Lectures on Malliavin calculus and its applications to finance

Eulalia Nualart

University of Paris 13

These lectures are offered on the basis of need or interest to graduate/ Ph.D. students, post-docs and other researchers of the University of Wisconsin (Madison), from June 23rd to July 2nd 2009. The recommended prior knowledge is an advance probability course. Some familiarity with Itô stochastic calculus is also recommended.

The aim of these lectures is to give an introduction to the stochastic calculus of variations, known as Malliavin calculus, and give one of its applications in Mathematical Finance to the computations of "Greeks", sensitivity parameters of opcion prices.

The Malliavin calculus is an infinite-dimensional differential calculus on the Wiener space, that was first introduced by Paul Malliavin in the 70's, with the aim of giving a probabilistic proof of Hörmander's theorem. This theory was then further developed, and since then, many new applications of this calculus have appeared.

We will start these lectures by defining in an abstract setting the concepts of isonormal Gaussian process, derivative operator and Sobolev spaces associated to differentiable random variables. We will then study the particular case of stochastic integrals with respect to Brownian motion. Secondly, we will define the dual operator, known as the Sokorohod integral, that will take us to one of the main tools of the Malliavin calculus which is the integration by parts formula. We will recall some of its applications to the study of probability laws of random variables on an abstract Wiener space. In particular we will study the case of diffusion processes and state Hörmander's theorem.

The second part of this course will discuss one of the applications of this calculus in Mathematical Finance. This application consists on using the integration by parts formula to give a probabilistic method for numerical computations of price sensitivities known as Greeks. We will particulary study the case where the option price process follows a Black-Scholes model, that will be previously introduced.

I would like to express my gratitude to Professor Thomas G. Kurtz for his kind invitation to give these lectures.

Plan of the lectures:

1 Introduction to Malliavin calculus

1.1 The isonormal Gaussian process

1.2 The Wiener chaos

1.3 Multiple stochastic integrals

1.4 The derivative operator

1.4.1 Definition and properties of the derivative operator

1.4.2 The derivative operator in the white noise case

1.5 The divergence operator

1.5.1 Definition and properties of the divergence operator

1.5.2 The Skorohod integral

1.5.3 The Clark-Ocone formula

1.6 Exercises

2. The integration by parts formula and applications to regularity of probability laws

2.1 The integration by parts formula

2.2 Existence and smoothness of densities

2.3 Application to diffusion processes: Hörmander's theorem

2.4 Exercises

3. Applications of Malliavin calculus in mathematical finance

3.1 Pricing and hedging financial options

3.2 The Black-Scholes model

3.3 Pricing and hedging financial options in the Black-Scholes model

3.4 Sensibility with respect to the parameters: the greeks

3.5 Application of the Clark-Ocone formula in hedging

3.5 Exercises

A Appendix

References

1 Introduction to Malliavin calculus

1.1 The isonormal Gaussian process

An isonormal Gaussian process is the set of :

a real and separable Hilbert space H (if $h, g \in H$, we denote its scalar product as $\langle f, g\rangle_H$ and its norm as $\|h\|_H$);

a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$;

a Gaussian process on H, $W = \{W(h), h \in H\}$, that is, W is a centered Gaussian family of random variables such that $E[W(h)W(g)] = \langle f, g\rangle_H$, for all $h, g \in H$.

We remark that :

1. By Kolmogorov's theorem, given H, we can always construct $(\Omega, \mathcal{F}, \mathrm{P})$ and W with the conditions above.
2. Under the above conditions, the mapping $h \mapsto W(h)$ is linear. Indeed, let $h, g \in H$ and $\lambda, \mu \in \mathbb{R}$. We have that

$$\begin{aligned}\mathrm{E}[(W(\lambda h + \mu g) - \lambda W(h) - \mu W(g))^2] &= \|\lambda h + \mu g\|_H + \lambda^2\|h\|_H^2 + \mu^2\|g\|_H^2 \\ &- 2\lambda\langle \lambda h + \mu g, h\rangle_H - 2\mu\langle \lambda h + \mu g, g\rangle_H + 2\lambda\mu\langle h, g\rangle_H = 0.\end{aligned}$$

Hence, there is a linear isometry of H onto a closed sub-space of $L^2(\Omega, \mathcal{F}, \mathrm{P})$ that will be denoted by $\mathcal{H}_1$, whose elements are centered Gaussian random variables. Moreover, the linearity implies that it suffices to assume that each random variable $W(h)$ is centered and Gaussian in the definition of the Gaussian process W.

Example 1.1 The isonormal Gaussian process associated to the Brownian motion: *Let $(B(t) = (B^1(t), ..., B^d(t)), t \geq 0)$ be a d-dimensional Brownian motion defined on its canonical probability space $(\Omega, \mathcal{F}, \mathrm{P})$. That is, $\Omega = \mathcal{C}_0(\mathbb{R}_+; \mathbb{R}^d)$, P is the d-dimensional Wiener measure, and $\mathcal{F}$ is the completion of the Borel σ-field of Ω with respect to P. In this case, the underlying Hilbert space is $H = L^2(\mathbb{R}_+; \mathbb{R}^d)$, and for each $h \in H$, we define $W(h)$ to be the Wiener integral*

$$W(h) = \sum_{i=1}^{d} \int_{\mathbb{R}_+} h_i(s) dB^i(s).$$

Example 1.2 The isonormal Gaussian process associated to the fractional Brownian motion: *Fix $T > 0$ and let $B^\gamma = (B^\gamma(t), t \in [0, T])$ be a fractional Brownian motion with Hurst parameter $\gamma \in (0, 1)$ defined on its canonical probability space $(\Omega, \mathcal{F}, \mathrm{P})$, that is, B^γ is a centered Gaussian process with covariance*

$$R_\gamma(t, s) = \mathrm{E}[B^\gamma(t)B^\gamma(s)] = \frac{1}{2}(s^{2\gamma} + t^{2\gamma} - |t - s|^{2\gamma}).$$

We denote by $\mathcal{E}$ the step functions on $[0, T]$. Let H be the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the scalar product

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

The mapping $\mathbf{1}_{[0,t]} \to B^\gamma(t)$ *can be extended to an isometry between* H *and the Gaussian space* $\mathcal{H}_1$ *associated with* B^γ*. We denote this isometry by* $\varphi \to B^\gamma(\varphi)$*. Then* $\{B^\gamma(\varphi), \varphi \in H\}$ *is a Gaussian process on* H*.*

Observe that when $\gamma < \frac{1}{2}$*,* H *contains the set of Hölder continuous functions* $\mathcal{C}^\alpha([0,T])$ *with* $\alpha > \frac{1}{2} - \gamma$*, and when* $\gamma > \frac{1}{2}$*,* H *contains distributions and contains the space* $L^{1/\gamma}([0,T])$*. For a detailed definition and properties of this isonormal Gaussian processes see [N06, Chapter 5] and [N09, Section 8.2].*

Example 1.3 The isonormal Gaussian process associated to the Brownian sheet: *Let* $W = (W(t) \in \mathbb{R}_+^N)$ *be an* N*-parameter Brownian sheet, that is, a centered Gaussian process with covariance*

$$\mathbb{E}[W(s)W(t)] = \prod_{i=1}^{N}(s_i \wedge t_i), \quad s,t \in \mathbb{R}_+^N,$$

defined on its canonical probability space $(\Omega, \mathcal{F}, \mathrm{P})$*. Let* H *be the Hilbert space* $H = L^2(\mathbb{R}_+^N, \mathcal{B}(\mathbb{R}_+^N), \lambda)$*, where* λ *is the Lebesgue measure in* $\mathbb{R}_+^N$*. For each* $h \in H$ *we define* $W(h)$ *to be the multi-parameter Wiener integral* $\int_{\mathbb{R}_+^N} h(s)dW(s)$*. Recall that multi-parameter Wiener integrals are constructed as in the one-parameter case.*

1.2 The Wiener chaos

We define the Hermite polynomial of degree n and parameter $\lambda > 0$ by:

$$H_0(x,\lambda) = 1,$$
$$H_n(x,\lambda) = \frac{(-\lambda)^n}{n!} e^{x^2/2\lambda} \frac{d^n}{dx^n}(e^{-x^2/2\lambda}), \quad n \geq 1,\ x \in \mathbb{R}.$$

We will use the notation $H_n(x) := H_n(x,1)$.

We note that $H_1(x,\lambda) = x$ y $H_2(x,\lambda) = \frac{1}{2}(x^2 - \lambda)$. Moreover, the following properties hold:

$$\frac{\partial}{\partial x} H_n(x,\lambda) = H_{n-1}(x,\lambda), \quad n \geq 1; \tag{1.1}$$
$$(n+1)H_{n+1}(x,\lambda) = xH_n(x,\lambda) - \lambda H_{n-1}(x,\lambda), \quad n \geq 1; \tag{1.2}$$
$$H_n(-x,\lambda) = (-1)^n H_n(x,\lambda), \quad n \geq 1; \tag{1.3}$$
$$\frac{\partial}{\partial \lambda} H_n(x,\lambda) = -\frac{1}{2}\frac{\partial^2}{\partial x^2} H_n(x,\lambda), \quad n \geq 1. \tag{1.4}$$

In particular, (1.2) implies that the highest order term of $H_n(x,\lambda)$ is $\frac{x^n}{n!}$.

In order to prove these properties it suffices to note that the Hermite polynomials are the coefficients of the expansion in powers of t of the function:

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

For each $n \geq 1$, we define the Wiener chaos of order n, $\mathcal{H}_n$, as the closed linear sub-space of $L^2(\Omega, \mathcal{F}, \mathrm{P})$ generated by the family of random variables

$$\{H_n(W(h)),\ h \in H,\ \|h\|_H = 1\}.$$

$\mathcal{H}_0$ will be the set of constants. Note that $\mathcal{H}_1 = W$.

The next lemma shows that $\mathcal{H}_n$ and $\mathcal{H}_m$ are orthogonal if $n \neq m$.

Lemma 1.4 *Let X y Y two centered random variables with variance 1 and Gaussian joint distribution. Then, for all $n, m \geq 0$,*

$$\mathrm{E}[H_n(X)H_m(Y)] = \begin{cases} 0, & \text{if } n \neq m, \\ \frac{1}{n!}(\mathrm{E}[XY])^n, & \text{if } n = m. \end{cases}$$

Proof. For all $s, t \in \mathbb{R}$, we have

$$\mathrm{E}\left[\exp(sX - \frac{s^2}{2})\exp(tY - \frac{t^2}{2})\right] = \exp(st\mathrm{E}[XY]).$$

Taking the $(n+m)$-th partial derivative $\frac{\partial^{n+m}}{\partial s^n \partial t^m}$ at $s = t = 0$ in both sides of the above equality, we obtain:

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

△

Moreover, the following orthogonal decomposition holds.

Theorem 1.5 *Let $\mathcal{G}$ be the σ-algebra generated by W. Then,*

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

Proof. Let $X \in L^2(\Omega, \mathcal{G}, \mathrm{P})$ be orthogonal to $\mathcal{H}_n$ for all $n \geq 0$. That is, $\mathrm{E}[XH_n(W(h))] = 0$ for all $h \in H$ such that $\|h\|_H = 1$. We note that x^n can be expressed as linear combination of $H_r(x)$, $0 \leq r \leq n$. This implies that $\mathrm{E}[XW(h)^n] = 0$ for all $n \geq 0$, and hence $\mathrm{E}[X\exp(W(h))] = 0$ for all $h \in H$ such that $\|h\|_H = 1$. The linearity of the mapping $h \mapsto W(h)$ implies that for all $t_1, ..., t_m \in \mathbb{R}$ and $h_1, ..., h_m \in H$ such that $\|h_i\|_H = 1$, $m \geq 1$,

$$\mathrm{E}[X\exp(\sum_{i=1}^m t_i W(h_i))] = 0. \tag{1.5}$$

We next show that (1.5) implies that $X = 0$. Write $X = X^+ - X^-$ and define the measures

$$\nu^{+,-}(B) = \mathrm{E}[X^{+,-}1_B(W(h_1), ..., W(h_m))], \quad B \in \mathcal{B}(\mathbb{R}^m).$$

Then ν^+ and ν^- are finite measures on $\mathbb{R}^m$ and their Laplace transforms are given by

$$\varphi_{\nu^{+,-}}(\lambda) = \int_{\mathbb{R}^m} \exp(\lambda \cdot x)\nu^{+,-}(dx) = \mathrm{E}[X^{+,-}\exp(\sum_{i=1}^m \lambda_i W(h_i))], \qquad \lambda \in \mathbb{R}^m.$$

Hence, (1.5) implies that the Laplace transform of the signed measure

$$\nu(B) = \mathrm{E}[X1_B(W(h_1), ..., W(h_m))], \quad B \in \mathcal{B}(\mathbb{R}^m)$$

is zero. Consequently, this measure is zero and thus $\mathrm{E}[X1_G] = 0$, for all $G \in \mathcal{G}$. So $X = 0$ and the proof is completed. △

Example 1.6 *We consider* $H = \mathbb{R}$, $(\Omega, \mathcal{F}, \mathrm{P}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, *where* μ *is the Gaussian law* $N(0,1)$. *For each* $h \in H$, *we define* $W(h) = hX$, *where* X *is a Gaussian random variable* $N(0,1)$. *From property (1.3) we deduce that* $\mathcal{H}_n$ *has dimension one and is generated by* $H_n(X)$. *Moreover, Lemma 1.4 and Theorem 1.5 imply that the Hermite polynomials* $\{\sqrt{n!}H_n(x), n \geq 0\}$ *form a complete orthonormal system in* $L^2(\mathbb{R}, \mu)$.

1.3 Multiple stochastic integrals

Tha aim of this section is to define multiple Wiener-Itô integrals with respect to a Brownian motion. Multiple Wiener-Itô integrals can also be defined for a general white noise based on a general measured space $(T, \mathcal{B}, \mu)$, where μ is a σ-finite measure without atoms (see [N06, Section 1.1.2]). To simplify the exposition, in these lectues we will consider a one-dimensional Brownian motion $(B(t), t \in T)$, $T = [a,b]$, where $0 \leq a < b < \infty$, and its isonormal Gaussian process defined in Example 1.1, that is, $H = L^2([a,b];\mathbb{R})$ and $W(h) = \int_{[a,b]} h(s)dB(s)$, for each $h \in H$.

We recall the definition of the Itô integral with respect to the Brownian motion:

For each $t \geq 0$, let $\mathcal{F}_t$ be the σ-algebra generated by the random variables $\{B(s), 0 \leq s \leq t\}$ and the P-measure zero sets of $\mathcal{F}$. We say that a stochastic process $\{u(t), t \geq 0\}$ is adapted to $\mathcal{F}_t$ if $u(t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$.

Let $L^2_a(T \times \Omega, \mathcal{B}(T) \times \mathcal{F}, \lambda \times \mathrm{P})$ be the set of square integrable and adapted processes. Consider the set of elementary and adapted processes $\mathcal{E}$:

$$\mathcal{E} = \{u(t) = \sum_{i=1}^{n} F_i \mathbf{1}_{(t_i, t_{i+1}]}(t), 0 \leq t_1 < \cdots < t_{n+1}, t_i \in T, F_i \in \mathcal{F}_{t_i} \text{ square integrable}\}.$$

The set $\mathcal{E}$ is dense in $L^2_a(T \times \Omega)$. We define the Itô integral of $u \in \mathcal{E}$ with respect to the Brownian motion by:

$$\int_T u(t)dB(t) = \sum_{i=1}^{n} F_i(B(t_{i+1}) - B(t_i)).$$

This is a linear functional with the following properties:

$$\mathrm{E}\Big[\int_T u(t)dB(t)\Big] = 0, \qquad \mathrm{E}\Big[\Big(\int_T u(t)dB(t)\Big)^2\Big] = \mathrm{E}\Big[\int_T u^2(t)dt\Big].$$

Using the isometry property we can extend the Itô integral to the class $L^2_a(T \times \Omega)$ as the limit in $L^2(\Omega)$ of the integral of processes in $\mathcal{E}$ with the same properties above.

