\square$

11.3 The Malliavin-Stein's approach

Let $B = \{B(h), h \in H\}$ be an isomormal Gaussian process on a Hilbert space H defined on a probability space $L^2(\Omega, \mathcal{F}, P)$. Assume that $\mathcal{F}$ is generated by B. The following results connects Stein's method with Malliavin calculus.

Theorem 11.4 (Nourdin-Peccati)**.** *Suppose that $F \in \mathbb{D}^{1,2}$ satisfies $F = \delta(u)$, where u belongs to the domain in $L^2(\Omega)$ of the divergence operator δ. Then,*

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

Proof. It follows from

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

Therefore,

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

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

Example 11.5. *Consider the Brownian motion case and let $F = \int_0^T u_s dB_s$, where u is a progressively 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,

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

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

A particular example is

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

Remarks:

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

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

and, we obtain

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

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

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

Proof. We can write

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

Finally, using that

$$\mathbb{E}(\langle DF, u\rangle_H) = \mathbb{E}(F\delta(u)) = \mathbb{E}(F^2) = \sigma^2,$$

and applying Caucy-Schwarz inequality we get the second inequality. □

11.4 Normal approximation on a fixed Wiener chaos

Proposition 11.4. *Suppose $F \in \mathcal{H}_q$ for some $q \geq 2$ and $E(F^2) = \sigma^2$. Then,*

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

Proof. Using $L^{-1}F = -\frac{1}{q}F$, and taking $u = -DL^{-1}F$, we obtain

$$\langle DF, u\rangle_H = \langle DF, -DL^{-1}F\rangle_H = \frac{1}{q}\|DF\|_H^2,$$

which allows us to conclude the proof □

Proposition 11.5. *Suppose that $F = I_q(f) \in \mathcal{H}_q$, $q \geq 2$. Then,*

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

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

$$\boxed{\mathrm{Var}\left(\|DF\|_H^2\right) = \sum_{r=1}^{q-1} r^2 (r!)^2 \binom{q}{r}^4 (2q-2r)! \|f \widetilde{\otimes}_r f\|_{H^{\otimes(2q-2r)}}^2} \tag{84}$$

Proof of (84): We have $D_t F = qI_{q-1}(f(\cdot, t))$, and using the product formula for multiple stochastic integrals we obtain

$$\begin{aligned} \|DF\|_H^2 &= q^2 \int_0^T I_{q-1}(f(\cdot,t))^2 dt \\ &= q^2 \sum_{r=0}^{q-1} r! \binom{q-1}{r}^2 I_{2q-2r-2}(f \widetilde{\otimes}_{r+1} f) \\ &= q^2 \sum_{r=1}^{q} (r-1)! \binom{q-1}{r-1}^2 I_{2q-2r}(f \widetilde{\otimes}_r f) \\ &= qq! \|f\|_{H^{\otimes q}}^2 + q^2 \sum_{r=1}^{q-1} (r-1)! \binom{q-1}{r-1}^2 I_{2q-2r}(f \widetilde{\otimes}_r f). \end{aligned} \tag{85}$$

Then, (84) follows from the isometry property of multiple integrals.
The second formula is the following one:

$$\boxed{E[F^4] - 3\sigma^4 = \frac{3}{q} \sum_{r=1}^{q-1} r(r!)^2 \binom{q}{r}^4 (2q-2r)! \|f \widetilde{\otimes}_r f\|_{H^{\otimes(2q-2r)}}^2} \tag{86}$$

Proof on (86): Using that $-L^{-1}F = \frac{1}{q}F$ and $L = -\delta D$ we can write

$$\begin{aligned} \mathbb{E}[F^4] &= \mathbb{E}[F \times F^3] = \mathbb{E}[(-\delta D L^{-1} F) F^3] = \mathbb{E}[\langle -DL^{-1}F, D(F^3)\rangle_H] \\ &= \frac{1}{q}\mathbb{E}[\langle DF, D(F^3)\rangle_H] = \frac{3}{q}\mathbb{E}[F^2 \|DF\|_H^2]. \end{aligned} \tag{87}$$

By the product formula of multiple integrals,

$$F^2 = I_q(f)^2 = q! \|f\|_{H^{\otimes q}}^2 + \sum_{r=0}^{q-1} r! \binom{q}{r}^2 I_{2q-2r}(f \widetilde{\otimes}_r f). \tag{88}$$

Then (86) follows from (87), (88), (85) and the isometry property of multiple integrals. $\square$

11.5 Fourth Moment theorem

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

Theorem 11.6 ([19, 20]). *Fix $q \geq 2$. Let $F_n = I_q(f_n) \in \mathcal{H}_q$, $n \geq 1$ be such that*

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

The following conditions are equivalent:

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

(ii) $\mathbb{E}(F_n^4) \to 3\sigma^4$, *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 (88) we get

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

Then, we use the fact that $(2q)!\|f_n \widetilde{\otimes} f_n\|^2_{H^{\otimes 2q}}$ equals to $2(q!)^2 \|f_n\|^4_{H^{\otimes q}}$ 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

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

□

11.6 Multivariate Gaussian approximation

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

Theorem 11.7 ([22]). *Let $d \geq 2$ and $1 \leq q_1 < \cdots < q_d$. Consider random vectors*

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

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

$$\lim_{n\to\infty} \mathbb{E}[(F_n^i)^2] = \sigma_i^2. \tag{89}$$

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

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

Proof. It suffices to show that (ii) implies (i). The sequence of random vectors $(F_n)_{n\geq 1}$ is tight by condition (89). Then, it suffices to show that the limit in distribution of any converging subsequence is $N_d(0,\Sigma)$. We can then assume that the sequence $(F_n)_{n\geq 1}$ converges in law to some random vector F_∞ and it only remains to show that the law of F_∞ is $N_d(0,\Sigma)$. For every n and $t \in \mathbb{R}^d$, define $\varphi_n(t) = \mathbb{E}(e^{i\langle t, F_n\rangle})$. Let $\varphi(t)$ be the characteristic function of the random vector F_∞. We know that for all $j = 1, \ldots, d$,

$$\frac{\partial \varphi_n}{\partial t_j}(t) = i\mathbb{E}\left(F_n^j e^{i\langle t, F_n\rangle}\right) \underset{n\to\infty}{\to} i\mathbb{E}\left(F_\infty^j e^{i\langle t, F_\infty\rangle}\right) = \frac{\partial \varphi}{\partial t_j}(t). \tag{90}$$

On the other hand, using the definition of the operator L, Proposition 7.4, and the duality relation between D and δ, we have

$$\begin{aligned}
\mathbb{E}\left(F_n^j e^{i\langle t, F_n\rangle}\right) &= -\frac{1}{q_j}\mathbb{E}\left(LF_n^j e^{i\langle t, F_n\rangle}\right) = -\frac{1}{q_j}\mathbb{E}\left(-\delta D(F_n^j) e^{i\langle t, F_n\rangle}\right) \\
&= \frac{1}{q_j}\mathbb{E}\left(\langle DF_n^j, D(e^{i\langle t, F_n\rangle})\rangle_H\right) = \frac{i}{q_j}\sum_{h=1}^d t_h \mathbb{E}\left(e^{i\langle t, F_n\rangle}\gamma_{F_n}^{jh}\right),
\end{aligned}$$

where γ_{F_n} is the Malliavin matrix of the random vector F_n, introduced in (53). Therefore,

$$\frac{\partial \varphi_n}{\partial t_j}(t) = -\frac{1}{q_j}\sum_{h=1}^d t_h \mathbb{E}\left(e^{i\langle t, F_n\rangle}\gamma_{F_n}^{jh}\right). \tag{91}$$

We claim that the following convergences hold in $L^2(\Omega)$ for all $1 \leq j, h \leq d$, as n tends to infinity:

$$\gamma_{F_n}^{jh} \to 0, \quad j \neq h, \tag{92}$$

and

$$\gamma_{F_n}^{jj} \to q_j \sigma_j^2. \tag{93}$$

In fact, (93) follows from condition (iii) in Theorem 11.6 and the fact that $\gamma_{F_n}^{jj} = \|DF_n^j\|_H^2$. On the other hand, using the product formula for multiple stochastic integrals and assuming $j < h$ yields

$$\begin{aligned}
\mathbb{E}\left(\langle DF_n^j, DF_n^h\rangle_H^2\right) &= q_j^2 q_h^2 \mathbb{E}\left(\left\{\int_0^\infty I_{q_j-1}(f_n^j(\cdot,t)) I_{q_h-1}(f_n^h(\cdot,t))dt\right\}^2\right) \\
&= q_j^2 q_h^2 \mathbb{E}\left(\left\{\sum_{r=0}^{q_j-1} \binom{q_j-1}{r}\binom{q_h-1}{r} r! I_{q_j+q_h-2-2r}(f_n^j \otimes_{r+1} f_n^h)\right\}^2\right) \\
&= q_j^2 q_h^2 \sum_{r=0}^{q_j-1} \binom{q_j-1}{r}^2 \binom{q_h-1}{r}^2 (r!)^2 \|f_n^j \tilde{\otimes}_{r+1} f_n^h\|_{H^{\otimes(q_j+q_h-2-2r)}}^2 \\
&\le q_j^2 q_h^2 \sum_{r=1}^{q_j} \binom{q_j-1}{r-1}^2 \binom{q_h-1}{r-1}^2 ((r-1)!)^2 \|f_n^j \otimes_r f_n^h\|_{H^{\otimes(q_j+q_h-2r)}}^2.
\end{aligned}$$

Then, to show (92) it suffices to check that $\|f_n^j \otimes_r f_n^h\|_{H^{\otimes(q_j+q_h-2r)}}^2$ converges to zero for all $1 \le r \le q_j$. This follows from

$$\begin{aligned}
\|f_n^j \otimes_r f_n^h\|_{H^{\otimes(q_j+q_h-2r)}}^2 &= \langle f_n^j \otimes_{q_j-r} f_n^j, f_n^h \otimes_{q_h-r} f_n^h\rangle_{H^{\otimes 2r}} \\
&\le \|f_n^j \otimes_{q_j-r} f_n^j\|_{H^{\otimes 2r}} \|f_n^h \otimes_{q_h-r} f_n^h\|_{H^{\otimes 2r}},
\end{aligned}$$

and the fact that $\|f_n^h \otimes_{q_h-r} f_n^h\|_{H^{\otimes 2r}}$ converges to zero by condition (iv) in Theorem 11.6 because $1 \le q_h - r \le q_h - 1$. Finally, (92) and (93) allows us to take the limit as $n \to \infty$ in the right-hand side of (91), and in view of (90), we obtain

$$\frac{\partial \varphi}{\partial t_j}(t) = -t_j \sigma_j \varphi(t),$$

for $j = 1, \dots, d$. As a consequence, φ is the characteristic function of the law $N_d(0, \Sigma)$. □

11.7 Chaotic Central Limit Theorem

Theorem 11.8. *Let $F_n = \sum_{q=1}^\infty I_q(f_{q,n})$, $n \ge 1$. Suppose that:*

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

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

(iii) $q!\|f_{q,n}\|^2 \le \delta_q$, 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)': $\lim_{n\to\infty} \mathbb{E}(I_q(f_{q,n})^4) = 3\sigma_q^4$, $q \geq 2$. 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.

Proof. Let ξ_q, $q \geq 1$, be independent centered Gaussian random variables with variances σ_q^2. For every $N \geq 1$, set $F_n^N = \sum_{q=1}^N I_q(f_{q,n})$ and $\xi^N = \sum_{q=1}^N \xi_q$. By Theorems 11.6 and 11.7, for each fixed N, F_n^N converges in law to ξ^N as n tends to infinity. Define also $\xi = \sum_{q=1}^{\infty} \xi_q$. Let f be a $C_b^1(\mathbb{R})$ function such that $\|f\|_\infty$ and $\|f'\|_\infty$ are bounded by one. Then,

$$\begin{aligned}
|\mathbb{E}(f(F_n)) - \mathbb{E}(f(\xi))| &\leq \left|\mathbb{E}(f(F_n)) - \mathbb{E}\left(f(F_n^N)\right)\right| \\
&\quad + \left|\mathbb{E}\left(f(F_n^N)\right) - \mathbb{E}\left(f(\xi^N)\right)\right| + \left|\mathbb{E}\left(f(\xi^N)\right) - \mathbb{E}(f(\xi))\right| \\
&\leq \left(\sum_{q=N+1}^{\infty} q!\|f_{q,n}\|^2_{H^{\otimes q}}\right)^{1/2} + \left|\mathbb{E}\left(f(F_n^N)\right) - \mathbb{E}\left(f(\xi^N)\right)\right| \\
&\quad + \left|\mathbb{E}\left(f(\xi^N)\right) - \mathbb{E}(f(\xi))\right|.
\end{aligned}$$

Taking first the limit as n tends to infinity, and then the limit as N tends to infinity, and applying conditions (i), (iii), we finish the proof. $\square$

Exercises

1. Show that the sequence $F^{(n)} = \int_0^1 u_t^{(n)} dB_t$ where B is a Brownian motion and

$$u_t^{(n)} = \sqrt{2n}t^n \exp(B_t(1-t)), \quad 0 \leq t \leq 1,$$

   converges in law to the distribution $N(0,1)$.

2. Let $B = (B_t)_{t\in[0,1]}$ be a standard Brownian motion. Consider the sequence of Itô's integrals

$$F_n = \sqrt{n}\int_0^1 t^n B_t dB_t, \quad n \geq 1.$$

   Show that sequence F_n converges stably to ηS as $n \to \infty$, where η is a random variable independent of B with law $N(0,1)$ and $S = \frac{|B_1|}{\sqrt{2}}$.

   Hint: A sequence of random variables F_n is said to converge *stably* to a random variable F defined on an extended probability space $(\Omega, \tilde{\mathcal{F}}, P)$ if one of the following equivalent conditions hold true:

(i) $\mathbb{E}[Ze^{i\lambda F_n}] \to \mathbb{E}[Ze^{i\lambda F}]$ for any $\lambda \in \mathbb{R}$ and for any bounded and $\mathcal{F}$-measurable random variable Z.

(ii) (F_n, Z) converges in law to $(F < Z)$, and for any bounded and $\mathcal{F}$-measurable random variable Z.

You need to show that (B, F_n) converges in law to $(B, \eta S)$ in $C([0,1]) \times \mathbb{R})$. Then, use the decomposition $\int_0^1 = \int_{1-\delta}^1 + \int_0^{1-\delta}$ for $\delta \in (0,1)$ and let $n \to \infty$ and $\delta \to 0$.

3. Let $q \geq 2$, and let $F_n = I_q(f_n)$ be a sequence of multiple stochastic integrals such that $E[F_n^2] = 1$ for all n. Denote $\phi_n(\lambda) = E[e^{i\lambda F_n}]$ the characteristic function of F_n.

   (i) Show that, for all λ real,

   $$\phi_n'(\lambda) = -\frac{\lambda}{q} E[e^{i\lambda F_n} \|DF\|_H^2].$$

   (ii) Using (i) as well as the fact that the mapping $t \to E[e^{itZ}]$, $Z \sim N0, 1)$ is the unique solution to the differential equation $\phi'(t) + t\phi(t) = 0$, $\phi(0) = 1$, show that $\mathrm{Var}\|DF\|_H^2 \to 0$ implies that F_n converges in law to a centered normal random variable.

4. Let $F = (F^{(1)}, \dots, F^{(m)})$ be a random vector such that $F^{(i)} = \delta(v^{(i)})$ for $v^{(i)} \in \mathrm{Dom}\,\delta$, $i = 1, \dots, m$. Assume $F^{(i)} \in \mathbb{D}^{1,2}$ for $i = 1, \dots, m$. Let Z be an m-dimensional Gaussian centered vector with covariance matrix $(C_{i,j})_{1 \leq i,j \leq m}$. Show that for any C^2 function $h : \mathbb{R}^m \to \mathbb{R}$ with bounded second partial derivatives, we have

$$|\mathbb{E}(h(F)) - \mathbb{E}(h(Z))| \leq \frac{1}{2}\|h''\|_\infty \sqrt{\sum_{i,j=1}^m \mathbb{E}\left[(C_{i,j} - \langle DF^{(i)}, v^{(j)}\rangle_H)^2\right]},$$

where

$$\|h''\|_\infty = \max_{1 \leq i,j \leq m} \sup_{x \in \mathbb{R}^m} \left| \frac{\partial^2 h}{\partial x_i \partial x_j}(x) \right|.$$

Hint: Assume that Z is independent of the underlying isonormal Gaussian process B. Introduce the function $\Psi(t) = \mathbb{E}[h(\sqrt{1-t})F + \sqrt{t}Z)]$, $t \in [0,1]$. Then

$$\mathbb{E}(h(F)) - \mathbb{E}(h(Z)) = \int_0^1 \Psi'(t)dt.$$

Then, compute the derivative of the function Ψ and integrate by parts with respect to Z and use the representation of the components of F as a divergence and the duality formula.

12 Central limit theorems for stationary sequences

Suppose that $(X_n)_{n\geq 1}$ is a sequence of random variables, defined in some probability space $(\Omega, \mathcal{F}, P)$, which are independent and identically distributed, with zero mean and finite variance $\mathbb{E}[X_1^2] = \sigma^2$. Then, the classical central limit theorem says that

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

That means, for any $a \leq b$,

$$\lim_{n\to\infty} P\left(a \leq \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} dx.$$

In this section we will apply the Malliavin-Stein's method to establish quantitative versions of the central limit theorem, where the independence property is replaced by stationarity and we consider random variables which are functionals of a given Gaussian process.

12.1 Breuer-Major theorem

Recall that γ denotes the standard normal distribution and H_q is the q-th Hermite polynomial. A function $g \in L^2(\mathbb{R}, \gamma)$ has *Hermite rank* $d \geq 1$ if

$$g(x) = \sum_{q=d}^{\infty} c_q H_q(x), \quad c_d \neq 0.$$

For example, $g(x) = |x|^p - \int_{\mathbb{R}} |x|^p d\gamma(x)$, $p \geq 1$, has Hermite rank 2.

Suppose that

$$\boxed{X_n = g(\theta_n)}, \quad n \in \mathbb{Z},$$

where

- $\theta = (\theta_k)_{k\in\mathbb{Z}}$ is a centered stationary Gaussian sequence with unit variance and covariance
$$\rho(v) = \mathbb{E}[\theta_0 \theta_v], \quad v \in \mathbb{Z}.$$
- $g \in L^2(\mathbb{R}, \gamma)$ has Hermite rank $d \geq 1$.

Theorem 12.1 (Breuer and Major [3]). *Let $g \in L^2(\mathbb{R}, \gamma)$ with Hermite rank $d \geq 1$ and assume*

$$\sum_{k\in\mathbb{Z}} |\rho(k)|^d < \infty.$$

Then,

$$F_n := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} g(\theta_k) \xrightarrow{\mathcal{L}} N(0, \sigma^2),$$

as $n \to \infty$, where

$$\sigma^2 = \sum_{m=d}^{\infty} m! c_m^2 \sum_{k\in\mathbb{Z}} \rho(k)^m. \tag{94}$$

Note that $|\rho(k)| \le 1$, so

$$\sigma^2 \le \sum_{m=d}^{\infty} c_m^2 m! \sum_{k\in\mathbb{Z}} |\rho(k)|^d = \|g\|^2_{L^2(\mathbb{R},\gamma)} \sum_{k\in\mathbb{Z}} |\rho(k)|^d < \infty.$$

Proof. From the chaotic Central Limit Theorem, it suffices to consider the case $g = c_q H_q$, $q \ge d$. There exists a sequence $(e_k)_{k\ge1}$ in $H = L^2([0,T])$ such that for each $j, k \ge 1$,

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

Let $(B_t)_{t\in[0,T]}$ be a Brownian motion. The sequence $(B(e_k))_{k\ge1}$ has the same law as $(\theta_k)_{k\ge1}$, and we may replace $F_n = \frac{c_q}{\sqrt{n}} \sum_{k=1}^n H_q(\theta_k)$ by

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

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

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

and by the dominated convergence theorem

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

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

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

We have

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

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

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

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

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

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 $n \to \infty$. □

12.2 Convergence in law in $C([0,T])$

In the framework of the Breuer-Major theorem, we are interested in the convergence in law in the space $C([0,T])$.

Definition 12.2. *Let $Y_n := (Y_n(t))_{t \in [0,T]}$ be a sequence of continuous stochastic processes. We say that this sequence converges in law to a continuous process Y if for any continuous and bounded function $\varphi : C([0,T]) \to C([0,T])$, we have*

$$\lim_{n \to \infty} \mathbb{E}(\varphi(Y_n)) = \mathbb{E}(\varphi(Y)).$$

A basic ingredient to show convergence in law in an infinite dimensional space is the notion of tightness.

Definition 12.3. *A sequence of probabilities $(\nu_n)_{n \ge 1}$ on $C([0,T])$ is tight if for any $\epsilon > 0$ there exists a compact set $K_\epsilon \subset C([0,T])$ such that*

$$\sup_{n \ge 1} \nu_n(K_\epsilon^c) \le \epsilon.$$

By the Arzelà-Ascoli theorem, a set of functions $K \subset C([0,T])$ is compact if and only if :

(i) $\sup_{x \in K, t \in [0,T]} |x(t)| < \infty$.

(ii) K is an equicontinuous set: For all $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in K$,

$$|s-t| < \delta \Rightarrow |x(s) - x(t)| \le \epsilon.$$

This leads to the folliwing criterion for tightness on $C([0,T])$:

Proposition 12.1. *Consider a sequence of continuous stochastic processes $Y_n = (Y_n(t))_{t\in[0,T]}$ defined in a probability space $(\Omega, \mathcal{F}, P)$ and set*

$$\nu_n(B) := P(Y_n^{-1}(B)),$$

for any Borel set in $C([0,T])$. The sequence $\nu_n = P \circ Y_n^{-1}$ is tight if

(i) $\sup_{n\geq 1} P(|Y_n(0)| > a) \to 0$ *as* $a \uparrow \infty$.