The goal of this section is to define the multiple Wiener-Itô integral with respect to B:

$$\int_{T^n} f(t_1, t_2, ..., t_n)dB(t_1)dB(t_2)\cdots dB(t_n),$$

of a function $f \in L^2(T^n)$. The crucial idea of Itô (1951) was to define an integral for elementary functions that vanish on the diagonal and then approximate a function $f \in L^2(T^n)$ by a function of this type and pass to the limit of the corresponding integral.

Consider the *diagonal set* of T^n:

$$D = \{(t_1, t_2, ..., t_n) \in T^n; \exists\, i \neq j : t_i = t_j\}.$$

Let $\mathcal{E}_n$ be the vector space formed by the set of elementary functions on T^n that vanish over D:

$$f(t_1, ..., t_n) = \sum_{i_1,...,i_n=1}^{k} a_{i_1 \cdots i_n} \mathbf{1}_{[\tau_{i_1-1}, \tau_{i_1}) \times \cdots \times [\tau_{i_n-1}, \tau_{i_n})}(t_1, ..., t_n), \tag{1.6}$$

where $a = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_k = b$, such that

$$a_{i_1 \cdots i_n} = 0 \;\text{ if }\; i_p = i_q \;\text{ for some }\; p \neq q. \tag{1.7}$$

For $f \in \mathcal{E}_n$, we define

$$I_n(f) = \sum_{i_1,...,i_n=1}^{k} a_{i_1 \cdots i_n} \xi_{i_1} \cdots \xi_{i_n}, \tag{1.8}$$

where $\xi_{i_p} = B(\tau_{i_p}) - B(\tau_{i_p - 1})$ for $p = 1, ..., n$. We observe that $I_n(f)$ is well-defined, that is, its definition does not depend on the particular representation of f. Moreover, it is linear over $\mathcal{E}_n$.

We define the symmetrization $\tilde{f}$ of f by

$$\tilde{f}(t_1, ..., t_n) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} f(t_{\sigma(1)}, ..., t_{\sigma(n)}),$$

where $\mathcal{S}_n$ is the set of all permutations $\{1, ..., n\}$. Because the Lebesgue measure is symmetric, we have that for all $\sigma \in \mathcal{S}_n$

$$\int_{T^n} |f(t_1, ..., t_n)|^2 dt_1 \cdots dt_n = \int_{T^n} |f(t_{\sigma(1)}, ..., t_{\sigma(n)})|^2 dt_1 \cdots dt_n.$$

Therefore, from the triangle inequality, we deduce that

$$\|\tilde{f}\|_{L^2(T^n)} \leq \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \|f\|_{L^2(T^n)} = \|f\|_{L^2(T^n)}. \tag{1.9}$$

The following properties hold:

Lemma 1.7 *If $f \in \mathcal{E}_n$, then $I_n(f) = I_n(\tilde{f})$.*

Proof. Because I_n is linear, it suffices to prove the lemma for a function of the form

$$f = \mathbf{1}_{[t_1^{(1)}, t_1^{(2)}) \times \cdots \times [t_n^{(1)}, t_n^{(2)})},$$

where the intervals $[t_i^{(1)}, t_i^{(2)}) \subset T$, $i = 1, ..., n$, are disjoint. We have that

$$I_n(f) = \prod_{i=1}^{n} (B(t_i^{(2)}) - B(t_i^{(1)})).$$

On the other hand,

$$I_n(\tilde{f}) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \prod_{i=1}^{n} (B(t_{\sigma(i)}^{(2)}) - B(t_{\sigma(i)}^{(1)})).$$

We observe that for all $\sigma \in \mathcal{S}_n$, $\prod_{i=1}^{n} (B(t_{\sigma(i)}^{(2)}) - B(t_{\sigma(i)}^{(1)})) = \prod_{i=1}^{n} (B(t_i^{(2)}) - B(t_i^{(1)}))$. The result is now proved. $\triangle$

Lemma 1.8 *Let $f \in \mathcal{E}_n$ and $g \in \mathcal{E}_m$, $n, m \geq 1$. Then* $\mathrm{E}[I_n(f)] = 0$ *and*

$$\mathrm{E}[I_n(f)I_m(g)] = \begin{cases} 0, & \text{if } n \neq m, \\ n!\langle \tilde{f}, \tilde{g}\rangle_{L^2(T^n)}, & \text{if } n = m. \end{cases} \tag{1.10}$$

Proof. Let $f \in \mathcal{E}_n$ defined in equation (1.6). Then $I_n(f)$ is given in equation (1.8). Because the function f satisfies condition (1.7), the coefficients $a_{i_1 i_2 \cdots i_n}$ are 0 if the intervals $[\tau_{i_1-1}, \tau_{i_1}), ..., [\tau_{i_n-1}, \tau_{i_n})$ are not disjoint. On the other hand, when these intervals are disjoint, the corresponding product $\xi_{i_1} \cdots \xi_{i_n}$ has zero mean. Thus, we deduce that $\mathrm{E}[I_n(f)] = 0$.

By Lemma 1.8 it suffices to assume that f and g are symmetric to prove (1.10). Let $f \in \mathcal{E}_n$ defined as in (1.6). It suffices to assume that $f \in \mathcal{E}_n$ and $g \in \mathcal{E}_m$ are associated to the same partition $[\tau_{i_1-1}, \tau_{i_1}), ..., [\tau_{i_k-1}, \tau_{i_k})$, that is,

$$g(t_1, ..., t_m) = \sum_{i_1, ..., i_m = 1}^{k} b_{i_1 \cdots i_m} \mathbf{1}_{[\tau_{i_1-1}, \tau_{i_1}) \times \cdots \times [\tau_{i_m-1}, \tau_{i_m})}(t_1, ..., t_m),$$

where $b_{i_1 \cdots i_m} = 0$ if $i_p = i_q$ for some $p \neq q$.

On the other hand, for all $\sigma \in \mathcal{S}_n$, we have that $a_{i_1 i_2 \cdots i_n} = a_{\sigma(i_1)\sigma(i_2)\cdots\sigma(i_n)}$, and hence,

$$I_n(f) = n! \sum_{1 \leq i_1 < ... < i_n \leq k} a_{i_1 \cdots i_n} \xi_{i_1} \cdots \xi_{i_n},$$

and the same holds for g.

Therefore,

$$\mathrm{E}[I_n(f)I_m(g)] = n!m! \sum_{1 \leq i_1 < ... < i_n \leq k} \sum_{1 \leq j_1 < ... < j_m \leq k} a_{i_1 \cdots i_n} b_{j_1 \cdots j_m} \mathrm{E}[\xi_{i_1} \cdots \xi_{i_n} \xi_{j_1} \cdots \xi_{j_m}].$$

We observe that for a fixed set of indices $i_1 < ... < i_n$,

$$\mathrm{E}[\xi_{i_1} \cdots \xi_{i_n} \xi_{j_1} \cdots \xi_{j_m}] = \begin{cases} \prod_{p=1}^{n} (\tau_{i_p} - \tau_{i_p-1}), & \text{if } n = m \text{ y } j_1 = i_1, ..., j_n = i_n, \\ 0, & \text{else.} \end{cases}$$

We deduce that, if $n \neq m$, $\mathrm{E}[I_n(f)I_m(g)] = 0$, and if $n = m$,

$$\begin{aligned} \mathrm{E}[I_n(f)I_n(g)] &= (n!)^2 \sum_{1 \leq i_1 < ... < i_n \leq k} a_{i_1 \cdots i_n} b_{i_1 \cdots i_n} \prod_{p=1}^{n} (\tau_{i_p} - \tau_{i_p-1}) \\ &= n! \sum_{1 \leq i_1, ..., i_n \leq k} a_{i_1 \cdots i_n} b_{i_1 \cdots i_n} \prod_{p=1}^{n} (\tau_{i_p} - \tau_{i_p-1}) \\ &= n! \int_{T^n} f(t_1, ..., t_n) g(t_1, ..., t_n) dt_1 \cdots dt_n. \end{aligned}$$

△

The next result will allow to extend the integral $I_n(f)$ to $L^2(T^n)$.

Lemma 1.9 *The space $\mathcal{E}_n$ is dense in $L^2(T^n)$. That is, for all $f \in L^2(T^n)$, there exists a sequence $\{f_k\}_{k \geq 1}$, $f_k \in \mathcal{E}_n$, that converges towards f in $L^2(T^n)$, that is*

$$\lim_{k \to \infty} \int_{T^n} |f_k(t_1, ..., t_n) - f(t_1, ..., t_n)|^2 dt_1 \cdots dt_n = 0.$$

Proof. Because the usual set of elementary functions is dense in $L^2(T^n)$, and the set D has Lebesgue measure zero, the proof is immediate. △

Let $f \in L^2(T^n)$ and $f_k \in \mathcal{E}_n$ as in Lemma 1.9. Using Lemma 1.8 and inequality (1.9), we have that

$$\mathrm{E}[(I_n(f_p) - I_n(f_q))^2] = n!\|\tilde{f}_p - \tilde{f}_q\|^2_{L^2(T^n)} \leq n!\|f_p - f_q\|^2_{L^2(T^n)} \to 0,$$

as $p, q \to \infty$. Therefore, the sequence $\{I_n(f_k)\}_{k\geq 1}$ is Cauchy in $L^2(\Omega)$.

We define the Wiener-Itô integral as the limit of the sequence $\{I_n(f_k)\}_{k\geq 1}$ in $L^2(\Omega)$ and we denote it by

$$I_n(f) = \int_{T^n} f(t_1, t_2, ..., t_n) dB(t_1) dB(t_2) \cdots dB(t_n).$$

We observe that the definition does not depend on the choosed sequence. Note also that $I_1(f) = W(f)$. Moreover, Lemmas 1.7 and 1.8 can be extended to functions in $L^2(T^n)$ using Lemma 1.9:

Theorem 1.10 *Let $f \in L^2(T^n)$ and $g \in L^2(T^m)$, $n, m \geq 1$. We have that*

(i) $I_n(f) = I_n(\tilde{f})$,

(ii) $\mathrm{E}[I_n(f)] = 0$,

(iii)

$$\mathrm{E}[I_n(f)I_m(g)] = \begin{cases} 0, & \text{if } n \neq m, \\ n!\langle \tilde{f}, \tilde{g}\rangle_{L^2(T^n)}, & \text{if } n = m. \end{cases}$$

The next result will allow us to write the Wiener-Itô integral as an iterated Itô integral:

Theorem 1.11 *Let $f \in L^2(T^n)$, $n \geq 1$. Then*

$$I_n(f) = n! \int_a^b \cdots \int_a^{t_{n-2}} \left(\int_a^{t_{n-1}} \tilde{f}(t_1, ..., t_n) dB(t_n) \right) dB(t_{n-1}) \cdots dB(t_1).$$

Proof. We observe that it suffices to show the theorem in the case where f is a characteristic function on a disjoint rectangle with the set D. That is,

$$f(t_1, ..., t_n) = \mathbf{1}_{[t_1^{(1)}, t_1^{(2)}) \times \cdots \times [t_n^{(1)}, t_n^{(2)})}(t_1, ..., t_n),$$

where $t_n^{(1)} < t_n^{(2)} \leq t_{n-1}^{(1)} < t_{n-1}^{(2)} \leq \cdots \leq t_2^{(1)} < t_2^{(2)} \leq t_1^{(1)} < t_1^{(2)}$. Then,

$$I_n(f) = \prod_{i=1}^n (B(t_i^{(2)}) - B(t_i^{(1)})).$$

On the other hand, we note that $\tilde{f} = \frac{1}{n!} f$ on the set $t_n < t_{n-1} < \cdots < t_1$. Therefore,

$$\int_a^{t_{n-1}} \tilde{f}(t_1, ..., t_n) dB(t_n) = \frac{1}{n!} \mathbf{1}_{[t_1^{(1)}, t_1^{(2)}) \times \cdots \times [t_{n-1}^{(1)}, t_{n-1}^{(2)})}(t_1, ..., t_{n-1})(B(t_n^{(2)}) - B(t_n^{(1)})),$$

which is an $\mathcal{F}_{t^{(1)}_{n-1}}$-measurable random variable and can be considered as a "constant" stochastic process by integrating on the interval $[t^{(1)}_{n-1}, t^{(2)}_{n-1}]$ with respect to $dB(t_{n-1})$.

Iterating this procedure, we obtain that

$$\int_a^b \cdots \int_a^{t_{n-2}} \left(\int_a^{t_{n-1}} \tilde{f}(t_1, ..., t_n) dB(t_n) \right) dB(t_{n-1}) \cdots dB(t_n) = \frac{1}{n!} \prod_{i=1}^{n} (B(t_i^{(2)}) - B(t_i^{(1)})).$$

$\triangle$

The following relationship between the Wiener-Itô integral and the Hermite polynomials holds:

Theorem 1.12 *For all $f \in L^2(T)$ and $n \geq 1$,*

$$I_n(f^{\otimes n}) = n! H_n(W(f), \|f\|^2_{L^2(T)}),$$

where $f^{\otimes n}$ is the n variables (symmetric) function $f^{\otimes n}(t_1, ..., t_n) = f(t_1) \cdots f(t_n)$.

Proof. We prove this result by induction over n. The case $n = 1$ is immediate. We assume that the result holds for $1, ..., n$. Using Theorem 1.11 we get that

$$I_{n+1}(f^{\otimes n+1}) = (n+1)! \int_a^b f(t_1) X_{t_1} \, dB(t_1),$$

where

$$X_t = \int_a^t \cdots \left(\int_a^{t_n} f(t_2) \cdots f(t_{n+1}) \, dB(t_{n-1}) \right) \cdots dB(t_2).$$

Appealing again to Theorem 1.11 and the induction hypothesis, we have that

$$\begin{aligned} X_t &= \frac{1}{n!} \int_{[a,t]^n} f(t_2) \cdots f(t_{n+1}) dB(t_2) \cdots dB(t_{n+1}) \\ &= H_n \left(\int_a^t f(s) dB(s), \int_a^t f^2(s) ds \right). \end{aligned}$$

Hence, we obtain the equality

$$I_{n+1}(f^{\otimes n+1}) = (n+1)! \int_a^b f(t_1) H_n \left(\int_a^{t_1} f(s) dB(s), \int_a^{t_1} f^2(s) ds \right) dB(t_1). \tag{1.11}$$

On the other hand, if we apply Itô's formula to the function $H_{n+1}(x, \lambda)$, it yields that

$$\begin{aligned} dH_{n+1} \left(\int_a^t f(s) dB(s), \int_a^t f^2(s) ds \right) &= \left(\frac{\partial}{\partial x} H_{n+1} \right) f(t) dB(t) + \frac{1}{2} \left(\frac{\partial^2}{\partial x^2} H_{n+1} \right) f^2(t) dt \\ &\quad + \left(\frac{\partial}{\partial \lambda} H_{n+1} \right) f^2(t) dt. \end{aligned}$$

Using properties (1.1) and (1.4), we obtain that

$$dH_{n+1} \left(\int_a^t f(s) dB(s), \int_a^t f^2(s) ds \right) = f(t) H_n \left(\int_a^t f(s) dB(s), \int_a^t f^2(s) ds \right) dB(t).$$

Integrating over $[a, b]$, we get that

$$H_{n+1}(W(f), \|f\|_{L^2(T)}) = \int_a^b f(t) H_n\left(\int_a^t f(s)dB(s), \int_a^t f^2(s)ds\right) dB(t). \quad (1.12)$$

Using inequalities (1.11) and (1.12) together, we conclude the desired result for $n+1$. $\triangle$

As a consequence of Theorem 1.12 we obtain the following version of Theorem 1.5 on the Wiener chaos expansion.