(ii) For some $\alpha > 1$ and $p \geq 1$.

$$\mathbb{E}[|Y_n(t) - Y_n(s)|^p] \leq c|t-s|^\alpha,$$

It is well known that proving convergence in law in $C([0,T])$ requires showing convergence of the finite-dimensional distributions and proving tightness. This is the contents of the next theorem.

Theorem 12.4. *Let $Y_n = (Y_n(t))_{t\in[0,T]}$ be a sequence of continuous stochastic processes. Then Y_n converges in law to a continuous process Y if:*

(i) The finite-dimensional distributions of Y_n converge in law to those of Y. That is, for any $0 \leq t_1 < \cdots \leq t_M \leq T$, we have, as $n \to \infty$,

$$(Y_n(t_1), \dots, Y_n(t_M)) \xrightarrow{\mathcal{L}} (Y(t_1), \dots, Y(t_M))$$

(ii) The sequence of laws $\nu_n = P \circ Y_n^{-1}$ is tight.

In view of Proposition 12.1, a sufficient condition for (ii) is: $\sup_{n\geq 1} \mathbb{E}(|Y_n(0)|^p) < \infty$ and

$$\mathbb{E}[|Y_n(t) - Y_n(s)|^p] \leq c|t-s|^\alpha,$$

for some $\alpha > 1$ and $p \geq 1$,

12.3 Functional version of the Breuer-Major theorem

In the framework of the Breuer-Major theorem, define Y_n as the continuous process on $[0,T]$ obtained by linear interpolation from

$$Y_n(N/n) = \frac{1}{\sqrt{n}} \sum_{k=1}^N g(\theta_k), \quad N = 1, 2, \dots$$

The following result is a functional version of the Breuer-Major theorem.

Theorem 12.5 (Nourdin and Nualart [10]). *Let $\theta = (\theta_n)_{n\in\mathbb{Z}}$ be a centered Gaussian stationary sequence with unit variance and covariance ρ. Let $g \in L^2(\mathbb{R}, \gamma)$ with Hermite rank $d \geq 1$. Suppose that:*

(i) $\sum_{k\in\mathbb{Z}} |\rho(k)|^d < \infty$.

(ii) $g \in L^p(\mathbb{R},\gamma)$ *for some* $p > 2$.

Then, as $n \to \infty$,

$$(Y_n(t))_{t\in[0,T]} \xrightarrow{\mathcal{L}} \sigma B,$$

where B *is a Brownian motion, the convergence holds in law in* $C([0,T])$ *and* σ *is defined in (94).*

When the random variables θ_n are independent, this is the classical Donsker theorem: The random walk converges in distribution to the Brownian motion.

Remark: Chambers and Slud [4] and Ben-Hariz [1] proved the functional version of the Breuer-Major theorem under the following condition on the rate of the convergence of the coefficients:

(ii') $m! c_m^2 \le C\alpha^m, \quad m \ge d,$

for some $C > 0$ and $\alpha < 1$.

By the hypercontractivity property, (ii') implies that $g \in L^p(\mathbb{R},\gamma)$ for $2 < p < \frac{1}{\alpha} + 1$. In fact,

$$\begin{aligned} \|g\|_{L^p(\mathbb{R},\gamma)} &\le \sum_{m=d}^{\infty} |c_m| \, \|H_m\|_{L^p(\mathbb{R},\gamma)} \\ &\le \sum_{m=d}^{\infty} |c_m| (p-1)^{\frac{m}{2}} \sqrt{m!} \\ &\le \sqrt{C} \sum_{m=d}^{\infty} (\alpha(p-1))^{\frac{m}{2}} < \infty. \end{aligned}$$

The converse is not true. For instance, the function $g(x) = |x| - \sqrt{\frac{2}{\pi}}$ belongs to $L^p(\mathbb{R},\gamma)$ for all $p \ge 2$ and its Hermite coefficients satisfy

$$(2m)! c_{2m}^2 = \frac{2(2m)!}{\pi 2^{2m} (m!)^2 (2m-1)^2} \sim \frac{2}{(2m-1)^2 \pi \sqrt{\pi m}}.$$

Proof Theorem 12.5. For the convergence of the law of Y_n to the law of σB in $C([0,T])$, we need:

(A) Convergence in law of the finite-dimensional distributions: this follows again from the Fourth Moment Theorem. Here we only need $g \in L^2(\mathbb{R},\gamma)$ and condition (i).

(B) Tightness: This follows from the moment estimate:

$$\|Y_n(t) - Y_n(s)\|_{L^p(\Omega)} \le c|t-s|^{1/2}, \tag{95}$$

for some $p > 2$, which is proved using techniques of Malliavin calculus.

Notice that when $g = \sum_{m=d}^{M} c_m H_m$, by hypercontractivity,

$$\|Y_n(t) - Y_n(s)\|_{L^p(\Omega)} \le c_{M,p} \|Y_n(t) - Y_n(s)\|_{L^2(\Omega)}$$

and condition (ii) is not needed.

The proof of the moment estimates (95) will be done in several steps:

Step 1: Embedding the sequence $(\theta_k)_{k\ge 1}$ in the Wiener space.

Recall that $\theta = (\theta_n)_{n\in\mathbb{Z}}$ is a centered Gaussian stationary sequence with unit variance and covariance

$$\rho(k) = \mathbb{E}[\theta_n \theta_{n+k}], \quad k \in \mathbb{Z}.$$

Consider a sequence $e_k \in L^2([0,T])$ such that for each $j,k \in \mathbb{Z}$,

$$\rho(j-k) = \langle e_j, e_k \rangle_{L^2([0,T])} = \int_0^T e_j(s) e_k(s) ds.$$

Then,

$$(\theta_k)_{k\in\mathbb{Z}} \overset{\text{law}}{=} (B(e_k))_{k\in\mathbb{Z}}.$$

So, we can assume that $\theta_k = B(e_k) = \int_0^T e_k(s) dB_s$.

Step 2: Shift operator.

Let $g \in L^2(\mathbb{R}, \gamma)$ be a function of Hermite rank $d \ge 1$ and expansion

$$g(x) = \sum_{m=d}^{\infty} c_m H_m(x).$$

For any $1 \le k \le d$, consider the function g_k defined by

$$g_k(x) = \sum_{m=d}^{\infty} c_m H_{m-k}(x).$$

From $H_m = \delta^k H_{m-k}$, we deduce that

$$\delta^k g_k(x) = g(x).$$

Step 3: Representation of $g(\theta_i)$ as an iterated divergence.

We have

$$g(\theta_i) = \delta^d (g_d(\theta_i) e_i^{\otimes d}).$$

Proof:

$$g(\theta_i) = \sum_{m=d}^{\infty} c_m H_m(\theta_i).$$

Then, for every $m \geq d$, using that multiple stochastic integrals are iterated divergences, we obtain

$$\begin{aligned} H_m(\theta_i) &= H_m(B(e_i)) = I_m(e_i^{\otimes m}) = \delta^m(e_i^{\otimes m}) \\ &= \delta^d \left(\delta^{m-d}(e_i^{\otimes(m-d)}) e_i^{\otimes d} \right) = \delta^d \left(I_{m-d}(e_i^{\otimes(m-d)}) e_i^{\otimes d} \right) \\ &= \delta^d \left(H_{m-d}(\theta_i) e_i^{\otimes d} \right). \end{aligned}$$

Step 4: Regularization property of the shift.

Lemma 12.6. *If $g \in L^p(\mathbb{R}, \gamma)$ for some $p \geq 2$ with Hermite rank d, then, for any $i \in \mathbb{Z}$ and $k = 1, \dots, d$, $g_k(\theta_i) \in \mathbb{D}^{k,p}$.*

Proof. Suppose $d = 1$. We have

$$DL^{-1} H_m(\theta_i) = -\frac{1}{m} H_m'(\theta_i) e_i = -H_{m-1}(\theta_i) e_i.$$

This implies

$$g_1(\theta_i) e_i = -DL^{-1} g(\theta_i).$$

The result follows from the equivalence in L^p of the operators D and $(-L)^{1/2}$ (Meyer inequalities):

$$\mathbb{E}(\|D^2 L^{-1} g(\theta_i)\|_{L^2([0,T]^2)}^p) \leq c_p \mathbb{E}(|g(\theta_i)|^p).$$

□

Step 5: Proof of tightness using Malliavin calculus.

Assume $s = \frac{N_1}{n} \leq t = \frac{N_2}{n}$. We can write, using the continuity of the iterated divergence,

$$\begin{aligned} \|Y_n(t) - Y_n(s)\|_{L^p(\Omega)} &= \frac{1}{\sqrt{n}} \left\| \sum_{i=N_1+1}^{N_2} g(\theta_i) \right\|_{L^p(\Omega)} \\ &= \frac{1}{\sqrt{n}} \left\| \sum_{i=N_1+1}^{N_2} \delta^d \left(g_d(\theta_i) e_i^{\otimes d} \right) \right\|_{L^p(\Omega)} \\ &\leq c_{p,d} \sum_{k=0}^{d} \frac{1}{\sqrt{n}} \left\| \sum_{i=N_1+1}^{N_2} D^k \left(g_d(\theta_i) e_i^{\otimes d} \right) \right\|_{L^p(\Omega; L^2([0.T]^{k+d})} \\ &= c_{p,d} \sum_{k=0}^{d} \left\| \frac{1}{n} \sum_{i,j=N_1+1}^{N_2} \left\langle D^k(g_d(\theta_i)), D^k(g_d(\theta_j)) \right\rangle_{L^2([0,T]^k)} \rho(i-j)^d \right\|_{L^{\frac{p}{2}}(\Omega)}^{1/2}. \end{aligned}$$

Notice that, from Lemma 12.6,

$$\sup_j \mathbb{E}[\|D^k(g_d(\theta_i))\|_{L^2([0,T]^k)}^p] < \infty, \quad 0 \leq k \leq d$$

because $\mathbb{E}[|g(\theta_i)|^p] < \infty$. As a consequence, deduce

$$\|Y_n(t) - Y_n(s)\|_{L^p(\Omega)} \leq C \left(\frac{1}{n} \sum_{i,j=N_1+1}^{N_2} |\rho(i-j)|^d \right)^{1/2}.$$

Finally, the change of indices $(i,j) \to (i, j+h)$ leads to

$$\frac{1}{n} \sum_{i,j=N_1+1}^{N_2} |\rho(i-j)|^d \leq C \frac{N_2 - N_1}{n} \sum_{h \in \mathbb{Z}} |\rho(h)|^d = C(t-s).$$

This completes the proof of the moment estimates. □

12.4 Rate of convergence

Recall that $(\theta_k)_{k\in\mathbb{Z}}$ is a centered Gaussian stationary sequence with unit variance and covariance $\rho(k) = \mathbb{E}[\theta_n \theta_{n+k}]$, $k \in \mathbb{Z}$. Let

$$F_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} g(\theta_k),$$

where g has Hermite rank $d \geq 2$ and $\sum_{k\in\mathbb{Z}} |\rho(k)|^d < \infty$. Then, we know that

$$F_n := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} g(\theta_k) \xrightarrow{\mathcal{L}} N(0, \sigma^2),$$

where $\sigma^2 = \sum_{m=d}^{\infty} c_m^2 m! \sum_{k\in\mathbb{Z}} \rho(k)^m$. Assume $\sigma^2 > 0$. Define $\sigma_n^2 = \mathrm{Var}(F_n)$ and consider the normalized sequence

$$Z_n = \frac{F_n}{\sigma_n}.$$

We are interested in the rate of convergence to zero of $d_{TV}(Z_n, Z)$ where $Z \sim N(0,1)$. A basic estimated is given by the following proposition (see (2)).