Theorem 1.13 *Any random variable $F \in L^2(\Omega, \mathcal{G}, \mathrm{P})$ admits a decomposition of the form $F = \sum_{n=0}^{\infty} I_n(f_n)$, where $f_0 = \mathrm{E}[F]$, I_0 is the identity mapping on the constants, and the functions $f_n \in L^2(T^n)$ are symmetric and uniquely determined by F.*

1.4 The derivative operator

The goal of this chapter is to define the notion of derivative operator DF of a square integrable random variable $F : \Omega \mapsto \mathbb{R}$, where the derivative is taken with respect to the parameter $\omega \in \Omega$. We will use a notion of derivative in a weak sense, without assuming any topological structure on the space Ω.

1.4.1 Definition and properties of the derivative operator

We consider the set of smooth random variables $\mathcal{S}$:

$$\mathcal{S} = \{F = f(W(h_1), ..., W(h_n)), f \in \mathcal{C}_p^{\infty}(\mathbb{R}^n), h_i \in H, n \geq 1\},$$

where $\mathcal{C}_p^{\infty}(\mathbb{R}^n)$ is the set of functions $f : \mathbb{R}^n \to \mathbb{R}$ that are $\mathcal{C}^{\infty}$ and such that f and all its partial derivatives have polynomial growth. We will use the notation $\mathcal{S}_b$ for the case where $f \in \mathcal{C}_b^{\infty}(\mathbb{R}^n)$, that is, f and all its partial derivatives are bounded. Observe that $\mathcal{S}_b \subset \mathcal{S}$ and $\mathcal{S}_b$ is dense in $L^2(\Omega)$.

If $F \in \mathcal{S}$ we define the derivative of F, DF, as the H-valued random variable:

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

For example: $D(W(h)) = h$.

The derivative DF can be interpreted as the a directional derivative: for any $h \in H$, we have

$$\langle DF, h \rangle_H = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(f(W(h_1) + \epsilon \langle h_1, h \rangle_H, ..., W(h_n) + \epsilon \langle h_n, h \rangle_H) - F),$$

that is, $\langle DF, h \rangle_H$ is the derivative at $\epsilon = 0$ of F composed with the shifted process $\{W(g) + \epsilon \langle g, h \rangle_H, g \in H\}$.

Example 1.14 *Let $(B(t), t \in [0, 1])$ be a one-dimensional Brownian motion defined on its canonical probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Let $H = L^2([0, 1], \mathcal{B}([0, 1]), \lambda)$, where λ is the Lebesgue measure in $[0, 1]$. For each $h \in H$, set $W(h) = \int_{[0,1]} h(s)dB(s)$.*

Let $F = f(W(\mathbf{1}_{[0,t_1]}), ..., W(\mathbf{1}_{[0,t_n]})) \in \mathcal{S}$, $0 \leq t_1 < \cdots < t_n \leq 1$. Then, for each $h \in H$,

$$\langle DF, h \rangle_H = \sum_{i=1}^{n} \partial_i f(B(t_1), ..., B(t_n)) \int_0^{t_i} h(t)dt = \frac{d}{d\epsilon} F\left(\omega + \epsilon \int_0^{\cdot} h(t)dt\right)\Big|_{\epsilon=0}.$$

We define the Cameron-Martin space as the subspace H^1 of Ω:

$$H^1 = \{\tilde{h} : [0,1] \mapsto \mathbb{R} \text{ continuous} : \tilde{h}(t) = \int_0^t h(s)ds, h \in H\}.$$

Then, for any $h \in H$, the product $\langle DF, h\rangle_H$ is the directional derivative of F in the direction of the element $\int_0^{\cdot} h(s)ds \in H^1$. We observe that H^1 is a Hilbert space isomorphic to H, writing

$$\langle \tilde{h}, \tilde{g}\rangle_{H^1} = \langle h, g\rangle_H = \int_0^1 h(s)g(s)ds.$$

The following integration by parts formula holds:

Lemma 1.15 *Let $F \in \mathcal{S}$ and $h \in H$. Then,*

$$\mathrm{E}[\langle DF, h\rangle_H] = \mathrm{E}[FW(h)].$$

Proof. It is sufficient to consider the case where $\|h\|_H = 1$. There exists and orthonormal family of H, $\{e_1, ...e_n\}$ such that $h = e_1$ y $F = f(W(e_1), ..., W(e_n))$, where $f \in \mathcal{C}_p^\infty(\mathbb{R}^n)$. Let $\phi(x)$ denote the standard normal distribution on $\mathbb{R}^n$ standard, that is,

$$\phi(x) = (2\pi)^{-n/2} \exp(-\frac{1}{2}\sum_{i=1}^n x_i^2).$$

Then, using the usual integration by parts formula, we have

$$\mathrm{E}[\langle DF, h\rangle_H] = \int_{\mathbb{R}^n} \partial_1 f(x)\phi(x)dx = \int_{\mathbb{R}^n} f(x)\phi(x)x_1 dx = \mathrm{E}[FW(e_1)] = \mathrm{E}[FW(h)].$$

△

Applying the previous result to FG and using that $D(FG) = DFG + FDG$, we obtain:

Lemma 1.16 *Let $F, G \in \mathcal{S}$ and $h \in H$. Then,*

$$\mathrm{E}[G\langle DF, h\rangle_H] = -\mathrm{E}[F\langle DG, h\rangle_H] + \mathrm{E}[FGW(h)].$$

As a consequence of the above lemma we obtain the following result:

Proposition 1.17 *For any $p \geq 1$, the operator D is closable from $L^p(\Omega)$ to $L^p(\Omega; H)$.*

Proof. Let $\{F_N, N \geq 1\}$ be a sequence of random variables in $\mathcal{S}$ such that F_N converges to zero in $L^p(\Omega)$ and the sequence of derivatives DF_N converges to η in $L^p(\Omega; H)$.

Then, for any $h \in H$ and $F \in \mathcal{S}_b$ such that $FW(h)$ is bounded (for example, $F = Ge^{-\epsilon W(h)^2}$, where $G \in \mathcal{S}_b$ and $\epsilon > 0$), we have that

$$\mathrm{E}[F\langle \eta, h\rangle_H] = \lim_{N\to\infty} \mathrm{E}[F\langle DF_N, h\rangle_H] = \lim_{N\to\infty} \mathrm{E}[-F_N\langle DF, h\rangle_H + F_N FW(h)] = 0,$$

because F_N converges to zero in $L^p(\Omega)$ when N converges to infinity and the random variables $\langle DF, h\rangle_H$ and $FW(h)$ are bounded. This implies that $\eta = 0$. △

From now on, we will denote the closed extension of the derivative operator $\bar{D}$ also by D. For any $p \geq 1$, the domain of this operator D is the space $\mathbb{D}^{1,p}$, defined as the completion of $\mathcal{S}$ with respect to the norm:

$$\|F\|_{1,p} = \left\{\mathrm{E}[|F|^p + \mathrm{E}[\|DF\|_H^p]\right\}^{\frac{1}{p}}.$$

We observe that $\mathbb{D}^{1,2}$ is a Hilbert space with respect to the scalar product

$$\langle F,G\rangle_{1,2} = \mathrm{E}[FG] + \mathrm{E}[\langle DF, DG\rangle_H].$$

We define the <u>second iterated derivative</u> of $F \in \mathcal{S}$, $D^2F = D(DF)$, as the $H\otimes H$-valued random variable defined by

$$D^2F = \sum_{i,j=1}^{n} \partial^2_{ij} f(W(h_1),...,W(h_n))(h_i \otimes h_j),$$

where $\otimes$ denotes the tensor product.

In general, we define the <u>k-th iterated derivative</u> of F, D^kF, for any natural number $k \geq 1$, as the $H^{\otimes k}$-valued random variable obtained iterating k times the operator D.

We have the following extension of Proposition 1.17, which can be proved in the same way iterating Lemma 1.16 (see Exercise 1.40):

Proposition 1.18 *For any $p \geq 1$ and $k \geq 1$ natural number, the operator D^k is closable from $\mathcal{S}$ to $L^p(\Omega; H^{\otimes k})$.*

Again, we will use the notation D^k for the closed extension of the iterated derivative $\bar{D}^k$. The domain of the operator D^k is the space $\mathbb{D}^{k,p}$ defined as the completion of $\mathcal{S}$ with respect to the norm:

$$\|F\|_{k,p} = \left\{ \mathrm{E}[|F|^p + \sum_{j=1}^{k} \mathrm{E}[\|D^jF\|^p_{H^{\otimes j}}] \right\}^{\frac{1}{p}}.$$

For $k=0$ we write $\|\cdot\|_{0,p} = \|\cdot\|_p$ and $\mathbb{D}^{0,p} = L^p(\Omega)$.

The following property holds:

Lemma 1.19 *For any $F \in \mathcal{S}$, $1 \leq p \leq q$ and $0 \leq k \leq j$, $k,j \in \mathbb{N}$, $\|F\|_{k,p} \leq \|F\|_{j,q}$. In particular, $\mathbb{D}^{k+1,p} \subset \mathbb{D}^{k,q}$ for all $k \geq 0$ y $p > q$.*

Proof. The case $p=q$ is trivial. If $k=j$, we use Hölder's inequality. $\triangle$

The next result is the <u>chain rule</u>:

Proposition 1.20 *Let $g : \mathbb{R}^d \mapsto \mathbb{R}$ be a function in $\mathcal{C}^1$ with bounded partial derivatives. Let $p \geq 1$ and $F = (F^1,...,F^d)$ a random vector such that $F^i \in \mathbb{D}^{1,p}$ for any $i = 1,...,d$. Then $g(F) \in \mathbb{D}^{1,p}$ and*

$$D(g(F)) = \sum_{i=1}^{d} \partial_i g(F) DF^i. \tag{1.13}$$

Proof. We will only prove the case $d=1$. The case $d>1$ can be proved in the same way.

Let $g_\epsilon(x) = g * \psi_\epsilon(x)$, where ψ_ϵ is an approximation of the identity, that is, $\psi_\epsilon(x) = \epsilon^{-1}\psi(\frac{x}{\epsilon})$, $\epsilon > 0$, $x \in \mathbb{R}$, where ψ is a $\mathcal{C}^\infty$ positive function with support in $[-1,1]$ and such that $\int_{\mathbb{R}} \psi(x)dx = 1$. We observe that $g_\epsilon \in \mathcal{C}^\infty$, and is bounded with bounded partial derivatives.

On the other hand, because $F \in (\mathbb{D}^{1,p})^d$, there exists a sequence $\{F_k\}_{k\geq 1}$, $F_k \in \mathcal{S}$ ($F_k = f_k(W(h_1),...,W(h_{n_k}))$, $f_k \in \mathcal{C}_p^\infty(\mathbb{R}^n)$) that converges to F in $L^p(\Omega)$ as $k \to \infty$, and the sequence DF_k converges to DF in $L^p(\Omega; H)$ as $k \to \infty$. Then, using the definition of the derivative, we have that

$$D(g_\epsilon(F_k)) = \sum_{i=1}^{n_k} \partial_i (g_\epsilon \circ f_k)(W(h_1), ..., W(h_{n_k}))h_i = g'_\epsilon(F_k)DF_k.$$

On the other hand, using the triangle inequality, it yields that

$$\begin{aligned}&\|g'_\epsilon(F_k)DF_k - g'(F)DF\|_{L^p(\Omega;H)} \leq \|g'_\epsilon(F_k)(DF_k - DF)\|_{L^p(\Omega;H)} \\ &+ \|(g'_\epsilon(F_k) - g'(F_k))DF\|_{L^p(\Omega;H)} + \|(g'(F_k) - g'(F))DF\|_{L^p(\Omega;H)} := (1) + (2) + (3).\end{aligned}$$

We observe that for any $\epsilon > 0$ and $k \geq 1$ $g'_\epsilon(F_k)$ is bounded a.s. by a constant not depending on ϵ and k, and hence, (1) converges to zero as $k \to \infty$. On the other hand, the dominated convergence theorem implies that for any $k \geq 1$, (2) converges to zero as $\epsilon \to 0$. In the same way, (3) converges to zero as $k \to \infty$.

Thus, $D(g_\epsilon(F_k))$ converges to $g'(F)DF$ in $L^p(\Omega; H)$ as $\epsilon \to 0$ and $k \to \infty$. On the other hand, $g'_\epsilon(F_k)$ converges to $g(F)$ in $L^p(\Omega)$ as $\epsilon \to 0$ and $k \to \infty$. Finally, the closability of the operator D (Proposition 1.17) implies that $g(F) \in \mathbb{D}^{1,p}$ and that $D(g(F)) = g'(F)DF$, which concludes the proof of the proposition for $d = 1$. △

The chain rule can be extended in the case where g is a Lipschitz function (see Proposition 1.23).

The definition of the derivative operator can be extended to a family of smooth random variables taking values on a real separable Hilbert space V:

$$\mathcal{S}_V = \left\{ u = \sum_{j=1}^{n} F_j h_j : F_j \in \mathcal{S}, h_j \in V, n \geq 1 \right\}.$$

For each $k \geq 1$, we define the derivative D^k of $u \in \mathcal{S}_V$ by

$$D^k u = \sum_{j=1}^{n} F^k F_j \otimes h_j.$$

Then D^k is a closable operator from $\mathcal{S}_V$ to $L^p(\Omega; H^{\otimes k} \otimes V)$, for any $p \geq 1$ and we will note the closed extension also by D^k. The domain of the operator D^k is the space $\mathbb{D}^{k,p}(V)$ defined as the completion of $\mathcal{S}_V$ with respect to the norm:

$$\|u\|_{k,p,V} = \left\{ \mathrm{E}[|u|_V^p + \sum_{j=1}^{k} \mathrm{E}[\|D^j u\|^p_{H^{\otimes j} \otimes V}] \right\}^{\frac{1}{p}}.$$

For $k = 0$ we write $\|u\|_{0,p,V} = \{\mathrm{E}[\|u\|_V]\}^{1/p}$ and $\mathbb{D}^{0,p}(V) = L^p(\Omega; V)$.

1.4.2 The derivative operator in the white noise case

As in Section 1.3 we will restrict ourselves in the case of a one-dimensional Brownian motion $(B(t), t \in T)$, $T = [a, b]$, $H = L^2(T)$, and $W(h) = \int_{[a,b]} h(s)dB(s)$ for $h \in H$ (see [N06, Section 1.2.1] for the general white noise case).

Using the identification $L^2(\Omega; H) \cong L^2(T \times \Omega)$ we have that the derivative of a random variable $F \in \mathbb{D}^{1,2}$ is a stochastic process $\{D_t F, t \in T\}$ defined almost surely with respect to the measure $\lambda \times \mathrm{P}$. In this case, we observe that

$$\|DF\|^2_{L^2(\Omega;H)} = \mathrm{E}\left[\int_T (D_t F)^2 dt\right] = \int_T \mathrm{E}[(D_t F)^2] dt = \mathrm{E}[\|D_\cdot F\|^2_H].$$

For example, $D_t(W(h)) = h(t)$, $h \in H$, $t \in T$, and

$$\|D(W(h))\|^2_{L^2(\Omega;H)} = \int_T h(t)^2 dt = \|h\|^2_H.$$

In the same way, we have that $D^k F = \{D^k_{t_1,\dots,t_k} F, t_i \in T\}$ is a $T^k \times \Omega$-measurable function defined $\lambda^k \times \mathrm{P}$-almost everywhere.