Proposition 12.2. *Suppose that $F \in \mathbb{D}^{1,2}$ satisfies $F = \delta(u)$, where u belongs to* Domδ *and $\mathbb{E}[F^2] = 1$. Then, if $Z \sim N(0,1)$,*

$$d_{TV}(F, Z) \leq 2\sqrt{\mathrm{Var}\langle DF, u\rangle_H}.$$

We will discuss here only the case of Hermite rank $d = 2$. Consider the particular function $g(x) = H_2(x) = x^2 - 1$. With the assumption $\theta_k = B(e_k)$, we can write

$$Z_n = \frac{1}{\sigma_n \sqrt{n}} \sum_{j=1}^{n} [B^2(e_j) - 1] = \frac{1}{\sigma_n \sqrt{n}} \sum_{j=1}^{n} \delta(B(e_j) e_j).$$

So, $Z_n = \delta(u_n)$, where

$$u_n = \frac{1}{\sigma_n \sqrt{n}} \sum_{j=1}^{n} B(e_j) e_j.$$

Then,

$$DZ_n = \frac{2}{\sigma_n \sqrt{n}} \sum_{j=1}^{n} B(e_j) e_j$$

and

$$\langle DZ_n, u_n \rangle_H = \frac{2}{n\sigma_n^2} \sum_{i,j=1}^{n} B(e_i) B(e_j) \rho(i-j).$$

As a consequence

$$\begin{aligned}
&\mathrm{Var}(\langle DZ_n, u_n \rangle_H) \\
&= \frac{4}{n^2 \sigma_n^4} \sum_{i_1,i_2,i_3,i_4=1}^{n} \mathrm{Cov}(B(e_{i_1})B(e_{i_2}), B(e_{i_3})B(e_{i_4})) \rho(i_1 - i_2)\rho(i_3 - i_4) \\
&= \frac{4}{n^2 \sigma_n^4} \sum_{i_1,i_2,i_3,i_4=1}^{n} \big[\rho(i_1 - i_3)\rho(i_2 - i_4)\rho(i_1 - i_2)\rho(i_3 - i_4) \\
&\quad + \rho(i_1 - i_4)\rho(i_2 - i_3)\rho(i_1 - i_2)\rho(i_3 - i_4) \big] \\
&= V_n^{(1)} + V_n^{(2)}.
\end{aligned}$$

Both summands are similar and can be treated in the same way. Focusing only on the first summand, we make the change of variable $i_1 - i_2 = k_1$, $i_3 - i_4 = k_2$ and $i_1 - i_3 = k_3$ and we obtain

$$V_n^{(1)} \leq \frac{C}{n} \sum_{|k_i| \leq n, i=1,2,3} |\rho(k_1)\rho(k_2)\rho(k_3)\rho(k_2 + k_3 - k_1)|.$$

Set $\rho_n(k) = |\rho(k)| \mathbf{1}_{\{|k| \leq n\}}$. Then, by Hölder's inequality,

$$\begin{aligned}
V_n^{(1)} &\leq \frac{C}{n} \langle \rho_n, (\rho_n * \rho_n) * \rho_n \rangle_{\ell^2(\mathbb{Z})} \\
&\leq \frac{C}{n} \|\rho_n\|_{\ell^{4/3}(\mathbb{Z})} \|(\rho_n * \rho_n) * \rho_n\|_{\ell^4(\mathbb{Z})}.
\end{aligned}$$

Applying Young's convolution inequality, yields

$$V_n^{(1)}) \leq \frac{C}{n} \|\rho_n\|_{\ell^{4/3}(\mathbb{Z})}^4 \leq \frac{C}{n} \left(\sum_{|k| \leq n} |\rho(k)|^{\frac{4}{3}} \right)^3. \tag{96}$$

Therefore in the case $g = H_2$ we have proved the following estimate

$$d_{TV}(Z_n, Z) \leq \frac{C}{\sqrt{n}} \left(\sum_{|k| \leq n} |\rho(k)|^{\frac{4}{3}} \right)^{\frac{3}{2}}.$$

To handle the case of a general function g, let us recall the definition of the Sobolev spaces in the context of the one-dimensional Gaussian analysis. For any $m \geq 1$ and $p \geq 2$, $\mathbb{D}^{m,p}(\mathbb{R}, \gamma)$ is the class of functions $g \in L^p(\mathbb{R}, \gamma)$ which are m times weakly differentiable and $g^{(k)} \in L^p(\mathbb{R}, \gamma)$ for all $1 \leq k \leq m$.

Theorem 12.7 (Nourdin, Nualart and Peccati [11])**.** *If $g \in \mathbb{D}^{1,4}$ has Hermite rank 2, then*

$$d_{TV}(Z_n, Z) \leq \frac{C}{\sqrt{n}} \left(\sum_{|k| \leq n} |\rho(k)| \right)^{\frac{1}{2}} + \frac{C}{\sqrt{n}} \left(\sum_{|k| \leq n} |\rho(k)|^{\frac{4}{3}} \right)^{\frac{3}{2}}. \tag{97}$$

Some ingredients of the proof of Theorem 12.7: We can represent $g(B(e_i))$ as

$$g(B(e_i)) = \delta(g_1(B(e_i))e_i),$$

where $g_1(x) = \sum_{m=2}^{\infty} c_m H_{m-1}(x)$. With this representation, we can write $Z_n = \delta(u_n)$, where

$$u_n = \frac{1}{\sigma_n \sqrt{n}} \sum_{i=1}^{n} g_1(B(e_i))e_i$$

and

$$\langle DZ_n, u_n \rangle_H = \frac{1}{n\sigma_n^2} \sum_{i,j=1}^{n} g'(B(e_i))g_1(B(e_j))\rho(i-j).$$

Then, we can write

$$\begin{aligned}\operatorname{Var}(\langle DZ_n, u_n \rangle_H) = \frac{1}{n^2 \sigma_n^4} \sum_{i_1,i_2,i_3,i_4=1}^{n} &\operatorname{Cov}\left(g'(B(e_{i_1}))g_1(B(e_{i_2})), g'(B(e_{i_3}))g_1(B(e_{i_4}))\right) \\ &\times \rho(i_1 - i_2)\rho(i_3 - i_4),\end{aligned} \tag{98}$$

and use Gebelein's inequality (Theorem 12.8 below) to estimate the covariances.

Theorem 12.8 (Gebelein inequality)**.** *Let H_1 and H_2 be subspaces of H and denote by B_1 and B_2 the restrictions of B to H_1 and H_2, respectively. Consider centered and square integrable functionals $F_i(B_i)$, $i = 1, 2$. Then,*

$$\begin{aligned}|\mathbb{E}[F_1(B_1)F_2(B_2)]| \leq &\sup_{\varphi_1 \in H_1, \varphi_2 \in H_2, \|\varphi_1\|_H = \|\varphi_2\|_H = 1} |\langle \varphi_1, \varphi_2 \rangle_H| \\ &\times \sqrt{\operatorname{Var}(F_1(B_1))\operatorname{Var}(F_2(B_2))}.\end{aligned}$$

Applying Gebelein's inequality in (98), we have

$$\begin{aligned}\operatorname{Var}(\langle DZ_n, u_n \rangle_H) = \frac{C}{n^2} &\sum_{i_1,i_2,i_3,i_4=1}^{n} |\rho(i_1 - i_2)\rho(i_3 - i_4)| \\ &\times \max\{|\rho(i_1 - i_3)|, |\rho(i_1 - i_4)|, |\rho(i_2 - i_3)|, |\rho(i_2 - i_4)|\} \\ &\leq \frac{C}{n} \sum_{|k_i| \leq n, i=1,2,3} |\rho(k_1)\rho(k_2)\rho(k_3)| = \frac{C}{n} \left(\sum_{|k_|\leq n} |\rho(k)| \right)^3.\end{aligned}$$

This gives the rate

$$d_{TV}(Z_n, Z) \leq \frac{C}{\sqrt{n}} \left(\sum_{|k| \leq n} |\rho(k)| \right)^{\frac{3}{2}},$$

which was proved by Nourdin, Peccati and Yang in [14]. The proof of (97) is more involved (see [11]]).

Example: Suppose that $g(x) = |x|^p - \mathbb{E}[|Z|^p]$, for any $p \geq 1$. Consider the particular case where $\rho(k) \sim k^{-\alpha}$ for some $\alpha > 0$. Then, condition $\sum_{k \in \mathbb{Z}} \rho(k)^2 < \infty$ means $\alpha > \frac{1}{2}$, and the rate is

$$d_{TV}(Z_n, Z) \leq C \times \begin{cases} n^{-\frac{1}{2}} & \text{if } \alpha > 1 \\ n^{-\frac{1}{2}} \sqrt{\log n} & \text{if } \alpha = 1 \\ n^{-\frac{\alpha}{2}} & \text{if } \alpha \in [\frac{2}{3}, 1) \\ n^{1-2\alpha} & \text{if } \alpha \in (\frac{1}{2}, \frac{2}{3}). \end{cases}$$

In [13] the following optimal version of the fourth moment theorem was derived.

Theorem 12.9. *Consider a sequence of random variables $(F_n)_{n \geq 1}$ in the chaos $\mathcal{H}_q$, where $q \geq 2$. Suppose that $\mathbb{E}(F_n^2) = 1$. Let X be a $N(0,1)$ random variable. Then, there exist positive constants c and C such that*

$$cM(F_n) \leq d_{\mathrm{TV}}(F_n, X) \leq CM(F_n),$$

where $M(F_n) = \max(|\mathbb{E}[F_n^3]|, \mathbb{E}[F_n^4] - 3)$.

In the famework of the Breuer-Major theorem, this yields (see Biermé, Bonami, Nourdin and Peccati [**?**]):

$$d_{TV}(F_n, Z) \sim \frac{c}{\sqrt{n}} \left(\sum_{|k| \leq n} |\rho(k)|^{\frac{3}{2}} \right)^2.$$

12.5 Fractional Brownian motion

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

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

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

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

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

For $\mathbf{H} = \frac{1}{2}$, $B^{\frac{1}{2}}$ 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 $(a^{-\mathbf{H}} B^{\mathbf{H}}_{at})_{t\geq 0}$ and $(B^{\mathbf{H}}_t)_{t\geq 0}$ have the same probability distribution (they are fractional Brownian motions with Hurst parameter $\mathbf{H}$).
2) Unlike Brownian motion, the fractional Brownian motion has correlated increments. More precisely, for $\mathbf{H} \neq \frac{1}{2}$, we can write

$$\begin{aligned}\rho(n) &= \mathbb{E}(B^{\mathbf{H}}_1(B^{\mathbf{H}}_{n+1} - B^{\mathbf{H}}_n)) = \frac{1}{2}\left((n+1)^{2\mathbf{H}} + (n-1)^{2\mathbf{H}} - 2n^{2\mathbf{H}}\right) \\ &\sim \mathbf{H}(2\mathbf{H}-1)n^{2\mathbf{H}-2},\end{aligned}$$

as $n \to \infty$.