The next result gives the decomposition of the derivative in the Wiener chaos (see Exercise 1.41 for the iterated derivative case).

Proposition 1.21 *Let $F \in \mathbb{D}^{1,2}$ be a square integrable random variable with the Wiener chaos decomposition of Theorem 1.13, that is, $F = \sum_{n=0}^\infty I_n(f_n)$, where $f_n \in L^2(T^n)$ are symmetric. Then, F belongs to $\mathbb{D}^{1,2}$ if and only if*

$$\sum_{n=1}^\infty n n! \|f_n\|^2_{L^2(T^n)} < +\infty, \tag{1.14}$$

and in this case we have that

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

Proof. Assume first that $F = I_n(f_n)$, with $f_n \in \mathcal{E}_n$ symmetric. Then

$$D_t F = \sum_{j=1}^n \sum_{i_1,\dots,i_n=1}^k a_{i_1\cdots i_n} \xi_{i_1} \cdots 1_{[\tau_{i_j-1}, \tau_{i_j})}(t) \cdots \xi_{i_n} = n I_{n-1}(f_n(\cdot, t)), \tag{1.15}$$

where $\xi_{i_p} = B(\tau_{i_p}) - B(\tau_{i_p - 1})$ for $p = 1, ..., n$. Note that the symmetry is used in the second equality.

Let $F \in \mathbb{D}^{1,2}$, $F = \sum_{n=0}^\infty I_n(f_n)$, $f_n \in L^2(T^n)$ symmetric. Consider the sequence of partial sums $F_N = \sum_{n=0}^N I_n(f_n)$, $N \geq 0$. Then F_N converges to F in $L^2(\Omega)$ as $N \to \infty$. This implies that the sequence DF_N converges to DF in $L^2(\Omega; H)$ as $N \to \infty$. On the other hand, for each $n \geq 1$, $I_n(f_n)$ is the limit in $L^2(\Omega)$ of the sequence $\{I_n(f_n^k)\}_{k\geq 1}$, where $f_n^k \in \mathcal{E}_n$ and the sequence f_n^k converges to f_n in $L^2(T^n)$ as $k \to \infty$. Hence, the sequence $D(I_n(f_n^k))$ converges to $D(I_n(f_n))$ in $L^2(\Omega; H)$ as $k \to \infty$. Moreover, the sequence $I_{n-1}(f_n^k(\cdot, *))$ converges to $I_{n-1}(f_n(\cdot, *))$ in $L^2(\Omega; H)$ as $k \to \infty$. Finally, using the linearity of the operator D and (1.15) we obtain the desired result. △

We now prove the following technical result.

Lemma 1.22 *Let $\{F_k, k \geq 1\}$ be a sequence of random variables in $\mathbb{D}^{1,2}$ that converges to F in $L^2(\Omega)$ and such that*

$$\sup_k \mathrm{E}[\|DF_k\|^2_H] < +\infty.$$

Then F belongs to $\mathbb{D}^{1,2}$, and the sequence of derivatives $\{DF_k, k \geq 1\}$ converges to DF in the weak topology of $L^2(\Omega; H)$ (that is, for any $G \in L^2(\Omega; H)$, $\langle DF_k, G\rangle_{L^2(\Omega;H)}$ converges to $\langle DF, G\rangle_{L^2(\Omega;H)}$).

Proof. By Proposition 1.21, to show that $F \in \mathbb{D}^{1,2}$ it suffices to check that (1.14) holds true. Consider the Wiener chaos decompositions $F = \sum_{n=0}^{\infty} I_n(f_n)$ and $F_k = \sum_{n=0}^{\infty} I_n(f_{n,k})$. Then, because F_k converges to F in $L^2(\Omega)$ as $k \to \infty$, for all $n \geq 1$, $f_{n,k}$ converges to f_n in $L^2(T^n)$ as $k \to \infty$ (as $n!\|f_{n,k} - f_n\|^2_{L^2(T^n)} \leq \mathrm{E}[(F_k - F)^2]$). Hence, by Fatou's lemma,

$$\begin{aligned}\sum_{n=1}^{\infty} nn!\|f_n\|^2_{L^2(T^n)} &= \sum_{n=1}^{\infty} nn! \lim_{k\to\infty} \|f_{n_k}\|^2_{L^2(T^n)} \\ &\leq \mathrm{liminf}_{k\to\infty} \sum_{n=1}^{\infty} nn!\|f_{n,k}\|^2_{L^2(T^n)} \leq \sup_k \mathrm{E}[\|DF_k\|^2_H] < +\infty,\end{aligned}$$

which proves (1.14), and thus $F \in \mathbb{D}^{1,2}$.

Moreover, because the sequence DF_k is bounded in $L^2(\Omega; H)$, there exists a subsequence $\{F_{k(j)}, j \geq 1\}$ such that the sequence of derivatives $DF_{k(j)}$ converges in the weak topology of $L^2(\Omega; H)$ to some element $\alpha \in L^2(\Omega; H)$. We claim that $\alpha = DF$. Indeed, it suffices to prove that for any random variable G in the Nth Wiener chaos, $N \geq 0$, and for any $h \in H$,

$$\mathrm{E}[\langle \alpha, h\rangle_H G] = \mathrm{E}[\langle DF, h\rangle_H G],$$

as the set of linear combinations of the form $\sum_{i=1}^{p} h_i G_i$ is dense in $L^2(\Omega; H)$.

Set $G = \sum_{m=0}^{N} I_m(g_m)$. Then, using Proposition 1.21, the isometry of Theorem 1.10 (iii) and the convergence of $F_{k(j)}$ towards F in $L^2(\Omega)$, we get that

$$\begin{aligned}\mathrm{E}[\langle \alpha, h\rangle_H G] &= \lim_{j\to\infty} \mathrm{E}[\langle DF_{k(j)}, h\rangle_H G] \\ &= \lim_{j\to\infty} \mathrm{E}\Big[\Big(\sum_{n=1}^{\infty} nI_{n-1}\Big(\int_a^b f_{n,k(j)}(\cdot,t)h(t)dt\Big)\Big)\Big(\sum_{m=0}^{N} I_m(g_m)\Big)\Big] \\ &= \lim_{j\to\infty} \sum_{n=1}^{N} n(n-1)!\Big\langle \int_a^b f_{n,k(j)}(\cdot,t)h(t)dt, g_{n-1}\Big\rangle_{L^2(T^{n-1})} \\ &= \sum_{n=1}^{N} n!\Big\langle \int_a^b f_n(\cdot,t)h(t)dt, g_{n-1}\Big\rangle_{L^2(T^{n-1})} \\ &= \mathrm{E}[\langle DF, h\rangle_H G].\end{aligned}$$

Hence, $DF = \alpha$. Finally, for any weakly convergent subsequence of $\{DF_n, n \geq 1\}$ the limit must be equal to DF by the preceeding argument, which implies the weak convergence of the whole sequence to DF (recall that a bounded sequence with a unique accumulation point converges to this point in a Banach metric space). △

The next result is the extension of the chain rule for Lipschitz functions in the Brownian motion case.

Proposition 1.23 *Let $g : \mathbb{R}^d \mapsto \mathbb{R}$ be a Lipschitz function, that is, for some constant $K > 0$,*

$$|g(x) - g(y)| \leq K\|x - y\|, \quad \textit{for all } x, y \in \mathbb{R}^d.$$

Suppose that $F = (F^1, ..., F^d)$ is a random vector such that $F^i \in \mathbb{D}^{1,2}$ for any $i = 1, ..., d$. Then $g(F) \in \mathbb{D}^{1,2}$, and there exists a random vector $G = (G_1, ..., G_d)$ bounded by K such that

$$D(g(F)) = \sum_{i=1}^{d} G_i DF^i. \tag{1.16}$$

Proof. If $g \in \mathcal{C}^1(\mathbb{R}^d)$, then the result reduces to that of Proposition 1.20 with $G_i = \partial_i g(F)$.

For all $n \geq 1$, let $\alpha_n(x) = n^d \alpha(nx)$, where α is a nonnegative function in $\mathcal{C}^\infty(\mathbb{R}^d)$ whose support is the unit ball and such that $\int_{\mathbb{R}^d} \alpha(x)dx = 1$. Set $g_n = g * \alpha_n$. Then

(i) $\lim_{n\to\infty} g_n(x) = g(x)$ uniformly with respect to x. Indeed,

$$|g_n(x) - g(x)| \leq \int_{\mathbb{R}^d} |g(x-y) - g(x)|\alpha_n(y)dy \leq K \int_{\mathbb{R}^d} \|y\|\alpha_n(y)dy \leq K.$$

(ii) For all $n \geq 1$, $g_n \in \mathcal{C}^\infty$ which follows trivially by the definition of α_n.

(iii) For all $n \geq 1$ and $x \in \mathbb{R}^d$, $\|\nabla g_n(x)\| \leq K$. Indeed, we first observe that because g is Lipschitz,

$$|g(x) - g_n(y)| \leq \int_{\mathbb{R}^d} |g(x-z) - g(y-z)|\alpha_n(z)dz \leq K\|x-y\|.$$

Hence,

$$\|\nabla g_n(x)\| = \sup_{\|\xi\|=1} |\nabla g_n(x) \cdot \xi| = \sup_{\|\xi\|=1} \lim_{h\to 0} \left| \frac{g_n(x) - g_n(x+h\xi)}{h} \right| \leq K.$$

Then for any $n \geq 1$, we have that

$$D(g_n(F)) = \sum_{i=1}^{d} \partial_i g_n(F) DF^i. \tag{1.17}$$

On the other hand, the sequence $g_n(F)$ converges to $g(F)$ in $L^2(\Omega)$ as n tends to infinity, and the sequence $\{D(g_n(F)), n \geq 1\}$ is bounded in $L^2(\Omega; H)$. Hence, by Lemma 1.22, $g(F) \in \mathbb{D}^{1,2}$ and $D(g_n(F))$ converges in the weak topology of $L^2(\Omega; H)$ to $D(g(F))$.

Furthermore, the sequence $\{\nabla g_n(F), n \geq 1\}$ is bounded by K. Hence, there exists a subsequence $\{\nabla g_{n(k)}(F), k \geq 1\}$ that converges to some random vector $G = (G_1, ..., G_d)$ in the weak topology $\sigma(L^2(\Omega; \mathbb{R}^d))$. This implies that G is bounded by K a.s. Indeed, for any $\xi \in \mathbb{R}^d$, $\|\xi\| = 1$ and $A \in \mathcal{B}(\mathbb{R})$,

$$\int_A |G \cdot \xi| \, d\mathrm{P} = \lim_{k\to\infty} \int_A |g_{n(k)}(F) \cdot \xi| \, d\mathrm{P} \leq K\mathrm{P}(A).$$

This implies that

$$\|G\| = \sup_{\|\xi\|=1} |G \cdot \xi| \leq K, \quad \text{a.s.}$$

Finally, taking the limit in (1.17), we obtain (1.16). △

Remark 1.24 *If the law of the random vector is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$, then $G^i = \partial g(F)$ in (1.16).*

We will also need the following technical result which is an extension of Lemma 1.22 (see [N06, Lemma 4.6]).

Lemma 1.25 *Let $\{F_n, n \geq 1\}$ be a sequence of random variables converging to F in $L^p(\Omega)$ for some $p > 1$. Suppose that $\sup_n \|F_n\|_{k,p} < +\infty$, for some $k \geq 1$. Then $F \in \mathbb{D}^{k,p}$.*

We next prove a local property of the derivative operator. For any interval $[c,d] \subset [a,b]$, let $\mathcal{F}_{[c,d]}$ be the σ-algebra generated by the random variables $\{B(s), s \in [c,d]\}$.

Lemma 1.26 *Let F be a random variable such that $F \in \mathbb{D}^{1,2} \cap L^2(\Omega, \mathcal{F}_{[c,d]^c}, \mathrm{P})$. Then $D_t F = 0$, $(\lambda \times \mathrm{P})$-for almost every $(t,\omega) \in [c,d] \times \Omega$.*

Proof. Let $F \in \mathcal{S}$, $F = f(W(|1_{[a_1,b_1]}), ..., W(\mathbf{1}_{[a_n,b_n]}))$, where $f \in \mathcal{C}_p^\infty(\mathbb{R}^n)$, $[a_i, b_i] \subset [c,d]^c$, and $i = 1, ..., n$. Then,

$$D_t F = \sum_{i=1}^n \partial_i f(W(\mathbf{1}_{[a_1,b_1]}), ..., W(\mathbf{1}_{[a_n,b_n]}))\mathbf{1}_{[a_i,b_i]}(t),$$

and hence, $D_t F = 0$, for all $(t,\omega) \in [c,d] \times \Omega$.

In general, if $F \in \mathbb{D}^{1,2} \cap L^2(\Omega, \mathcal{F}_{[c,d]^c}, \mathrm{P})$, there exists a sequence of random variables $\{F_k\}_{k \geq 1}$, $F_k \in \mathcal{S}$ of the above type that converge to F in $L^2(\Omega)$ as $k \to \infty$. In particular, the sequence DF_k converges to DF in $L^2(\Omega, H)$ as $k \to \infty$. Using the argument above we have that $D_t F_k = 0$, for all $(t,\omega) \in [c,d] \times \Omega$, $k \geq 1$, and hence $D_t F = 0$ for all $(t,\omega) \in [c,d] \times \Omega$. $\triangle$

For example, $D_t W([c,d]^c) = 0$, $(\lambda \times \mathrm{P})$-for almost all $(t,\omega) \in [c,d] \times \Omega$. Moreover, if $h \in H$, $D_t W(h) = 0$, $(\lambda \times \mathrm{P})$-for almost every $(t,\omega) \in [a,b]^c \times \Omega$.

We end this section by computing the derivative of the supremum of the Brownian motion on an interval. We start proving that a Brownian motion in $[0,1]$ attains it maximum a.s. on a unique point.