(ii) If $\mathbf{H} > \frac{1}{2}$, then $\rho(n) > 0$ and $\sum_n \rho(n) = \infty$ (*long memory*).

(iii) If $\mathbf{H} < \frac{1}{2}$, then $\rho(n) < 0$ (*intermittency*) and $\sum_n |\rho(n)| < \infty$.

3) The fractional Brownian motion has finite $\frac{1}{\mathbf{H}}$-variation: Fix $T > 0$. Set $t_i = \frac{iT}{n}$ for $1 \leq i \leq n$ and define $\Delta B^{\mathbf{H}}_{t_i} = B^{\mathbf{H}}_{t_i} - B^{\mathbf{H}}_{t_{i-1}}$. Then, as $n \to \infty$,

$$\sum_{i=1}^n |\Delta B^{\mathbf{H}}_{t_i}|^{\frac{1}{\mathbf{H}}} \overset{L^2(\Omega),a.s.}{\longrightarrow} c_{\mathbf{H}} T,$$

where $c_{\mathbf{H}} = \mathbb{E}[|B^{\mathbf{H}}_1|^{\frac{1}{\mathbf{H}}}]$.

Proof. By the self-similarity, $\sum_{i=1}^n |\Delta B^{\mathbf{H}}_{t_i}|^{\frac{1}{\mathbf{H}}}$ has the same law as

$$\frac{T}{n}\sum_{i=1}^n |B^{\mathbf{H}}_i - B^{\mathbf{H}}_{i-1}|^{\frac{1}{H}}.$$

The sequence $(B^{\mathbf{H}}_i - B^{\mathbf{H}}_{i-1})_{i\geq 1}$ is stationary and ergodic. Therefore, the Ergodic Theorem implies the desired convergence. □

Exercises

1. Let $B^{\mathbf{H}} = (B^{\mathbf{H}}_t)_{t\geq 0}$ be a fractional Brownian motion of Hurst parameter $\mathbf{H}$. Define $\theta_k = B^{\mathbf{H}}_k - B^{\mathbf{H}}_{k-1}$ (*fractional noise*). Show the following statements:

   (i) The sequence $(\theta_k)_{k\geq 1}$ is Gaussian, stationary and centered with covariance

   $$\rho_{\mathbf{H}}(k) = \frac{1}{2}\left(|k+1|^{2\mathbf{H}} + |k-1|^{2\mathbf{H}} - 2|k|^{2\mathbf{H}}\right).$$

   (ii) For any integer $d \geq 2$ such that $\mathbf{H} < 1 - \frac{1}{2d}$, we have

   $$\sum_{v\in\mathbb{Z}} |\rho_{\mathbf{H}}(v)|^d < \infty.$$

(ii) By the Breuer-Major theorem if $g \in L^p(\mathbb{R}, \gamma)$, $p > 2$, has Hermite rank d and $\mathbf{H} < 1 - \frac{1}{2d}$

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} g(n^{\mathbf{H}}(B^{\mathbf{H}}_{\frac{k}{n}} - B^{\mathbf{H}}_{\frac{k-1}{n}})) \xrightarrow{\mathcal{L}} \sigma_{\mathbf{H}} B,$$

where $\sigma_{\mathbf{H}}^2 = \sum_{m=d}^{\infty} c_m^2 m! \sum_{k \in \mathbb{Z}} \rho_{\mathbf{H}}(k)^m$ and B is a Brownian motion.

2. Let $B^{\mathbf{H}} = (B_t^{\mathbf{H}})_{t \geq 0}$ be a fractional Brownian motion of Hurst parameter $\mathbf{H}$. For a real $q \geq 1$, set $c_q = \mathbb{E}[|Z|^q]$, where $Z \sim N(0,1)$. Show that the normalized p-variation of the fractional Brownian motion satisfies

$$n^{p\mathbf{H}-1} \sum_{k=1}^{n} |B^{\mathbf{H}}_{\frac{k}{n}} - B^{\mathbf{H}}_{\frac{k-1}{n}}|^p \overset{\mathcal{L}}{=} \frac{1}{n} \sum_{k=1}^{n} |B^{\mathbf{H}}_{k} - B^{\mathbf{H}}_{k-1}|^p \xrightarrow{a.s.} c_p.$$

3. Let $B^{\mathbf{H}} = (B_t^{\mathbf{H}})_{t \geq 0}$ be a fractional Brownian motion of Hurst parameter $\mathbf{H}$. If $\mathbf{H} < \frac{3}{4}$, show that

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} \left[n^{p\mathbf{H}} |B^{\mathbf{H}}_{\frac{k}{n}} - B^{\mathbf{H}}_{\frac{k-1}{n}}|^p - c_p \right] \xrightarrow{\mathcal{L}} \sigma_{\mathbf{H},p} B,$$

where B is a Brownian motion.

4. Let $B^{\mathbf{H}} = (B_t^{\mathbf{H}})_{t \geq 0}$ be a fractional Brownian motion of Hurst parameter $\mathbf{H}$. Show that the sequence $Z_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \left[(B^{\mathbf{H}}_{k} - B^{\mathbf{H}}_{k-1})^2 - 1 \right]$, with $\mathrm{Var}(Z_n) = 1$, satisfies:

$$d_{TV}(Z_n, Z) \sim c_H \times \begin{cases} n^{-\frac{1}{2}} & \text{if } \mathbf{H} \in (0, \frac{2}{3}) \\ n^{-\frac{1}{2}} (\log n)^2 & \text{if } \mathbf{H} = \frac{2}{3} \\ n^{6H - \frac{9}{2}} & \text{if } \mathbf{H} \in (\frac{2}{3}, \frac{3}{4}). \end{cases}$$

13 Spatial averaging of stochastic partial differential equations

13.1 Stochastic heat equation

Consider the one-dimensional stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \sigma(u)\frac{\partial^2 W}{\partial t \partial x}, \quad x \in \mathbb{R}, \quad t \geq 0,$$

with initial condition $u_0(x) = 1$, driven by a space-time white noise $\frac{\partial^2 W}{\partial t \partial x}$. We assume that σ is a Lipschitz function such that $\sigma(1) \neq 0$.

The space-time white noise is formally defined as an isonormal Gaussian process W on the Hilbert space $H = L^2(\mathbb{R}_+ \times \mathbb{R})$. Writing $W(A) = W(\mathbf{1}_A)$ for $A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$ with $|A| < \infty$, we obtain a Gaussian centered family of random variables $\{W(A), A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}), |A| < \infty\}$ with covariance given by

$$\mathbb{E}[W(A)W(B)] = |A \cap B|.$$

We will make use of the notation

$$W(h) = \int_{\mathbb{R}_+ \times \mathbb{R}} h(t,x) W(dt, dx),$$

Formally, $\frac{\partial^2 W}{\partial t \partial x}$ is the density of the random measure W with respect to the Lebesgue measure. Moreover, $W(s,t) = W([0,s] \times [0,t])$, $s,t \geq 0$ is a two-parameter Brownian motion.

13.2 Stochastic integration

For any $t \geq 0$, let $\mathcal{F}_t$ be the σ-algebra generated by $\{W(A), A \in \mathcal{B}([0,t] \times \mathbb{R})\}$. A random field $u = \{u(t,x), t \geq 0, x \in \mathbb{R}\}$ is *adapted* if $u(t,x)$ is $\mathcal{F}_t$-measurable for each $t \geq 0$ and $x \in \mathbb{R}$. We can define the Itô-Walsh stochastic integral of adapted and measurable processes u such that

$$\int_{\mathbb{R}_+ \times \mathbb{R}} \mathbb{E}(u^2(t,x)) dt dx < \infty,$$

in such a way that the following isometry property holds:

$$\mathbb{E}\left(\left| \int_{\mathbb{R}_+ \times \mathbb{R}} u(t,x) W(dt, dx) \right|^2 \right) = \int_{\mathbb{R}_+ \times \mathbb{R}} \mathbb{E}(u^2(t,x)) dt dx.$$

Suppose that δ is the divergence operator in the framework of the Malliavin calculus for the isonormal Gaussian process $\{W(h), h \in H\}$. Then, if $v \in L^2(\Omega \times \mathbb{R}_+ \times \mathbb{R})$ is a square integrable adapted random field, v belongs to the domain of δ and $\delta(v)$ coincides with the Itô-Walsh integral of v:

$$\delta(v) = \int_{\mathbb{R}_+ \times \mathbb{R}} v(s,y) W(ds, dy).$$

13.3 Mild solution

The following is the fundamental result on the existence and uniqueness of a mild solution to the stochastic heat equation.

Theorem 13.1 (Walsh [25]). *There is a unique mild solution, which is an adapted random field u such that for all $p \geq 2$,*

$$\sup_{x \in \mathbb{R}} \sup_{0 \leq t \leq T} \mathbb{E}[|u(t,x)|^p] < \infty, \tag{99}$$

and u satisfies the integral equation:

$$u(t,x) = 1 + \int_0^t \int_{\mathbb{R}} p_{t-s}(x-y)\sigma(u(s,y))W(ds,dy),$$

where $p_t(x) = \frac{1}{\sqrt{2\pi t}} \exp(-x^2/2t)$.

Moreover, we have

$$\sup_{t \in [0,T]} \mathbb{E}(|u(t,x) - u(s,y)|^p) \leq C_{T,p}(|s-t|^{p/4} + |x-y|^{p/2}).$$

13.4 Malliavin differentiability of the solution

We will first show that the random variable $u(t,x)$ belongs to the space $\mathbb{D}^{1,p}$ and obtain bounds for the p-moments of the derivative.

Proposition 13.1. *For each $(t,x) \in \mathbb{R}_+ \times \mathbb{R}$, $u(t,x) \in \mathbb{D}^{1,p}$ for all $p \geq 2$ and the derivative $D_{s,y}u(t,x)$ satisfies*

(i) *$D_{s,y}u(t,x) = 0$ if $s > t$ and for $s \leq t$:*

$$\begin{aligned} D_{s,y}u(t,x) &= p_{t-s}(x-y)\sigma(u(s,y)) \\ &\quad + \int_s^t \int_{\mathbb{R}} p_{t-r}(x-z)\Sigma(r,z)D_{s,y}u(r,z)W(dr,dz) \end{aligned}$$

where $\Sigma(r,z)$, $(r,z) \in \mathbb{R}_+ \times \mathbb{R}$, is an adapted and bounded process.

(ii) *For all $t \in [0,T]$, and $p \geq 2$,*

$$.\|D_{s,y}u(t,x)\|_p \leq C_{T,p}\, p_{t-s}(x-y) \tag{100}$$

for almost all $(s,y) \in [0,t] \times \mathbb{R}$.

Remark: If the coefficient σ is of class C^1 with bounded derivative, then $\Sigma(r,z) = \sigma'(u(r,z))$.

Proof. The proof will be done in several steps:

Step 1: Consider the Picard approximations defined by $u_0(t,x)=1$, and for each integer $n\geq 0$ set

$$u_{n+1}(t,x):=1+\int_0^t\int_{\mathbb{R}}p_{t-s}(x-y)\sigma(u_n(s,y))W(ds,dy). \tag{101}$$

We know that, for each $p\geq 2$,

$$\lim_{n\to\infty}\mathbb{E}(|u_n(t,x)-u(t,x)|^p)=0.$$

and also $\sup_n\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}}\mathbb{E}(|u_n(t,x)|^p)<\infty$.

By induction we will show that for each $(t,x)\in[0,T]\times\mathbb{R}$ and for each $p\geq 2$, $u_n(t,x)\in\mathbb{D}^{1,p}$ and

$$\|D_{s,y}u_n(t,x)\|_p\leq C_{T,p}p_{t-s}(x-y). \tag{102}$$

for almost all $(s,y)\in[0,t]\times\mathbb{R}$.

Suppose that the induction hypothesis holds for all integers less than or equal to n. By the chain rule for Lipschitz functions (see Proposition 65), for each $(t,x)\in[0,T]\times\mathbb{R}$, $\sigma(u_n(t,x))$ belongs to $\mathbb{D}^{1,p}$ and there exists a random variable Σ_n bounded by the Lipschitz constant of σ, such that

$$D(\sigma(u_n(t,x)))=\Sigma_n Du_n(t,x).$$

If σ is continuously differentiable, then $\Sigma_n=\sigma'(u_n(t,x))$. Then, applying the properties of the divergence operator (see (31), we deduce that $\int_0^t\int_{\mathbb{R}}p_{t-s}(x-y)\sigma(u_n(s,y))W(ds,dy)$ is in $\mathbb{D}^{1,p}$ and

$$\begin{aligned}D_{s,y}u_{n+1}(t,x)&=D_{s,y}\left(\int_0^t\int_{\mathbb{R}}p_{t-r}(x-z)\sigma(u_n(r,z))W(dr,dz)\right)\\&=p_{t-s}(x-y)\sigma(u_n(s,y))+\int_r^t\int_{\mathbb{R}}p_{t-r}(x-z)\Sigma_n D_{s,y}u_n(r,z)W(dr,dz)\end{aligned}$$

Now, let us estimate the L^p norm of $D_{s,y}u_{n+1}(t,x)$.

$$\begin{aligned}\left\|D_{s,y}u_{n+1}(t,x)\right\|_p&\leq\sup_{r,z}\|\sigma(u_n(r,z))\|_p p_{t-s}(x-y)\\&\quad+\left\|\int_s^t\int_{\mathbb{R}}p_{t-r}(x-z)\Sigma_n D_{s,y}u_n(r,z)W(dr,dz)\right\|_p.\end{aligned}$$

The Burkholder-David-Gundy and Minkowski's inequalities yield, with a constant c_p,

$$\begin{aligned}&\mathbb{E}\left(\left|\int_s^t\int_{\mathbb{R}}p_{t-r}(x-z)\Sigma_n D_{s,y}u_n(r,z)W(dr,dz)\right|^p\right)\\&\leq c_p^p\mathrm{Lip}(\sigma)^p\mathbb{E}\left(\left|\int_s^t\int_{\mathbb{R}}p_{t-r}^2(x-z)|D_{s,y}u_n(r,z)|^2dzdr\right|^{p/2}\right)\\&\leq c_p^p\mathrm{Lip}(\sigma)^p\left(\int_s^t\int_{\mathbb{R}}p_{t-r}^2(x-z)\|D_{s,y}u_n(r,z)\|_p^2dzdr\right)^{p/2},\end{aligned}$$

where $\mathrm{Lip}(\sigma)$ denotes the Lipschitz constant of σ. Therefore, the preceding displayed computation yields that

$$\|D_{s,y}u_{n+1}(t,x)\|_p^2 \leq K_{p,T}\, p_{t-s}^2(x-y) + K_{p,\sigma} \int_s^t \int_{\mathbb{R}} p_{t-r}^2(x-z) \|D_{s,y}u_n(r,z)\|_p^2 dzdr,$$

where

$$K_{p,T} := 2\sup_n \sup_{(s,y)\in[0,T]\times\mathbb{R}^d} \|\sigma(u_n(s,y))\|_p^2 < \infty.$$

and

$$K_{p,\sigma} = 2c_p^2 \mathrm{Lip}(\sigma)^2$$

By iterating this inequality, yields

$$\begin{aligned}&\|D_{s,y}u_{n+1}(t,x)\|_p^2\\ \leq& K_{p,T}p_{t-s}^2(x-y) + \sum_{k=1}^n K_{p,T}K_{p,\sigma}^k \int_{\Delta_k(s,t)} \int_{\mathbb{R}^k} \prod_{j=1}^{k+1} p_{r_{j-1}-r_j}^2(z_{j-1}-z_j) d\mathbf{z}_k d\mathbf{r}_k,\end{aligned}$$

with the convention $r_0 = t$, $z_0 = x$, $r_{k+1} = s$, $z_{k+1} = y$. Here $d\mathbf{z}_k$ means $dz_1 \cdots dz_k$, $d\mathbf{r}_k$ means $dr_1 \cdots dr_k$ and $\Delta_k(s,t)$ denotes the simplex $\{s < r_k < \cdots < r_1 < t\}$. We have also used the fact that $u_0(t,x) = 1$ so $D_{s,y}u_0(t,x) = 0$. Then, using the fact that

$$p_{r-s}^2(z) = \frac{1}{2\sqrt{\pi}}(t-s)^{-1/2} p_{\frac{r-s}{2}}(z)\,,$$

and

$$\int_{\mathbb{R}} p_s(x-y)p_t(y-z)dy = p_{s+t}(x-z)\,,$$

we have for the integrals of the product of heat kernels above,

$$\begin{aligned}&\int_{\Delta_k(s,t)} \int_{\mathbb{R}^k} \prod_{j=1}^{k+1} p_{r_{j-1}-r_j}^2(z_{j-1}-z_j) d\mathbf{z}_k d\mathbf{r}_k\\ &\leq (2\sqrt{\pi})^{-k} \int_{\Delta_k(s,t)} \int_{\mathbb{R}^k} \prod_{j=1}^{k+1} (r_{j-1}-r_j)^{-\frac{1}{2}} p_{(r_{j-1}-r_j)/2}(z_{j-1}-z_j) d\mathbf{z}_k d\mathbf{r}_k\\ =&(2\sqrt{\pi})^{-k} \frac{(t-s)^{\frac{k-1}{2}}\Gamma(\frac{1}{2})^{k+1}}{\Gamma(\frac{k+1}{2})} p_{\frac{t-s}{2}}(x-y).\end{aligned}$$

This allows us to complete the proof of (102).

Step 2: We know that $u_n(t,x)$ converges in $L^p(\Omega)$ to $u(t,x)$. The estimate (102) implies that

$$\sup_n \mathbb{E}(\|Du_n(t,x)\|_H^p) < \infty.$$

By Proposition 6.4, this implies that $u(t,x)$ belongs to $\mathbb{D}^{1,p}$. Applying the operator D to the equation satisfied by $u(t,x)$ we deduce point (i).

Step 3: Finally, it remains to show that the estimate (102) holds for $u(t,x)$. This follows from the fact that $Du_n(t,x)$ (by choosing a subsequence) converges in the weak topology of $L^p(\Omega; H)$ to $Du(t,x)$. $\square$

13.5 Space averages

Fix $t \geq 0$. We are interested in the asymptotic behavior as $R \to \infty$ of the random variable

$$\int_{-R}^{R} u(t,x)dx$$

Because the process $x \to u(t,x)$ is stationary and the covariance decays fast, we expect a central limit theorem to hold for these space averages.

The mean of this variable is given by

$$\mathbb{E}\left(\int_{-R}^{R} u(t,x)dx\right) = 2R.$$

Define

$$F_R = \int_{-R}^{R} u(t,x)d - 2R = \int_0^t \int_{\mathbb{R}} \left(\int_{-R}^{R} p_{t-s}(x-y)dx\right) \sigma(u(s,y))W(ds,dy).$$

Set

$$f_R(s,y) = \int_{-R}^{R} p_{t-s}(x-y)dx.$$

Notice that $f_R(s,y) \leq 1$ and

$$\begin{aligned}\int_{\mathbb{R}} f_R^2(s,y)dy &= \int_{\mathbb{R}} \int_{[-R,R]^2} p_{t-s}(x-y)p_{t-s}(x'-y)dxdx'dy \\ &= \int_{[-R,R]^2} p_{2(t-s)}(x-x')dxdx' \\ &\leq 2R. \end{aligned} \tag{103}$$

Then, using the isometry property of the stochastic integral we can compute the variance σ_R^2 of F_R:

$$\begin{aligned}\sigma_R^2 &:= \int_0^t \int_{\mathbb{R}} \left(\int_{-R}^{R} p_{t-s}(x-y)dx\right)^2 \mathbb{E}[\sigma(u(s,y))^2]dyds \\ &= \int_0^t \mathcal{E}(s) \int_{\mathbb{R}} f_R^2(s,y)dyds \\ &= \int_0^t \mathcal{E}(s) \int_{[-R,R]^2} p_{2(t-s)}(x-x')dxdx'ds \\ &\approx 2R \int_0^t \mathcal{E}(s)ds,\end{aligned}$$

as $R \to \infty$, with $\mathcal{E}(s) = \mathbb{E}[\sigma(u(s,y))^2]$.

The main result is the following quantitative version of the central limit theorem.

Theorem 13.2 (Huang, Nualart and Viitasaari [7])**.** *Let $Z \sim N(0,1)$. Then there exists a constant C, depending on t, such that*

$$d_{TV}\left(\frac{F_R}{\sigma_R}, Z\right) \le \frac{C}{\sqrt{R}},$$

where $Z \sim N(0,1)$.

Proof. The proof will be done in several steps.

Step (i) (see (1): By the Malliavin-Stein's approach, if $F_R \in \mathbb{D}^{1,2}$ satisfies $F_R = \delta(v_R)$, where $v_R \in \mathrm{Dom}\delta$. Then,

$$d_{TV}(\frac{F_R}{\sigma_R}, Z) \le \frac{2}{\sigma_R^2}\sqrt{\mathrm{Var}(\langle DF_R, v_R\rangle_H)}.$$

We obtain the representation

$$F_R = \delta(v_R),$$

by choosing

$$v_R(s,y) = \sigma(u(s,y)) f_R(s,y) \mathbf{1}_{[0,t]}(s),$$

taking into account that the stochastic integral of adapted and square integrable random fields coincides with the divergence operator.

Step (ii) In order to estimate the variance of $\langle DF_R, v_R\rangle_H)$ we use the following proposition.