Lemma 1.27 *With probability one the Brownian motion attains it maximum on $[0,1]$ on a unique point.*

Proof. We want to show that the set

$$G = \left\{ \omega : \sup_{t \in [0,1]} B(t) = B(t_1) = B(t_2) \text{ for some } t_1 \neq t_2 \right\}$$

has probability zero. For each $n \geq 0$, we denote by $\mathcal{I}_n$, the class of dyadic intervals of the form $[(j-1)2^{-n}, j2^{-n}]$, with $1 \leq j \leq 2^n$. The set G is included in the countable union

$$\bigcup_{n \geq 1} \bigcup_{I_1, I_2 \in \mathcal{I}_n, I_1 \cap I_2 = \emptyset} \left\{ \sup_{t \in I_1} B(t) = \sup_{t \in I_2} B(t) \right\}.$$

Finally, it suffices to check that for each $n \geq 1$ and for any couple of disjoint intervals I_1, I_2

$$\mathrm{P}\left\{ \sup_{t \in I_1} B(t) = \sup_{t \in I_2} B(t) \right\} = 0.$$

Fix a rectangle $[a,b] \subset [0,1]$. Then it suffices to show that the law of the random variable $\sup_{t\in[a,b]} B(t)$ conditioned on $\mathcal{F}_a$ is continuous. Note that by the reflection principle we have that for all $y > 0$,

$$\mathrm{P}\left\{\sup_{t\in[0,b]} B(t) \geq y\right\} = 2\mathrm{P}\{B(b) > y\} = 2 - 2\Phi\left(\frac{y}{\sqrt{b}}\right), \tag{1.18}$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. Hence, on the set $\{B(a) = x\}$, writing

$$\sup_{t\in[a,b]} B(t) = \sup_{t\in[a,b]} (B(t) - B(a)) + x,$$

we conclude the desired result. △

Lemma 1.28 *Let $B = \{B(t), t \in [0,1]\}$ be a Brownian motion. Consider the random variable $M = \sup_{t\in[0,1]} B(t)$. Then $M \in \mathbb{D}^{1,2}$ and $D_t M = \mathbf{1}_{[0,T]}(t)$, where T is the a.s. unique point where B attains its maximum.*

Proof. We start proving that $M \in \mathbb{D}^{1,2}$. Consider the approximation of M defined by

$$M_n = \max\{B(t_1), ..., B(t_n)\},$$

where $(t_n, n \geq 1)$ is a countable and dense subset of $[0,1]$. Because the function $g_n : \mathbb{R}^n \to \mathbb{R}$ defined as $g_n(x_1, ..., x_n) = \max\{x_1, ..., x_n\}$ is Lipschitz, Proposition 1.23 implies that M_n belongs to $\mathbb{D}^{1,2}$, for all $n \geq 1$. On the other hand, the sequence M_n converges to M in $L^2(\Omega)$. Thus, by Lemma 1.22, it suffices to show that the sequence DM_n is bounded in $L^2(\Omega; H)$. In order to evaluate the derivatives of M_n, we introduce the following sets:

$$A_1 = \{M_n = B(t_1)\},$$
$$\ldots$$
$$A_k = \{M_n \neq B(t_1), ..., M_n \neq B(t_{k-1}), M_n = B(t_k)\},\ 2 \leq k \leq n.$$

By the local property of the operator D (Lemma 1.26), on the set A_k, $DM_n = DB(t_k)$. Hence, we can write

$$DM_n = \sum_{k=1}^{n} \mathbf{1}_{A_k} DB(t_k).$$

Consequently,

$$\mathrm{E}[\|DM_n\|_H^2] \leq \mathrm{E}\left[\sup_{t\in[0,1]} \|DB(t)\|_H^2\right] = 1,$$

and by Lemma 1.22 we conclude that $M \in \mathbb{D}^{1,2}$.

In order to prove the second statement, consider the sequence $T_n = \sum_{k=1}^{n} \mathbf{1}_{A_k} t_k$ of bounded random variables taking values in $[0,1]$. Then T_n converges to T a.s, and DM_n converges to $DM = DB(t)|_{t=T}$, where T is the unique point where B attains it maximum. △

Remark 1.29 *Recall that Lemmas 1.27 and 1.28 also hold for the Brownian sheet defined in Example 1.3, but in this case the explicit form of the density of M is unknown. One can show that this density exists and is infintely differentiable using general criterium of the Malliavin calculus for existence and smoothness of densities that will be studied in Section 2.2. See [N06, Section 2.1.7] for the proof of all this facts.*

1.5 The divergence operator

In this section we introduce the divergence operator which is the dual of the derivative operator and the divergence operator in the white noise case which is called the Skorohod integral.

1.5.1 Definition and properties of the divergence operator

The divergence operator δ is the adjoint of the operator D. That is, δ is an unbounded operator from $L^2(\Omega; H)$ to $L^2(\Omega)$ such that:

(i) The domain of δ, Dom δ, is the set of random variables $u \in L^2(\Omega; H)$ such that for all $F \in \mathbb{D}^{1,2}$, $|\mathrm{E}[\langle DF, u\rangle_H]| \leq c_u \|F\|_{L^2(\Omega)}$.

(ii) If $u \in$ Dom δ, then $\delta(u) \in L^2(\Omega)$ and the following duality relation holds: for any $F \in \mathbb{D}^{1,2}$,

$$\mathrm{E}[F\delta(u)] = \mathrm{E}[\langle DF, u\rangle_H]. \tag{1.19}$$

The following properties of the divergence operator hold:

Proposition 1.30 (i) *If $u \in$ Dom δ, then* $\mathrm{E}[\delta(u)] = 0$.

(ii) *δ is a linear and closed operator in Dom δ.*

(iii) *If $u \in \mathcal{S}_H$, then $u \in$ Dom δ and*

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

(iv) *Let $u \in \mathcal{S}_H$, $F \in \mathcal{S}$ and $h \in H$. Then*

$$\langle D(\delta(u)), h\rangle_H = \langle u, h\rangle_H + \delta\left(\sum_{j=1}^{n} \langle DF_j, h\rangle_H h_j\right).$$

Proof. (i) is immediate applying (1.19) with $F = 1$. The operator δ is closed as the adjoint of an unbounded and densely defined operator. The linearity can be easily deduced from (1.19). This proves (ii). In order to prove (iii) we first observe that if $u \in \mathcal{S}_H$, using Lemma 1.16 we have that for all $F \in \mathbb{D}^{1,2}$,

$$\begin{aligned} \Big|\mathrm{E}[\langle DF, u\rangle_H]\Big| &= \Big|\sum_{j=1}^{n} \mathrm{E}[F_j \langle DF, h_j\rangle_H]\Big| \\ &\leq \sum_{j=1}^{n} \left(\Big|\mathrm{E}[F\langle DF_j, h_j\rangle_H]\Big| + \Big|\mathrm{E}[FF_j W(h_j)]\Big|\right) \\ &\leq c_u \|F\|_{L^2(\Omega)}. \end{aligned}$$

Hence, $u \in$ Dom δ. Moreover, using again Lemma 1.16 and (1.19), we obtain that for all $F \in \mathbb{D}^{1,2}$,

$$\mathrm{E}[F\delta(u)] = \mathrm{E}[\langle DF, u\rangle_H] = \mathrm{E}\left[F\left(\sum_{j=1}^{n} F_j W(h_j) - \sum_{j=1}^{n} \langle DF_j, h_j\rangle_H\right)\right],$$

which implies (iii). Finally, in order to prove (iv), we use (iii) to get that

$$\langle D(\delta(u)), h\rangle_H = \sum_{j=1}^n F_j \langle h_j, h\rangle_H + \sum_{j=1}^n \Big(W(h_j)\langle DF_j, h\rangle_H - \langle D(\langle DF_j, h\rangle_H), h_j\rangle_H \Big),$$

and using again (iii) we conclude the proof of (iv). △

The next result shows that the set of H-valued random variables $\mathbb{D}^{1,2}(H)$ is in the domain of the divergence. We observe that if $u \in \mathbb{D}^{1,2}(H)$, then Du is an $H \otimes H$-valued square integrable random variable, and the Hilbert space $H \otimes H$ can be identified as the space of Hilbert-Schmidt operators from H to H.

Proposition 1.31 *If $u \in \mathbb{D}^{1,2}(H)$, then*

$$\|\delta(u)\|_{0,2}^2 = \mathrm{E}[\delta(u)^2] \le \|u\|_{1,2,H}^2. \tag{1.20}$$

In particular, $\mathbb{D}^{1,2}(H) \subset$ Dom δ and δ is continuous from $\mathbb{D}^{1,2}(H)$ into $L^2(\Omega)$.

Proof. Let $u \in \mathcal{S}_H$ and let $\{e_i, i \ge 1\}$ a complete orthonormal system in H. Using the duality relation and Proposition 1.30(iv) it yields that

$$\begin{aligned}
\mathrm{E}[\delta(u)^2] &= \mathrm{E}[\langle D(\delta(u)), u\rangle_H] \\
&= \mathrm{E}\Big[\sum_{i=1}^\infty \langle D(\delta(u)), e_i\rangle_H \langle u, e_i\rangle_H\Big] \\
&= \mathrm{E}\Big[\sum_{i=1}^\infty \Big\{ \langle u, e_i\rangle_H + \delta\Big(\sum_{j=1}^n \langle DF_j, e_i\rangle_H h_j\Big)\Big\} \langle u, e_i\rangle_H\Big] \\
&= \mathrm{E}[\|u\|_H^2] + \mathrm{E}\Big[\sum_{i=1}^\infty \delta\Big(\sum_{j=1}^n \langle DF_j, e_i\rangle_H h_j\Big) \langle u, e_i\rangle_H\Big] \\
&= \mathrm{E}[\|u\|_H^2] + \mathrm{E}\Big[\sum_{i,k=1}^\infty \langle D(\langle u, e_k\rangle_H), e_i\rangle_H \langle D(\langle u, e_i\rangle_H), e_k\rangle_H\Big].
\end{aligned}$$

Applying the Cauchy-Schwarz inequality, the above expression is

$$\begin{aligned}
&\le \mathrm{E}[\|u\|_H^2] + \mathrm{E}\Big[\sum_{i,k=1}^\infty (\langle D(\langle u, e_i\rangle_H), e_k\rangle_H)^2\Big] \\
&= \mathrm{E}[\|u\|_H^2] + \mathrm{E}\Big[\sum_{i,k=1}^\infty (\langle Du, e_i \otimes e_k\rangle_{H\otimes H})^2\Big] \\
&= \mathrm{E}[\|u\|_H^2] + \mathrm{E}[\|Du\|_{H\otimes H}^2] = \|u\|_{1,2,H}^2.
\end{aligned}$$

Then, if $u \in \mathbb{D}^{1,2}(H)$, there exists a sequence $u_n \in \mathcal{S}_H$ that converges to u in $L^2(\Omega; H)$ and such that the sequence $D(u_n)$ converges to Du in $L^2(\Omega; H \otimes H)$. Hence, the sequence $\delta(u_n)$ converges in $L^2(\Omega)$ and its limit is $\delta(u)$. That is, $u \in$ Dom δ and (1.20) holds for all $u \in \mathbb{D}^{1,2}(H)$. △

More generally, using Meyer inequalities one can prove the following central result in Malliavin calculus (see [N06, Proposition 1.5.7]).

Theorem 1.32 *The operator δ is continuous from $\mathbb{D}^{k,p}(H)$ into $\mathbb{D}^{k-1,p}$, for all $p > 1$ and $k \geq 1$. That is, if $u \in \mathbb{D}^{k,p}(H)$, then*

$$\|\delta(u)\|_{k-1,p} \leq c_{k,p}\|u\|_{k,p,H}.$$

In particular, $\mathbb{D}^{k,p}(H) \subset$ Dom δ.

We will also use the following property of the divergence.

Proposition 1.33 *Let $F \in \mathbb{D}^{1,2}$ and $u \in$ Dom δ such that $Fu \in L^2(\Omega, H)$ and $F\delta(u) - \langle DF, u\rangle_H \in L^2(\Omega)$. Then $Fu \in$ Dom δ and*

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

Proof. For any random variable $G \in \mathcal{S}$, using (1.19), it holds that

$$\mathrm{E}[\langle Fu, DG\rangle_H] = \mathrm{E}[\langle u, D(FG) - GDF\rangle_H] = \mathrm{E}[(F\delta(u) - \langle u, DF\rangle_H)G].$$

This implies the desired result. △

1.5.2 The Skorohod integral

As in Section 1.4.2 we will restrict ourselves in the case of a one-domensional Brownian motion $(B(t), t \in T)$, $T = [a, b]$, $H = L^2(T)$, and $W(h) = \int_{[a,b]} h(s)dB(s)$ for $h \in H$ (see [N06, Section 1.3.2] for the general white noise case).

In this case, the elements of Dom $\delta \subset L^2(\Omega; H) \cong L^2(T \times \Omega)$ are square integrable stochastic processes and the divergence $\delta(u)$ is called the Skorohod integral of the process u.

An element $u \in L^2(T \times \Omega)$ has a Wiener chaos decomposition of the form

$$u(t) = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)), \tag{1.21}$$

where for each $n \geq 1$, $f_n \in L^2(T^{n+1})$ is symmetric in its first n variables.

The next results gives the decomposition of the Skorohod integral in the Wiener chaos.

Proposition 1.34 *Let $u \in L^2(T \times \Omega)$ with the decomposition given in (1.21). Then $u \in$ Dom δ if and only if the series*

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \tag{1.22}$$

converges in $L^2(\Omega)$, where

$$\tilde{f}_n(t_1, ..., t_n, t) = \frac{1}{n+1}\left(f_n(t_1, ..., t_n, t) + \sum_{i=1}^{n} f_n(t_1, ..., t_{i-1}, t, t_{i+1}, ..., t_n, t_i)\right).$$

Proof. Let $n \geq 1$ and $g \in L^2(T^n)$ a symmetric function. We have that

$$\begin{aligned}
\mathrm{E}\left[\int_T u_t D_t(I_n(g))dt\right] &= \sum_{m=0}^{\infty} \int_T \mathrm{E}[I_m(f_m(\cdot,t))nI_{n-1}(g(\cdot,t))]dt \\
&= n\int_T \mathrm{E}[I_{n-1}(f_{n-1}(\cdot,t))I_{n-1}(g(\cdot,t))]dt \\
&= n(n-1)!\int_T \langle f_{n-1}(\cdot,t), g(\cdot,t)\rangle_{L^2(T^{n-1})}dt \\
&= n!\langle f_{n-1}, g\rangle_{L^2(T^n)} = n!\langle \tilde{f}_{n-1}, g\rangle_{L^2(T^n)} \\
&= \mathrm{E}[I_n(\tilde{f}_{n-1})I_n(g)].
\end{aligned}$$

Assume first that $u \in$ Dom δ. Then using (1.19) and the computation above we obtain that for all $n \geq 1$ and $g \in L^2(T^n)$ symmetric,

$$\mathrm{E}[\delta(u)I_n(g)] = \mathrm{E}[\langle u, D(I_n(g))\rangle_H] = \mathrm{E}[I_n(\tilde{f}_{n-1})I_n(g)].$$

This implies that $I_n(g)$ is the projection of $\delta(u)$ in the n-th Wiener chaos, $\mathcal{H}_n$. Hence, the series (1.22) converges in $L^2(\Omega)$ to its sum which is equal to $\delta(u)$.

Reciprocally, we assume that this series converges in $L^2(\Omega)$ and we denote its sum by V. Let $F_N = \sum_{n=0}^N I_n(g_n)$, where $g_n \in L^2(T^n)$ are symmetric and $N \geq 1$. Using the computation above we obtain that for all $N \geq 1$,

$$\mathrm{E}\left[\int_T u_t D_t F_N dt\right] = \sum_{n=1}^{N} \mathrm{E}[I_n(\tilde{f}_{n-1})I_n(g_n)].$$

In particular,

$$\left|\mathrm{E}\left[\int_T u_t D_t F_N dt\right]\right| \leq \|V\|_{L^2(\Omega)}\|F_N\|_{L^2(\Omega)}.$$

Let $F \in \mathbb{D}^{1,2}$, $F = \sum_{n=0}^{\infty} I_n(g_n)$, where the $g_n \in L^2(T^n)$ are symmetric. Then F_N converges to F in $L^2(\Omega)$ as $N \to \infty$ and DF_N converges to DF in $L^2(\Omega, H)$ as $N \to \infty$. Therefore,

$$\left|\mathrm{E}\left[\int_T u_t D_t F dt\right]\right| \leq \|V\|_{L^2(\Omega)}\|F\|_{L^2(\Omega)},$$

which implies that $u \in$ Dom δ. △

The next result shows that when $u \in L^2(T \times \Omega)$ in an adapted processes, the Skorohod integral coincides with the Itô integral with respect to the Brownian motion.

Proposition 1.35 *If $u \in L^2(T \times \Omega)$ is an adapted process then $u \in$ Dom δ. Moreover, $\delta(u)$ coincides with the Itô integral with respect to the Brownian motion, that is,*

$$\delta(u) = \int_a^b u(s)dB(s).$$

Proof. Let u be an elementary adapted process of the form

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

where $F_j \in L^2(\Omega, \mathcal{F}_{t_j}, \mathrm{P})$ and $a \le t_1 < \cdots < t_{n+1} \le b$.

Then we have that for all $j = 1, ..., n$, $F_j \mathbf{1}_{(t_j, t_{j+1}]}(t) \in \text{Dom } \delta$. Indeed, assume first that $F_j \in \mathbb{D}^{1,2}$. Then Proposition 1.33 implies that $F_j \mathbf{1}_{(t_j,t_{j+1}]}(t) \in \text{Dom } \delta$, and appealing to Lemma 1.26 it yields that

$$\delta(F_j \mathbf{1}_{(t_j,t_{j+1}]}(\cdot)) = F_j \delta(\mathbf{1}_{(t_j,t_{j+1}]}(\cdot)) - \int_t D_t F_j \mathbf{1}_{(t_j,t_{j+1}]}(t) dt = F_j(B(t_{j+1}) - B(t_j)).$$

The general case can be proved using a density argument and the fact that δ is a closed operator. Hence, we have shown that $u \in \text{Dom } \delta$ and

$$\delta(u) = \sum_{j=1}^{n} F_j(B(t_{j+1}) - B(t_j)). \tag{1.23}$$

On the other hand, we know that any adapted process in $L^2(T \times \Omega)$ can be approximated by a sequence $\{u_n\}_{n\ge1}$ of elementary and adapted processes in $L^2(T \times \Omega)$. Then by (1.23), $\delta(u_n)$ is the Itô integral of u_n and converges in $L^2(\Omega)$ to the Itô integral of u. Because δ is a closed operator, we deduce that $u \in \text{Dom } \delta$ and $\delta(u)$ is the Itô integral of u. △

Finally, the next results will allow us to compute the derivative of the Skorohod integral of a process.

Proposition 1.36 *Let $u \in \mathbb{D}^{1,2}(H)$. Assume that for any $t \in T$, the process $(D_t u(s), s \in T)$ is in Dom δ and that there exists a version of the process $(\delta(D_t u(s)), t \in T)$ that belong to $L^2(T \times \Omega)$. Then $\delta(u) \in \mathbb{D}^{1,2}$ and for all $t \in T$,*

$$D_t(\delta(u)) = u(t) + \delta(D_t u). \tag{1.24}$$

Proof. Let

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

where for all $n \ge 1$, $f_n \in L^2(T^{n+1})$ is symmetric in its first n variables. Then, using Propositions 1.21 and 1.34 we have that

$$\begin{aligned} D_t(\delta(u)) &= D_t\left(\sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n)\right) = \sum_{n=0}^{\infty}(n+1) I_n(\tilde{f}_n(\cdot, t)) \\ &= u(t) + \sum_{n=0}^{\infty} I_n\left(\sum_{i=1}^{n} f_n(t_1, ..., t_{i-1}, t, t_{i+1}, ..., t_n, t_i)\right) \\ &= u(t) + \sum_{n=0}^{\infty} n I_n(\phi_n(\cdot, t, \cdot)), \end{aligned}$$

where $\phi_n(\cdot, t, \cdot)$ is the symmetrization of the function

$$(t_1, ..., t_n) \mapsto f_n(t_1, ..., t_{n-1}, t, t_n).$$

On the other hand, using again Propositions 1.21 and 1.34, we obtain that

$$\delta(D_t u) = \delta\left(\sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t, s))\right) = \sum_{n=0}^{\infty} n I_n(\phi_n(\cdot, t, \cdot)),$$

which shows the desired result. △

For example, if f is a $\mathcal{C}^1$ function with bounded partial derivatives, it holds that

$$\begin{aligned} D_t\Big(f\Big(\int_a^b B(s)dB(s)\Big)\Big) &= f'\Big(\int_a^b B(s)dB(s)\Big)\Big(B(t)+\int_t^b dB(s)\Big)\mathbf{1}_{[a,b]}(t) \\ &= f'\Big(\int_a^b B(s)dB(s)\Big)B(b)\mathbf{1}_{[a,b]}(t). \end{aligned}$$

1.5.3 The Clark-Ocone formula

Suppose that $B = \{B(t), t \geq 0\}$ is a one-dimensional Brownian motion. The following result is a basic result in Itô calculus which provides an integral representation of any square functional of the Brownian motion.

Theorem 1.37 *Let $F \in L^2(\Omega)$, measurable with respect to B. Then there exists a unique process $u \in L^2_a(\mathbb{R}_+ \times \Omega)$ such that*

$$F = \mathrm{E}[F] + \int_0^\infty u(t)dB(t). \tag{1.25}$$

Proof. It suffices to show that any zero-mean square integrable random variable G that is orthogonal to all stochastic integrals $\int_{\mathbb{R}_+} u(t)dB(t)$, $u \in L^2_a(\mathbb{R}_+ \times \Omega)$ must be zero. Let $u \in L^2_a(\mathbb{R}_+ \times \Omega)$ and set $M_u(t) = \exp(\int_0^t u(s)dB(s) - \frac{1}{2}\int_0^t u^2(s)ds)$. Using Itô's formula we deduce that

$$M_u(t) = 1 + \int_0^t M_u(s)u(s)dB(s).$$

Hence, such a random variable G is orthogonal to the exponentials

$$\mathcal{E}(h) = \exp\Big(\int_0^\infty h(s)dB(s) - \frac{1}{2}\int_0^\infty h^2(s)ds\Big), \qquad h \in L^2(\mathbb{R}_+).$$

Finally, because the random variables $\{e^{W(h)}, h \in L^2(\mathbb{R}_+)\}$ form a total subset of $L^2(\Omega)$ (see the proof of Theorem 1.5), we conclude the desired proof. △

When the random variable belongs to the space $\mathbb{D}^{1,2}$, it turn out that the process u can be identified as the optional projection of the derivative of F. This is called the Clark-Ocone representation formula:

Theorem 1.38 *Let $F \in \mathbb{D}^{1,2}$. Then*

$$F = \mathrm{E}[F] + \int_0^\infty \mathrm{E}[D_tF|\mathcal{F}_t]dB(t). \tag{1.26}$$

Proof. Suppose that F has the representation (1.25) with $u \in L^2_a(\mathbb{R}_+ \times \Omega)$. Then for any $v \in L^2_a(\mathbb{R}_+ \times \Omega)$, using the isometry property of the Itô integral, we can write

$$\mathrm{E}[\delta(v)F] = \int_0^\infty \mathrm{E}[v(s)u(s)]ds. \tag{1.27}$$

On the other hand, by the duality relationship (1.19), and taking into account that v is adapted we obtain

$$\mathrm{E}[\delta(v)F] = \mathrm{E}\left[\int_0^\infty v(t) D_t F dt\right] = \int_0^\infty \mathrm{E}[v(s)\mathrm{E}[D_t F|\mathcal{F}_t]]dt. \tag{1.28}$$

Finally, (1.27) and (1.28) imply that $u(t) = \mathrm{E}[D_t F|\mathcal{F}_t]$, which proves (1.26). $\triangle$

As an example of application of the Clark-Ocone formula, we will find the integral representation of the random variable $M = \sup_{t\in[0,1]} B(t)$.

In Lemmas 1.27 and 1.28 we have proved that $M \in \mathbb{D}^{1,2}$ and $D_t M = 1_{[0,T]}(t)$, where T is the a.s. unique point where B attains in maximum. Hence, using the Clark-Ocone formula (1.26), we have that

$$M = \mathrm{E}[M] + \int_0^1 \mathrm{E}[\mathbf{1}_{[0,T]}(t)|\mathcal{F}_t]dB(t).$$

We now compute the conditional expectation. Using the reflection principle (1.18), we have that

$$\begin{aligned}\mathrm{E}[\mathbf{1}_{[0,T]}(t)|\mathcal{F}_t] &= \mathrm{E}\left[\mathbf{1}_{\{\sup_{t\le s\le 1}(B(s)-B(t)) \ge \sup_{0\le s\le t}(B(s)-B(t))\}} \,\middle|\, \mathcal{F}_t\right] \\ &= 2 - 2\Phi\left(\frac{\sup_{0\le s\le t}(B(s)-B(t))}{\sqrt{1-t}}\right).\end{aligned}$$

We conclude that the integral representation of M is given by

$$M = \mathrm{E}[M] + \int_0^1 \left(2 - 2\Phi\left(\frac{\sup_{0\le s\le t}(B(s)-B(t))}{\sqrt{1-t}}\right)\right)dB(t).$$

1.6 Exercises

Exercise 1.39 *Show that the process $\{H_n(B_t,t), t\ge 0\}$ is a martingale, where $(B_t, t\ge 0)$ is a Brownian motion.*

Exercise 1.40 *Prove Proposition 1.17.*

In the next exercises we consider the isonormal Gaussian process associated to the Brownian motion as in Section 1.3.

Exercise 1.41 *Let $F = \sum_{n=0}^\infty I_n(f_n)$, $f_n \in L^2(T^n)$ symmetric, a random variable in the space $\mathbb{D}^{k,2}$, $k\ge 1$. Show that for all $k \ge 1$,*

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

and

$$\mathrm{E}\left[\|D^k F\|^2_{L^2(T^k)}\right] = \sum_{n=k}^\infty \frac{(n!)^2}{(n-k)!}\|f_n\|^2_{L^2(T^n)}.$$

Exercise 1.42 *Let $F = \sum_{n=0}^\infty I_n(f_n)$, $f_n \in L^2(T^n)$ symmetric, a random variable in the space $\mathbb{D}^{\infty,2} = \cap_k \mathbb{D}^{k,2}$. Show that, for all $n\ge 0$, $f_n = \frac{1}{n!}\mathrm{E}[D^n F]$.*

Exercise 1.43 *Let* $F = \exp(W(h) - \frac{1}{2}\int_{[a,b]} h^2(s)ds)$, $h \in L^2(T)$. *Compute the iterated derivatives of* F *and its expansion on the Wiener chaos.*

Exercise 1.44 *Let* $B = \{B(t), t \in [0,1]\}$ *be a Brownian motion. Compute the expansion on the Wiener chaos of the random variables:*

$$F = \int_0^1 (t^3 B(t)^3 + 2tB(t)^2)dB(t), \qquad G = \int_0^1 te^{B(t)}dB(t).$$

Exercise 1.45 *Using the Clark-Ocone formula find the stochastic integral representation of the random variable* $F = B^3(1)$.

2 The integration by parts formula and applications to regularity of probability laws

The integration by parts formula is a fundamental tool of the Malliavin calculus, which in particular has an important application which is the study of the absolutely continuity and smoothness of the density of the law of random variables on the Wiener space.

2.1 The integration by parts formula

In this section we prove the integration by parts formula in the one-dimensional case and give some applications to the existence and estimates of the density of a random variable on the Wiener space.

Proposition 2.1 *Let F, G two random variables such that $F \in \mathbb{D}^{1,2}$. Let u be an H-valued random variable such that $\langle DF, u\rangle_H \neq 0$ a.s. and $Gu(\langle DF, u\rangle_H)^{-1} \in$ Dom δ. Then for any function $f \in \mathcal{C}^1$ with bounded derivatives, we have that*

$$\mathrm{E}[f'(F)G] = \mathrm{E}[f(F)H(F,G)],$$

where $H(F,G) = \delta(Gu(\langle DF, u\rangle_H)^{-1})$.

Proof. Applying the chain rule (Proposition 1.20) we have that

$$\langle D(f(F)), u\rangle_H = f'(F)\langle DF, u\rangle_H.$$

Using the duality relation (1.19) we obtain that

$$\begin{aligned} \mathrm{E}[f'(F)G] &= \mathrm{E}[\langle D(f(F)), u\rangle_H (\langle DF, u\rangle_H)^{-1} G] \\ &= \mathrm{E}[\langle D(f(F)), u(\langle DF, u\rangle_H)^{-1} G\rangle_H] \\ &= \mathrm{E}[f(F)\delta(Gu(\langle DF, u\rangle_H)^{-1})], \end{aligned}$$

which concludes the desired proof. △

The following observations will be important for the application of Proposition 2.1.

1. If $u = DF$, then the conclusion of Proposition 2.1 is written as

$$\mathrm{E}[f'(F)G] = \mathrm{E}\left[f(F)\delta\left(\frac{GDF}{\|DF\|_H^2}\right)\right]. \tag{2.1}$$

2. If u is a deterministic process it suffices to assume that $G(\langle DF, u\rangle_H)^{-1} \in \mathbb{D}^{1,2}$, as this implies that $Gu(\langle DF, u\rangle_H)^{-1} \in \mathbb{D}^{1,2}(H) \subset \text{Dom } \delta$ (see Proposition 1.31).

An important application of the integration by parts formula is the following expression of the density of a random variable.

Proposition 2.2 *Let F be a random variable such that $F \in \mathbb{D}^{1,2}$. Assume that $\frac{DF}{\|DF\|_H^2} \in$ Dom δ. Then the law of F has a continuous and bounded density function given by*

$$f(x) = \mathrm{E}\left[\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right]. \tag{2.2}$$

Proof. Let $\psi : \mathbb{R} \mapsto \mathbb{R}_+$ be a $\mathcal{C}^1$ function with compact support and let $\phi(y) = \int_{-\infty}^y \psi(x)dx$. Using formula (2.1) with $G = 1$ and $f = \phi$, we have that

$$\mathrm{E}[\psi(F)] = \mathrm{E}\left[\phi(F)\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right].$$

Using an approximation argument this equality is valid for $\psi(x) = \mathbf{1}_{[a,b]}(x)$, where $a < b$. Therefore, by Fubini's theorem, we conclude that

$$\mathrm{P}(F \in [a,b]) = \mathrm{E}\left[\left(\int_{-\infty}^{F} \psi(x)dx\right)\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right] = \int_a^b \mathrm{E}\left[\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right] dx,$$

which implies the desired result. △

We observe that as $\mathrm{E}\left[\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right] = 0$ (Proposition 1.30(i)), we deduce form Proposition 2.2 the following expression for the density:

$$f(x) = -\mathrm{E}\left[\mathbf{1}_{\{F<x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right]. \tag{2.3}$$

As a consequence of Proposition 2.2 we obtain the following estimate of the density:

Lemma 2.3 *Let $F \in \mathbb{D}^{2,4}$ such that $\mathrm{E}[\|DF\|_H^{-8}] < +\infty$. Then the density $f(x)$ satisfies the following estimate:*

$$f(x) \leq (\mathrm{P}\{|F| > |x|\})^{1/2}\left(\|\|DF\|_H^{-1}\|_{0,2} + 3\|D^2F\|_{L^4(\Omega;H\otimes H)}\|\|DF\|_H^{-2}\|_{0,4}\right).$$

Proof. First observe that the hypotheses of the lemma and Proposition 1.31 imply that $\frac{DF}{\|DF\|_H^2} \in \mathbb{D}^{1,2}(H) \subset \mathrm{Dom}\,\delta$, and hence, the hypotheses of Proposition 2.2 hold. Applying the Cauchy-Schwarz inequality to the expression (2.2), we obtain that

$$f(x) \leq (\mathrm{P}\{F > x\})^{1/2}\left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{0,2}.$$

On the other hand, if we apply the Cauchy-Schwarz inequality to the expression (2.3), we get that

$$f(x) \leq (\mathrm{P}\{F < x\})^{1/2}\left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{0,2}.$$

Hence, we conclude that for all $x \in \mathbb{R}$,

$$f(x) \leq (\mathrm{P}\{|F| > |x|\})^{1/2}\left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{0,2}.$$

Using Proposition 1.31, we deduce that

$$\left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{0,2} \leq \mathrm{E}[\|DF\|_H^{-2}] + \mathrm{E}\left[\left\|D\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{H\otimes H}^2\right].$$

We have that

$$D\left(\frac{DF}{\|DF\|_H^2}\right) = \frac{D^2F}{\|DF\|_H^2} - 2\frac{\langle D^2F, DF\otimes DF\rangle}{\|DF\|_H^4},$$

and hence,

$$\left\| D\left(\frac{DF}{\|DF\|_H^2} \right) \right\|_{H\otimes H} \leq 3 \frac{\|D^2F\|_{H\otimes H}}{\|DF\|_H^2}.$$

Applying again the Cauchy-Schwarz inequality, we conclude the desired estimate for $f(x)$. $\triangle$

2.2 Existence and smoothness of densities

The following criterium was proved by Bouleau and Hirsch using techniques of geometric measure theory (see [N06, Section 2.1.3]).