Proposition 13.2. *Suppose that*

$$F = 1 + \int_{\mathbb{R}_+\times\mathbb{R}^d} V_{s,y} W(ds,dy),$$

where V is an adapted random field such that $V \in \mathbb{D}^{1,4}(H)$. Then

$$\mathrm{Var}(\langle DF, V\rangle_H) \le 4\mathbb{E}\left(\|DV_\bullet \otimes_1 V_\bullet\|_H^2\right) + 2\mathbb{E}\left(\|D_\bullet V \otimes_1 V_\bullet\|_H^2\right),$$

where $\otimes_1$ means the contraction of one index (the bullet).

Proof: Using that $D(\delta V) = V + \delta(DV)$, we get

$$D_{r,z}F = V_{r,z} + \int_{\mathbb{R}_+\times\mathbb{R}^d} D_{r,z}V_{s,y} W(ds,dy)$$

and

$$\langle DF, V\rangle_H = \|V\|_H^2 + \int_{\mathbb{R}_+\times\mathbb{R}^d} (D_\bullet V_{s,y} \otimes_1 V_\bullet) W(ds,dy).$$

The isometry property of the stochastic integral yields

$$\begin{aligned}\mathrm{Var}(\langle DF, V\rangle_H) &\le 2\mathrm{Var}\left(\|V\|_H^2\right) + 2\mathrm{Var}\left(\int_{\mathbb{R}_+\times\mathbb{R}^d} (D_\bullet V_{s,y}\otimes_1 V_\bullet) W(ds,dy)\right)\\ &= 2\mathrm{Var}\left(\|V\|_H^2\right) + 2E\left(\|D_\bullet V\otimes_1 V_\bullet\|_H^2\right).\end{aligned}$$

To estimate the variance of $\|V\|_H^2$ we use *Poincaré's inequality*:

$$\mathrm{Var}\left(\|V\|_H^2\right) \le \mathbb{E}\left(\|D\left(\|V\|_H^2\right)\|_H^2\right) = 4\mathbb{E}\left(\|DV_\bullet \otimes_1 V_\bullet\|_H^2\right).$$

This completes the proof.

Step (iii) Estimation of $\mathbb{E}\left(\|Dv_{R,\bullet}\otimes_1 v_{R,\bullet}\|_H^2\right)$: Assuming, to simplify, that σ is continuously differentiable, we have

$$D_{r,z}v_R(s,y) = \sigma'(u(s,y))D_{r,z}u(s,y)f_R(s,y)\mathbf{1}_{\{0\le r\le s\le t\}}.$$

As a consequence,

$$\begin{aligned}\|Dv_{R,\bullet}\otimes_1 v_{R,\bullet}\|_H^2 &= \int_{[0,t]\times\mathbb{R}}\left(\int_{[0,t]\times\mathbb{R}} (\sigma'\sigma)(u(s,y))D_{r,z}u(s,y)f_R^2(s,y)dsdy\right)^2 drdz\\ &\le t(\mathrm{Lip}\sigma)^2\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}\sigma(u(s,y))D_{r,z}u(s,y)f_R^2(s,y)dy\right)^2 dsdrdz.\end{aligned}$$

Taking the mathematical expectation, Minkowski's inequality yields

$$\mathbb{E}(Dv_{R,\bullet}\otimes_1 v_{R,\bullet}\|_H^2 \le t(\mathrm{Lip}\sigma)^2\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}\|\sigma(u(s,y))D_{r,z}u(s,y)\|_2 f_R^2(s,y)dy\right)^2 dsdrdz.$$

Using Hölder's inequality, the fact that σ has linear growth and the estimates (100) and (99), we can write

$$\begin{aligned}&\mathbb{E}(Dv_{R,\bullet}\otimes_1 v_{R,\bullet}\|_H^2\\ &\le C\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}} p_{s-r}(y-z)f_R^2(s,y)dy\right)^2 dsdrdz\\ &= C\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}\int_{[-R,R]^2} p_{s-r}(y-z)p_{t-s}(x-y)p_{t-s}(x'-y)dxdx'dy\right)^2 dsdrdz\end{aligned}$$

Using that $\int_{[-R,R]} p_{t-s}(x'-y)dx' \le 1$ and the semigroup property of the heat kernel, we obtain

$$\mathbb{E}(Dv_{R,\bullet}\otimes_1 v_{R,\bullet}\|_H^2 \le C\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{[-R,R]} p_{t-r}(x-z)dx\right)^2 drdz \le CR.$$

Step (iv) Estimation of $\mathbb{E}\left(\|D_\bullet v_R \otimes_1 v_{R,\bullet}\|_H^2\right)$: Assuming, to simplify, that σ is continuously differentiable, we have

$$D_{r,z}v_R(s,y) = \sigma'(u(s,y))D_{r,z}u(s,y)f_R(s,y)\mathbf{1}_{\{0\le r\le s\le t\}}.$$

As a consequence,

$$\begin{aligned}\|D_\bullet v_R \otimes_1 v_{R,\bullet}\|_H^2 =& \int_{[0,t]\times\mathbb{R}}\left(\int_{[0,t]\times\mathbb{R}} \sigma'(u(s,y))\sigma(u(r,z))D_{r,z}u(s,y)f_R(r,z)f_R(s,y)drdz\right)^2 dsdy\\ \le& t(\mathrm{Lip}\sigma)^2\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}\sigma(u(r,z))D_{s,y}u(r,z)f_R(r,z)f_R(s,y)dz\right)^2 drdy.\end{aligned}$$

Taking the mathematical expectation, Minkowski's inequality yields

$$\mathbb{E}(D_\bullet v_R \otimes_1 v_{R,\bullet}\|_H^2 \le t(\mathrm{Lip}\sigma)^2\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}\|\sigma(u(r,z))D_{s,y}u(r,z)\|_2 f_R(s,y)f_R(r,z)dz\right)^2 drdy.$$

Using Hölder's inequality, the fact that σ has linear growth and the estimates (100) and (99), we can write

$$\begin{aligned}&\mathbb{E}(D_\bullet v_R \otimes_1 v_{R,\bullet}\|_H^2\\ &\le C\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}p_{s-r}(y-z)f_R(s,y)f_R(r,z)dz\right)^2 drdy\\ &= C\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{\mathbb{R}}\int_{[-R,R]^2}p_{s-r}(y-z)p_{t-s}(x-y)p_{t-r}(x'-z)dxdx'dz\right)^2 drdy\end{aligned}$$

Using that $\int_{[-R,R]}p_{t-r}(x'-z)dx' \le 1$ and the semigroup property of the heat kernel, we obtain

$$\mathbb{E}(Dv_{R,\bullet} \otimes_1 v_{R,\bullet}\|_H^2 \le C\int_{[0,t]\times\mathbb{R}}\int_0^t\left(\int_{[-R,R]}p_{t-r}(x-z)dx\right)^2 drdz \le CR.$$

This completes the proof. $\square$

13.6 Functional Central Limit Theorem

The following result is a functional version of the central limit theorem for space averages.

Theorem 13.3 (Huang, Nualart and Viitasaari [7]). *Set $\xi(s) = \mathbb{E}[\sigma(u(s,y))^2]$ for any $s \ge 0$. Then*

$$\left(\frac{1}{\sqrt{R}}\left(\int_{-R}^R u(t,x)dx - 2R\right)\right)_{t\in[0,T]} \to \left(\int_0^t\sqrt{2\xi(s)}dB_s\right)_{t\in[0,T]},$$

as R tends to infinity, where B is a Brownian motion and the convergence is in law on the space of continuous functions $C([0,T])$.

Proof. The proof will be done in two steps.

(i) Proof of tightness: Tightness is a consequence of the estimate:

$$\mathbb{E}\left(\left|\int_{-R}^{R} u(t,x)dx - \int_{-R}^{R} u(s,x)dx\right|^p\right) \le C(T)R^{\frac{p}{2}}(t-s)^{\frac{p}{2}}.$$

for any $0 \le s < t \le T$ and any $p \ge 1$.

(ii) Convergence of the finite-dimensional distributions: We use the following result:

Proposition 13.3. *Let $F = (F^{(1)}, \dots, F^{(m)})$ be a random vector such that $F^{(i)} = \delta(v^{(i)})$ for $v^{(i)} \in \operatorname{Dom}\delta$, $i = 1, \dots, m$. Assume $F^{(i)} \in \mathbb{D}^{1,2}$ for $i = 1, \dots, m$. Let Z be an m-dimensional Gaussian centered vector with covariance matrix $(C_{i,j})_{1 \le i,j \le m}$. For any C^2 function $h : \mathbb{R}^m \to \mathbb{R}$ with bounded second partial derivatives, we have*

$$|\mathbb{E}(h(F_R)) - \mathbb{E}(h(Z))| \le \frac{m}{2}\|h''\|_\infty \sqrt{\sum_{i,j=1}^{m} \mathbb{E}\left[(C_{i,j} - \langle DF^{(i)}, v^{(j)}\rangle_{\mathfrak{H}})^2\right]},$$

where

$$\|h''\|_\infty = \max_{1 \le i,j \le m} \sup_{x \in \mathbb{R}^m} \left|\frac{\partial^2 h}{\partial x_i \partial x_j}(x)\right|.$$

Fix points $0 \le t_1 < \cdots < t_m \le T$ and consider the random variables

$$F_R^{(i)} = \frac{1}{\sqrt{R}}\left(\int_{-R}^{R} u(t_i, x)dx - 2R\right),$$

for $i = 1, \dots, m$. We can write $F_R^{(i)} = \delta(v_R^{(i)})$, where

$$v_R^{(i)}(s,y) = \mathbf{1}_{[0,t_i]}(s)\frac{1}{\sqrt{R}}\int_{-R}^{R} p_{t_i - s}(x-y)\sigma(u(s,y))dx.$$

Set $F_R = (F_R^{(1)}, \dots, F_R^{(m)})$ and let Z be an m-dimensional Gaussian centered vector with covariance

$$C_{i,j} := \mathbb{E}[Z^i Z^j] = 2\int_0^{t_i \wedge t_j} \xi(r)dr,$$

where we recall that $\xi(r) = \mathbb{E}[\sigma(2(r,x))^2]$.