Theorem 2.4 *Let $F = (F^1, ..., F^d)$ be a random vector satisfying the following two condtions:*

(i) $F^i \in \mathbb{D}^{1,2}$ *for all* $i = 1, ..., d$.

(ii) *The <u>Malliavin matrix</u> γ_F defined as*

$$\gamma_F^{ij} = \langle DF^i, DF^j \rangle_H, \;\; 1 \leq i, j \leq d$$

satisfies $\det \gamma_F > 0$ *a.s.*

Then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.

Remark 2.5 *Condition* (i) *in Theorem 2.4 implies that the measure* $((\det \gamma_F) \cdot P) \circ F^{-1}$ *is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$. In other words, the random vector F has an absolutely continous law conditioned by the set $\{\det \gamma_F > 0\}$, that is,*

$$\mathrm{P}\{F \in B, \det \gamma_F > 0\} = 0,$$

for any Borel subset B of $\mathbb{R}^d$ of zero Lebesgue measure.

The regularity of the density requires under stronger conditions, and for this we introduce the following definition.

Definition 2.6 *We say that a random vector $F = (F^1, ..., F^d)$ is <u>nondegenerate</u> if it satisfies the following conditions:*

(i) $F^i \in \mathbb{D}^\infty$, *for all* $i = 1, ..., d$.

(ii) *The matrix γ_F satisfies* $\mathrm{E}[(\det \gamma_F)^{-p}] < \infty$, *for all* $p \geq 2$.

Using the techniques of Malliavin calculus we will establish the following general criterion for the smoothness of densities.

Theorem 2.7 *Let $F = (F^1, ..., F^d)$ a nondegenerate random vector in the sense of Definition 2.6. Then the law of F possesses an infinitely differentiable density.*

In order to prove Theorem 2.7 we will need the following technical result.

Lemma 2.8 *Let γ be a $d \times d$ random matrix such that det $\gamma > 0$ a.s. and* $(\det \gamma)^{-1} \in L^p(\Omega)$, *for all $p \geq 1$. Suppose that the entries γ^{ij} of γ are in $\mathbb{D}^\infty$. Then $(\gamma^{-1})^{ij}$ belongs to $\mathbb{D}^\infty$ for all i, j, and*

$$D(\gamma^{-1})^{ij} = -\sum_{k,l=1}^{d} (\gamma^{-1})^{ik} (\gamma^{-1})^{lj} D\gamma^{kl}. \tag{2.4}$$

Proof. For any $\epsilon > 0$, define

$$\gamma_\epsilon^{-1} = \frac{\det\,\gamma}{\det\,\gamma + \epsilon}\gamma^{-1}.$$

Note that $(\det\,\gamma + \epsilon)^{-1} \in \mathbb{D}^\infty$ as it can be expressed as the composition of $\det\,\gamma$ with a function in $\mathcal{C}_p^\infty(\mathbb{R})$. Therefore, the entries of γ_ϵ^{-1} belong to $\mathbb{D}^\infty$. Furthermore, for any i, j, $(\gamma_\epsilon^{-1})^{ij}$ converges in $L^p(\Omega)$ to $(\gamma^{-1})^{ij}$ as ϵ tends to zero. Then, in order to check that the entries of γ^{-1} belong to $\mathbb{D}^\infty$ it suffices to show, taking into account Lemma 1.25, that the iterated derivatives of $(\gamma_\epsilon^{-1})^{ij}$ are bounded in $L^p(\Omega)$, uniformly with respecto to ϵ, for any $p \geq 1$. This boundedness in $L^p(\Omega)$ holds from the Leibnitz rule for the operator D^k, that is,

$$D^k_{t_1,\ldots,t_k}(FG) = \sum_{I\subset\{t_1,\ldots,t_k\}} D_I^{|I|}(F)D_{I^c}^{k-|I|}(G), \quad F, G \in \mathcal{S},$$

because $(\det\,\gamma)\gamma^{-1} \in \mathbb{D}^\infty$, and, on the other hand, $(\det\,\gamma + \epsilon)^{-1}$ has bounded $\|\cdot\|_{k,p}$ norms for all k, p, due to our hypotheses.

Finally, form the expression $\gamma_\epsilon^{-1}\gamma = \frac{\det\,\gamma}{\det\,\gamma+\epsilon}I$, we deduce equation (2.4) by first applying the derivative operator D and then letting ϵ tend to zero. △

We next state and prove the integration by parts formula in the multi-dimensional case.

Proposition 2.9 *Let $F = (F^1, \ldots, F^d)$ be a nondegenerate random vector. Let $G \in \mathbb{D}^\infty$ and $g \in \mathcal{C}_p^\infty(\mathbb{R}^d)$. Then for any multiindex $\alpha \in \{1, \ldots, d\}^k$, $k \geq 1$, there exists an element $H_\alpha(F, G) \in \mathbb{D}^\infty$ such that*

$$\mathrm{E}[\partial_\alpha g(F)G] = \mathrm{E}[g(F)H_\alpha(F, G)], \tag{2.5}$$

where the random variable $H_\alpha(F, G)$ is recursively given by

$$H_{(i)}(F, G) = \sum_{j=1}^d \delta\Big(G(\gamma_F^{-1})^{ij} DF^j\Big),$$
$$H_\alpha(F, G) = H_{\alpha_k}(F, H_{(\alpha_1,\ldots,\alpha_{k-1})}(F, G)).$$

Proof. By the chain rule (Proposition 1.20), we have, for all $j = 1, \ldots, d$

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

and, consequently, for all $i = 1, \ldots, d$,

$$\partial_i g(F) = \sum_{j=1}^d \langle D(g(F)), DF^j\rangle_H (\gamma_F^{-1})^{ij}.$$

Taking expectation and using the duality relation (1.19) we get

$$\mathrm{E}[\partial_i g(F)G] = \mathrm{E}[g(F)H_{(i)}(F, G)].$$

Notice that the continuity of the operator δ (Theorem 1.32) and Lemma 2.8 imply that $H_{(i)}(F, G) \in \mathbb{D}^\infty$ (note that Definition 2.6 (ii) implies that $\det\,\gamma_F > 0$ a.s.). Finally, the expression for $H_\alpha(F, G)$ follows by recurrence. △

Proof of Theorem 2.7. Equality (2.5) applied to the multiindex $\alpha = (1, ..., d)$ yields

$$\mathrm{E}[G\partial_\alpha g(F)] = \mathrm{E}[g(F)H_\alpha(F, G)].$$

Notice that

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

Hence, by Fubini's theorem we can write

$$\mathrm{E}[G\partial_\alpha g(F)] = \int_{\mathbb{R}^d} \partial_\alpha g(x) \mathrm{E}[\mathbf{1}_{\{F>x\}} H_\alpha(F, G)] \, dx. \tag{2.6}$$

We can take $\partial_\alpha g$ any function in $\mathcal{C}_0^\infty(\mathbb{R}^d)$. Then equation (2.6) implies that the random vector F has a density given by

$$p(x) = \mathrm{E}[\mathbf{1}_{\{F>x\}} H_\alpha(F, 1)].$$

Moreover, for any multiindex β, we have

$$\begin{aligned}\mathrm{E}[\partial_\beta \partial_\alpha g(F)] &= \mathrm{E}[g(F) H_\beta(F, H_\alpha(F, 1))] \\ &= \int_{\mathbb{R}^d} \partial_\alpha g(x) \mathrm{E}[\mathbf{1}_{\{F>x\}} H_\beta(H_\alpha)] \, dx.\end{aligned}$$

Hence, for any $f \in \mathcal{C}_0^\infty(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \partial_\beta f(x) p(x) dx = \int_{\mathbb{R}^d} \mathrm{E}[\mathbf{1}_{\{F>x\}} H_\beta(F, H_\alpha(F, 1))] \, dx.$$

Therefore, $p(x)$ is infinitely differentiable and for any multiindex β we have that

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

This completes the desired proof. △

2.3 Application to diffusion processes: Hörmander's theorem

Fix $T > 0$ and let $(B(t) = (B^1(t), ..., B^d(t)), t \in [0, T])$ be a d-dimensional Brownian motion defined on its canonical probability space $(\Omega, \mathcal{F}, \mathrm{P})$, and consider its associated isonormal Gaussian process with $H = L^2([0, T]; \mathbb{R}^d)$.

Let $X = (X(t), t \in [0, T])$ be the solution of the following d-dimensional system of SDEs:

$$dX_i(t) = \sum_{j=1}^d \sigma_{ij}(X(t)) dB^j(t) + b_i(X(t)) dt, \quad X_i(0) = x_0^i, \quad i = 1, ..., d, \tag{2.7}$$

where $\sigma_j : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are measurable functions satisfying the following globally Lipschitz condition:

$$\|\sigma_j(x) - \sigma_j(y)\| + \|b(x) - b(y)\| \le K\|x - y\|, \quad \text{for all } x, y \in \mathbb{R}^d. \tag{2.8}$$

Here σ_j denote the columns of the matrix $\sigma = (\sigma_{ij})_{1 \le i,j \le d}$.

The next results show that there is a unique continuous solution to this equation, such that for all $t \in [0, T]$ and for all $i = 1, ..., d$, the random variable $X_i(t)$ belongs to the space $\mathbb{D}^{1,p}$ for all $p \ge 2$ if the coefficients are continuously differentiable. Moreover, if the coefficients are infinitely differentiable with bounded partial derivatives, then $X_i(t) \in \mathbb{D}^\infty$. The first result is standard so it is left for exercise (see Exercise 2.15).

Theorem 2.10 *There exists a unique continuous solution $X = \{X(t), t \in [0,T]\}$ to equation* (2.7)*. Moreover,*

$$\mathrm{E}\left[\sup_{0\le t\le T} |X(t)|^p\right] \le C,$$

for any $p \ge 2$, where $C > 0$ is a positive constant depending on p, T, K.

Theorem 2.11 *Let $X = \{X(t), t \in [0,T]\}$ be the solution to equation* (2.7) *and assume that its coefficients are continuously differentiable functions. Then $X_i(t)$ belongs to $\mathbb{D}^{1,\infty}$ for any $t \in [0,T]$ and $i = 1, ..., d$. Moreover,*

$$\sup_{0\le r\le t} \mathrm{E}\left[\sup_{r\le s\le T} |D_r^j X_i(s)|^p\right] < \infty,$$

and the derivative $D_r^j X_i(s)$ satisfies the following linear equation:

$$D_r^j X_i(t) = \sigma_{ij}(X(r)) + \sum_{k,\ell=1}^{d} \int_r^t \partial_k \sigma_{i\ell}(X(s)) D_r^j(X_k(s)) dB^\ell(s) + \sum_{k=1}^{d} \int_r^t \partial_k b_i(X(s)) ds, \quad (2.9)$$

for $r \le t$ a.e., and $D_r^j X_i(t) = 0$ for $r > t$ a.e.

Furthermore, if the coefficients are assumed to be infinitely differentiable with bounded partial derivatives of all orders greater than or equal to one, then $X_i(t)$ belongs to $\mathbb{D}^\infty$ for all $t \in [0,T]$ and $i = 1, ..., d$.

Proof. Consider the Picard approximations given by

$$X_i^0(t) = x_0^i,$$
$$X_i^{n+1}(t) = x_0^i + \sum_{j=1}^{d} \int_0^t \sigma_{ij}(X^n(s)) dB^j(s) + \int_0^t b_i(X^n(s)) ds, \ n \ge 0.$$

We will prove the following property by induction on n:

(P) $X_i^n(t) \in \mathbb{D}^{1,\infty}$, for all $i = 1, ..., d$, $n \ge 0$, and $t \in [0,T]$; furthermore, for all $p > 1$, we have

$$\psi_n(t) := \sup_{0\le r\le t} \mathrm{E}\left[\sup_{s\in[r,t]} |D_r X^n(s)|^p\right] < \infty, \quad (2.10)$$

and for some constants c_1, c_2,

$$\psi_{n+1}(t) \le c_1 + c_2 \int_0^t \psi_n(s) ds. \quad (2.11)$$

Clearly, **(P)** holds for $n = 0$. Suppose it is true for n. Applying Proposition 1.20, we get

$$D_r(\sigma_j(X^n(s))) = \sum_{k=1}^{d} \partial_k \sigma_j(X^n(s)) D_r^j(X_k^n(s)) \mathbf{1}_{\{r\le s\}}, \ \text{and}$$
$$D_r(b(X^n(s))) = \sum_{k=1}^{d} \partial_k b(X^n(s)) D_r^j(X_k^n(s)) \mathbf{1}_{\{r\le s\}}.$$

Thus the processes $\{D_r(\sigma_j(X^n(s))), s \geq r\}$ and $\{D_r(b(X^n(s))), s \geq r\}$ are adapted and satisfy,

$$|D_r(\sigma_j(X^n(s)))| \leq C|D_r X^n(s)|, \qquad |D_r(b(X^n(s))) \leq |C|D_r X^n(s)|.$$

Using Proposition 1.36 we deduce that the Itô integral $\sum_{j=1}^d \int_0^t \sigma_{ij}(X^n(s))dB^j(s)$ belongs to $\mathbb{D}^{1,2}$, for each $i = 1, ..., d$. Indeed, the hypothesis that the process $(D_t u(s), s \in T)$ is in Dom δ follows from the square integrability, the adaptability and Proposition 1.35. On the other hand, the fact that the process $(\delta(D_t(u(s))), t \in T))$ follows from the Itô isometry and the induction hypothesis (2.10). Moreover, by (1.24), for $r \leq t$ and $\ell = 1, ..., d$,

$$D_r^\ell\Big(\int_0^t \sigma_{ij}(X^n(s))dB^j(s)\Big) = \delta_{\ell,j}\sigma_{\ell j}(X^n(r)) + \int_r^t D_r^\ell(\sigma_{ij}(X^n(s)))dB^j(s).$$

On the other hand, $\int_0^t b_i(X^n(s))ds \in \mathbb{D}^{1,2}$, and, for $r \leq t$,

$$D_r^\ell\Big(\int_0^t b_i(X^n(s))dB^j(s)\Big) = \int_r^t D_r^\ell(b_i(X^n(s)))ds.$$

From these equalities, applying Doob's maximal inequality (Proposition A.1), Burkhölder's inequality (Proposition A.2), and Hölder's inequality, we obtain that

$$\mathrm{E}\Big[\sup_{s\in[r,t]} |D_r^j X^{n+1}(s)|^p\Big] \leq c_p\Big(\gamma_p + T^{p-1}C^p \int_r^t \mathrm{E}[|D_r^j X^n(s)|^p]ds\Big),$$

where

$$\gamma_p := \sup_{n,j} \mathrm{E}\Big[\sup_{t\in[0,T]} |\sigma_j(X^n(t))|^p\Big] < \infty.$$

In particular, by hypothesis (2.10), $X_i^{n+1}(s) \in \mathbb{D}^{1,\infty}$, for all $t \in [0,T]$. So (2.10) and (2.11) hold for $n+1$. We know that

$$\mathrm{E}\Big[\sup_{s\leq T} |X^n(s) - X(s)|^p\Big] \to 0, \quad \text{as } n \to \infty.$$

Hence, by Gronwall's lemma applied to (2.11) we deduce that the derivatives of the sequence $X_i^n(t)$ are bounded in $L^p(\Omega, H)$ uniformly in n for all $p \geq 2$. Thus, from Lemma 1.25, we deduce that the random variables $X_i(t)$ belong to $\mathbb{D}^{1,2}$. Finally, applying the operator D to equation (2.7) and using Propositions 1.20 and 1.36 as above we deduce (2.9).