Then, applying the previous proposition,

$$|\mathbb{E}(h(F_R)) - \mathbb{E}(h(Z))| \le \frac{m}{2}\|h''\|_\infty \sqrt{\sum_{i,j=1}^{m} \mathbb{E}\left[\left(C_{i,j} - \langle DF_R^{(i)}, v_R^{(j)}\rangle_{\mathfrak{H}}\right)^2\right]}.$$

It suffices to show that for each i, j, $\langle DF_R^{(i)}, v_R^{(j)}\rangle_{\mathfrak{H}}$ converges in $L^2(\Omega)$, as R tends to infinity to $C_{i,j}$. This follows from the expression

$$\begin{aligned}\langle DF_R^{(i)}, v_R^{(j)}\rangle_{\mathfrak{H}} &= \frac{1}{R}\int_0^{t_i\wedge t_j}\int_{\mathbb{R}} f_R(t_i,s,y)f_R(t_j,s,y)\sigma^2(u(s,y))dyds\\ &\quad+\frac{1}{R}\int_0^{t_i\wedge t_j}\int_{\mathbb{R}} f_R(t_j,s,y)\sigma(u(s,y))\\ &\quad\times\left(\int_s^{t_i}\int_{\mathbb{R}} f_r(t_i,r,z)\Sigma(r,z)D_{s,y}u(r,z)W(dr,dz)\right)dyds,\end{aligned}$$

with the notation $f_R(t,s,y)=\int_{-R}^R p_{t-s}(x,y)dx$. □

13.7 Spatial colored noise

Consider the stochastic heat equation

$$\frac{\partial u}{\partial t}=\frac{1}{2}\Delta u+\sigma(u)\dot{W},$$

on $\mathbb{R}_+\times\mathbb{R}^d$ with initial condition $u(0,x)=1$. The function σ is Lipschitz continuous and we assume that $\sigma(1)\neq 0$.

The noise $\dot{W}(t,x)$ is a centered Gaussian random field with covariance

$$\mathbb{E}\big[\dot{W}(t,x)\dot{W}(s,y)\big]=\delta_0(t-s)\gamma(x-y),$$

where $\gamma:\mathbb{R}^d\to\mathbb{R}_+\cup\{\infty\}$ is nonnegative definite and the Fourier transform of γ is a tempered measure μ, that satisfies Dalang's condition:

$$\int_{\mathbb{R}^d}\frac{\mu(d\xi)}{1+|\xi|^2}<\infty.$$

We define formally the noise as a Gaussian centered family of random variables

$$W=\{W(\varphi),\varphi\in C_0^\infty([0,\infty)\times\mathbb{R}^d)\},$$

with covariance

$$\begin{aligned}\mathbb{E}[W(\phi)W(\varphi)] &= \int_0^\infty\int_{\mathbb{R}^{2d}}\phi(s,x)\varphi(s,y)\gamma(x-y)dxdyds\\ &=\int_0^\infty\int_{\mathbb{R}^d}\mathcal{F}\phi(s,\xi)\overline{\mathcal{F}\varphi(s,\xi)}\mu(d\xi)ds,\end{aligned}$$

where $\mathcal{F}\phi$ refers to the Fourier transform in the space variable. Let $\mathfrak{H}_0$ be the closure of $C_0^\infty(\mathbb{R}^d)$ under the inner product

$$\langle\varphi,\psi\rangle_{\mathfrak{H}_0}=\int_{\mathbb{R}^{2d}}\varphi(x)\psi(y)\gamma(x-y)dxdy=\int_{\mathbb{R}^d}\mathcal{F}\varphi(\xi)\overline{\mathcal{F}\psi(\xi)}\mu(d\xi).$$

Then, the Gaussian family W can be extended to the Hilbert space $\mathfrak{H} := L^2(\mathbb{R}_+, \mathfrak{H}_0)$, in such a way that $\{W(g), g \in L^2(\mathbb{R}_+; \mathfrak{H}_0)\}$ is an isonormal Gaussian process.

From Dalang [6] there is a unique mild solution, which is an adapted random field u such that for all $p \geq 2$,

$$\sup_{x\in\mathbb{R}} \sup_{0\leq t\leq T} \mathbb{E}[|u(t,x)|^p] < \infty,$$

and u satisfies the integral equation:

$$u(t,x) = 1 + \int_0^t \int_{\mathbb{R}^d} p_{t-s}(x-y)\sigma(u(s,y))W(ds,dy),$$

where $p_t(x) = (2\pi t)^{-d/2} e^{-|x|^2/2t}$. The stochastic integral is an extension of the Itô integral to adapted random fields g such that $\mathbb{E}[\|g\|_{\mathfrak{H}}^2] < \infty$.

The process $x \to u(t,x)$ is stationary and it is also ergodic provided $\mu(\{0\}) = 0$. Fix $t > 0$. We are interested in the asymptotic behavior as $R \to \infty$ of the the spacial averages

$$\int_{B_R} u(t,x)dx,$$

where $B_R = \{x \in \mathbb{R}^d : |x| \leq R\}$.

We put, using the mild formulation

$$F_R = \int_{B_R} (u(t,x)-1)dx = \int_0^t \int_{\mathbb{R}^d} \left(\int_{B_R} p_{t-s}(x-y)dx \right) \sigma(u(s,y))W(ds,dy).$$

Notice that F_R is a stochastic integral.

Consider the following assumptions on the covariance function γ:

(H1) Integrable covariance: $\int_{\mathbb{R}^d} \gamma(x)dx < \infty$.

(H2) Riesz covariance: $\gamma(x) = |x|^{-\beta}$, $\beta < \min(2,d)$.

The following quantitative central limit theorem has been proved in the references [8, 5].

Theorem 13.4. *Assume* **(H1)** *or* **(H2)**. *For each $t > 0$ there exist a constant C_t, depending on t, such that*

$$d_{TV}\left(\frac{F_R}{\sigma_R}, Z\right) \leq C_t \begin{cases} R^{-d/2} & \text{under } \textbf{(H1)} \\ R^{-\beta/2} & \text{under } \textbf{(H2)} \end{cases}$$

where $\sigma_R^2 = \mathrm{Var}\big(\int_{B_R}[u(t,x)-1]\,dx\big)$ and Z has the $N(0,1)$ law.

Moreover, as $R \to +\infty$, under **(H1)**

$$\sigma_R^2 \sim R^d \int_{\mathbb{R}^d} \mathrm{Cov}(u(t,x), u(t,0))dx$$

and, under **(H2)**

$$\sigma_R^2 \sim \left(k_\beta \int_0^t \eta^2(s)\, ds\right) R^{2d-\beta},$$

where $\eta(s) = \mathbb{E}[\sigma(u(s,y))]$ and $k_\beta := \int_{B_1^2} |x_1 - x_2|^{-\beta} dx_1 dx_2$.

In [8], the authors have also obtained a functional version of the above central limit theorem under **(H2)**. That means, as $R \to +\infty$, we have

$$\left(R^{\frac{\beta}{2}-d} \int_{B_R} \big[u(t,x) - 1\big]\, dx\right)_{t\in[0,T]} \Rightarrow \left(\sqrt{k_\beta} \int_0^t \eta(s) dB_s\right)_{t\in[0,T]},$$

where B is a Brownian motion and the convergence takes place on the space of continuous functions $C([0,T])$. If $\sigma(x) = x$, then the first chaos dominates (*non-chaotic behavior*). This is not true for space-time white noise and $\sigma(x) = x$.

References

[1] S. Ben Hariz: Limit theorems for the non-linear functionals of stationary Gaussian processes. *J. Mult. Anal.* **80**, 191-216, (2002).

[2] H. Biermé, A. Bonami, I. Nourdin and G. Peccati: Optimal Berry-Esseen rates on the Wiener space: the barrier of third and fourth cumulants. *ALEA* **9**, 473-500, (2012).

[3] P. Breuer and P. Major: Central limit theorems for non-linear functionals of Gaussian fields. *J. Mult. Anal.* **13**, 425-441, (1983).

[4] D. Chambers and E. Slud: Central limit theorems for nonlinear functionals of stationary Gaussian processes. *Probab. Theory Rel. Fields* **80**, 323-346, (1989).

[5] L. Chen, D. Khoshnevisan, D. Nualart and F. Pu: Central limit theorems for spatial averages of the stochastic heat equation via Malliavin-Stein's method. *Stoch. Partial Differ. Equ. Anal. Comput.* **11**, 122-176, (2023).

[6] R. Dalang: Extending martingale measure stochastic integral with applications to spatially homogeneous S. P. D. E's. *Electron. Journal Probab.* **4**, no. 6, 29 pp., (1999).

[7] J. Huang, D. Nualart and L. Viitasaari: A Central Limit Theorem for the stochastic heat equation. *Stoch. Process. Appl.* **130**, 7170-7184, (2020).

[8] J. Huang, D. Nualart, L. Viitasaari and G. Zheng: Gaussian fluctuations for the stochastic heat equation with colored noise. *Stoch. Partial Differ. Equ. Anal. Comput.* **8**, 402-421, (2020).

[9] P. Malliavin:Stochastic calculus of variation and hypoelliptic operators, In: *Proceedings of the International Symposium on Stochastic Differential Equations* (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), pp. 195–263, Wiley, New York-Chichester-Brisbane (1978).

[10] I. Nourdin and D. Nualart: The functional Breuer-Major theorem. *Probab. Theory Rel. Fields.* **176**, 203-218, (2020).

[11] I. Nourdin, D. Nualart and Peccati: The Breuer-Major Theorem in total variation: improved rates under minimal regularity. *Stoch. Process, Appl.* **131**, 1-20, (2021).

[12] I. Nourdin and G. Peccati: *Normal approximations with Malliavin calculus: from Stein's method to universality.* Cambridge tracts in Mathematics **192**, Cambridge University Press (2012).

[13] I. Nourdin and G. Peccati. The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* **143**, 3123–3133, (2015).

[14] I. Nourdin, G. Peccati and X. Yang: Berry-Esseen bounds in the Breuer-Major CLT and Gebelein's inequality. *Electron. Commun. Probab.* **24**, no. 34, 12 pp., (2019).

[15] D. Nualart: *The Malliavin calculus and related topics.* 2nd edition. Probability and Its Applications, Springer (2006).

[16] D. Nualart: Malliavin calculus and normal approximations. *Ensaios Matemáticos* **34**, 1-74, (2019).

[17] D. Nualart and E. Nualart: *Introduction to Malliavin Calculus.* Institute of Mathematical Statistics Textbooks, Cambridge University Press (2018).

[18] D. Nualart and E. Pardoux: Stochastic calculus with anticipating integrands. *Probab. Theory Rel. Fields* **78**, 535–581, (1988).

[19] D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33**, 177-193, (2005).

[20] D. Nualart and S. Ortiz-Latorre. Central limit theorems for multiple stochastic integrals. *Stochastic Proc. Appl.* **118** , 614–628, (2007).

[21] D. Nualart and H. Zhou. Total variation estimates in the Breuer-Major theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* **57**, 770-777, (2021).

[22] G. Peccati and C. Tudor. Gaussian limits for vector-valued multiple stochastic integrals . *Séminaire de Probabilités XXXVIII*, pag 247-262. Lecture Notes in Math., **1857**, Springer Verlag, Berlin, 2005.

[23] A. V. Skorohod: On a generalization of a stochastic integral. *Theory Probab. Appl.* **20**, 219–233., (1975).

[24] D. W. Stroock. Some applications of stochastic calculus to partial differential equations. In: *Ecole d'Eté de Probabilités de Saint Flour*, Lecture Notes in Math. **976**, 267–382., (1983).

[25] J. B. Walsh: An introduction to stochastic partial differential equations. In: *Ecole d'Eté de Probabilités de Saint Flour XIV*, Lecture Notes in Math. **1180** (1986) 265–438.

[26] S. Watanabe. *Lectures on Stochastic Differential Equations and Malliavin Calculus.* Tata Institute of Fundamental Research, Springer-Verlag (1984).