The proof that $X^i(t) \in \mathbb{D}^\infty$ if the coefficients are smooth follows by a recursive argument. See [N06, Theorem 2.2.2] for its proof. △

Remark 2.12 *If the coefficients are just Lipschitz (satisfy* (2.8)*), then we will still have that* $X_i(t) \in \mathbb{D}^{1,\infty}$*, and equation* (2.9) *holds with* $\partial_k\sigma_{i\ell}(X(s))$ *and* $\partial_k b_i(X(s))$ *replaced by some bounded and adapted processes. The proof of this extension follows in the same way as before (see [N06, Theorem 2.2.1]), but appealing to Proposition 1.23.*

We next introduce the Hörmander's condition. For this, we assume that the coefficients of equation (2.7) are infinitely differentiable with bounded partial derivatives of all orders greater than or equal to one.

To introduce this condition we consider the following vector fields on $\mathbb{R}^d$ associated with the coefficients of equation (2.7):

$$\sigma_j = \sum_{i=1}^{d} \sigma_{ij}(x)\frac{\partial}{\partial x_i},\ j = 1, ..., d,\ b = \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i}.$$

The covariant derivative of σ_k in the direction of σ_j is defined as the vector field

$$\sigma_j \nabla \sigma_k = \sum_{i,\ell=1}^{d} \sigma_{\ell j}\partial_l \sigma_{ik}\frac{\partial}{\partial x_i},$$

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

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

Consider the vector field on $\mathbb{R}^d$,

$$\sigma_0 = b - \frac{1}{2}\sum_{i=1}^{d} \sigma_i \nabla \sigma_i,$$

which appears when we write the Itô SDE (2.7) in terms of a Stratonovich integral:

$$X(t) = X_0 + \sum_{j=1}^{d} \int_0^t \sigma_j(X(s)) \circ dB^j(s) + \int_0^t \sigma_0(X(s))ds.$$

We are now ready to state Hörmander's condition:

(H) The vector space spanned by the vector fileds:

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

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

Example 2.13 *If $m = d = 1$, $\sigma_{11}(x) = \sigma(x)$, and $\sigma_{10}(x) = b(x)$, then Hörmander's condition means that $\sigma(x_0) \neq 0$ or $\sigma^n(x_0)b(x_0) \neq 0$ for some $n \geq 1$.*

The next theorem can be considered as a probabilistic version of Hörmander's theorem on the hypoellipticity of second-order differential operators.

Theorem 2.14 *Assume that Hörmander's condition* **(H)** *holds and that the coefficients of equation* (2.7) *are infinitely differentiable with bounded partial derivatives of all orders greater than or equal to one. Then for any $t > 0$, $X(t)$ has an infinitely differentiable density.*

The proof of this theorem is based on Norris Lemma ([N06, Lemma 2.3.2]) which essentially shows that when the quadratic variation or the bounded variation part of a continuous semimartingale is large, then the semimartingale is small with an exponentially small probability. The complete proof of Theorem 2.14 is given in [N06, Section 2.3], and [N09, Chapter 7] gives a useful sketch of the proof.

2.4 Exercises

Exercise 2.15 *Prove Theorem 2.10.*

Exercise 2.16 *Set* $M_t = \int_0^t u_s dB(s)$, *where* $B = \{B(s), s \in [0,T]\}$ *is a one-dimensional Brownian motion and* $u = \{u(t), t \in [0,T]\}$ *is an adapted process such that* $|u(t)| \geq \rho > 0$ *for some constant* ρ, $\mathrm{E}\left(\int_0^T u^2(t)dt \right) < \infty$, $u(t) \in \mathbb{D}^{2,2}$ *for each* $t \in [0,T]$, *and*

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

for some $p > 3$. *Show that the density of* M_t *denoted by* $p_t(x)$ *satisfies*

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

for all $t > 0$, *where* $q > p/(p-3)$ *and the constant* c *depends on* λ, ρ *and* p.

Exercise 2.17 *Show that the random variable* $F = \int_0^1 t^2 \arctan(B(t))\, dt$, *where* B *is a Brownian motion, has a* $\mathcal{C}^\infty$ *density.*

Exercise 2.18 *Let* $m = 3$, $d = 2$, *and* $x_0 = 0$, *and consider the vector fields:*

$$\sigma_1(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \sigma_2(x) = \begin{pmatrix} 0 \\ \sin x_2 \\ x_1 \end{pmatrix}, \ b(x) = \begin{pmatrix} 0 \\ \frac{1}{2}\sin x_2 \cos x_2 + 1 \\ 0 \end{pmatrix}.$$

Show that the solution to the SDE $X(t)$ *associated to these coefficients has a* $\mathcal{C}^\infty$ *density for any* $t > 0$.

3 Applications of Malliavin calculus in mathematical finance

3.1 Pricing and hedging financial options

An option is a financial contract that allows to buy or sell (if it is a put or call option) a certain quantity of a financial asset (which can be a financial stock, a currency, a benefit, etc...) at an exercise price K (*strike price*) and time exercise T (*maturity*), which are fixed in the contract. If the buyer exercises the right granted by the option, the writer has the obligation to purchase the underlying asset at the strike price. In exchange for having this option, the buyer pays the writer a fee x (the *option premium*).

The terms for exercise differ depending on the option style:

- An European option can only be exercised at the established time T.
- An American option allows exercise at any time before the expiration.

We will use the following notation:

- $(S_t, 0 \leq t \leq T)$ is the observed price in the market of the underlying asset at every instant.
- $(C_t, 0 \leq t \leq T)$ and $(P_t, 0 \leq t \leq T)$ are, respectively, the value of the call and put option at every instant.

We observe that the value of a call and a put option at the exercise time are, respectively,

$$C_T = (S_T - K)_+$$
$$P_T = (K - S_T)_+.$$

This type of options (call y put) are called standard or (*vanilla*) options. There exists more type of options, called exotic options which are more complex. Vanilla options are often used for hedging exotic options.

We will assume the existence of a non-risky asset with constant interest rate $r > 0$, exercise time T, and price at each time $0 \leq t \leq T$, $S_t^0 = e^{rt}$.

We now state the following two questions:

1. How to evaluate at time $t = 0$ the price of an option (premium), that is, the price in the contract where the buyer and the writer need to agree ? This problem is called *pricing* options.
2. How can one produce the value of an option at the maturity from the premium ? This is the problem is called *hedging* options.

In order to solve these problems we need to assume the following hypothesis on the market: absence of arbitrage opportunities, that is, its is impossible to obtain benefits without taking risks.

Lemma 3.1 *The following call-put parity relationship follows:*

$$C_t - P_t = S_t - Ke^{-r(T-t)}, \text{ for all } t < T.$$

Proof. Assume that $C_t - P_t > S_t - Ke^{-r(T-t)}$. At time t, if we sell a call, we buy a put and we buy an action we will obtain a net benefit of $Y_t = C_t - P_t - S_t$. If this quantity is positive, we can put this money at interest rate $r > 0$ till the maturity. On the other hand, if this quantity is negative, we can ask for a loan of this quantity at the same interest rates. When maturity arrives, we have two possibilities: If $S_T > K$, then the call is exercise and we obtain a benefit of $K + e^{r(T-t)}(C_t - P_t - S_t) > 0$. If $S_T \leq K$, the put is exercised and we obtain the same benefit. This is an example of arbitrage which contradict the hypothesis of our market. △

The dynamic pricing and hedging of options consists on the following: assume that the owner of the option can guarantee a flow of $h(S_T)$ at the maturity. Then he will try to use the premium to buy a portfolio of actions with price flow equal to the one of the option. It is called the *hedging portfolio* and its gestion is called the dynamic strategy of selling or buying actions or loan at the bank.

We denote the value at any time of the hedging portfolio by $(V_t, 0 \leq t \leq T)$. In this case the absence of arbitrage is written as: $V_0 = 0$, $V_T \geq 0$ and $\mathrm{P}\{V_T > 0\} > 0$. We observe that $V_0 = x$. Let β_t be the number of actions that the owner of the option has bought at time $0 \leq t \leq T$, and α_t the number of non-risky assets that he owns at time $0 \leq t \leq T$. We assume that the portfolio strategy is self-financing, that is, its manager does not take into account in his decision rule the value of the underlying asset when he renegotiates the value of the portfolio. That is, in small amount of time $[t, t+dt]$, the variation of the value of the portfolio only depends on the variation of the value of the option and the interest obtained on the inverted *cash* at the bank, which will be equal to $V_t - \beta_t S_t = \alpha_t e^{rt}$. Hence, we have that

$$dV_t = \beta_t dS_t + (V_t - \beta_t S_t) r dt = rV_t dt + \beta_t (dS_t - rS_t dt). \tag{3.1}$$

The problem of pricing and hedging the option is finding a self-financing portfolio strategy that replicates the terminal flow $h(S_T)$, that is, that $v(T, S_T) = h(S_T)$, and that each instant covers the derivative product. This problem can be mathematically translated to find two functions $v(t, x)$, $\beta(t, x)$ sufficiently regular and such that

$$\begin{cases} dv(t, S_t) &= v(t, S_t) r dt + \beta(t, S_t)(dS_t - rS_t dt), \\ v(T, S_T) &= h(S_T). \end{cases} \tag{3.2}$$

$\beta(t, S_t)$ is called the hedging portfolio of the derivative product $h(S_T)$.

- The problem of existence of a solution of (3.2) will depend on the model choosed to describe S_t. In the next sections, we will solve this equation under the Black-Scholes model.
- The unicity of the solution is a consequence of the absence of arbitrage hypothesis.
- A market in which for any terminal flow $h(S_T)$ exists a self-financing replicating portfolio is called a complete market.

Example 3.2 *Consider a market with a non-risky asset with $r = 0$ and a risky asset with initial price $S_0 = 10$ and such that*

$$\mathrm{P}\{S_1 = 20\} = p, \ \ \mathrm{P}\{S_1 = 7,5\} = 1 - p.$$

Consider a put with $K = 15$ y $T = 1$, where the time in this case is assumed to be discrete (for example, days). We want to obtain a replicating portfolio. For this, we need to give the initial capital V_0 and the value of the portfolio (α_1, β_1).

Under the self-financing assumption, we have that $V_0 = 10\beta_1 + \alpha_1$. On the other hand, because the portfolio needs to replicate the derivative, we have that

$$V_1 = \begin{cases} 5 & \text{if } S_1 = 20 \\ 0 & \text{if } S_1 = 7,5, \end{cases}$$

that is,

$$\begin{aligned} 20\beta_1 + \alpha_1 &= 5 \\ 7,5\beta_1 + \alpha_1 &= 0. \end{aligned}$$

Hence, we deduce that $\beta_1 = 0,4$, $\alpha_1 = -3$ and $V_0 = 1$. That is, the option price is 1. In order to cover the option with a capital 1, we construct a portfolio in the following way: we ask for a credit of 3 and invest 4 in actions. At time $T = 1$ we have two possibilities:

1. *If $S_1 = 20$, the option is exercised with a cost of 5. We next sell the actions and win $0,4 \times 20 = 8$, and with this money we reimburse the credit and the cost of the option.*
2. *If $S_1 = 7,5$, the option is not exercised. We sell the actions, win $0,4 \times 7,5 = 3$, and pay the credit.*

3.2 The Black-Scholes model

The Black-Scholes model consists in assuming that the price of the underlying asset observed in the market follows the following stochastic differential equation:

$$\begin{cases} dS_t &= S_t(\mu dt + \sigma dB_t), \quad t \in [0,T], \\ S_0 &= x, \end{cases} \tag{3.3}$$

where $(B_t, t \in [0,T])$ is a Brownian motion defined on its canonical probability space $(\Omega, \mathcal{F}, \mathrm{P})$, and μ and σ are constant parameters that can be financially interpreted as follows:

- μ represents the expected annual rate of return of the asset.
- σ is the volatility, it measures the risk and depends on the nature of the underlying asset.

The solution of the SDE (3.3) is given by the process:

$$S_t = S_0 \exp(\mu t - \frac{\sigma^2}{2} t + \sigma B_t). \tag{3.4}$$

Indeed, applying Itô's formula to the function $f(t,y) = x\exp(\mu t - \frac{\sigma^2}{2} t + \sigma y)$ at $y = B_t$ (remark that f is $\mathcal{C}^2$ on y, $\mathcal{C}^1$ on t and has continuous derivatives), we obtain that:

$$f(t, B_t) = f(0, B_0) + \int_0^t f'_s(s, B_s) ds + \int_0^t f'_y(s, B_s) dB_s + \frac{1}{2} \int_0^t f''_{yy}(s, B_s) d\langle B, B\rangle_s.$$

Using (3.4), we find that:

$$S_t = x + \int_0^t S_s \mu ds + \int_0^t S_s \sigma dB_s,$$

which is the integral version of the differential equation (3.3).

We observe the following:

- One can also prove the unicity of equation (3.3).
- In order to justify the integral writing we should prove that $\int_0^t (S_s\sigma)^2 ds < +\infty$ a.s. For this, it suffices to prove that $\mathrm{E}[\int_0^t (S_s\sigma)^2 ds] < +\infty$. This follows from the following result (see Exercise 3.6):

Lemma 3.3 *It holds that*

$$\mathrm{E}[S_t] = xe^{\mu t}; \;\; \mathrm{E}[S_t^2] = x^2 \exp((2\mu + \sigma^2)t).$$

3.3 Pricing and hedging options in the Black-Scholes model

The next result converts equation (3.2) in a partial derivatives equation under the Black-Scholes model, that will be solved below. In particular, this will show that the Black-Scholes model is complete.

Theorem 3.4 *Let h be a continuous function of at most linear growth. Assume that the following PDE admits a regular solution $v(t,y)$ over $]0,T]\times]0,+\infty[$:*

$$\left\{ \begin{array}{l} \frac{1}{2}\sigma^2 y^2 v''_{yy}(t,y) + ryv'_y(t,y) + v'_t(t,y) - rv(t,y) = 0, \\ v(T,y) = h(y). \end{array} \right. \tag{3.5}$$

Then, there exist a portfolio with valuer $v(t,S_t)$ at times t that replicates the flow $h(S_T)$. The value of this hedging portfolio is given by $\beta(t,S_t) = v'_y(t,S_t)$.

Proof. Applying Itô's formula to the function $v(t,S_t)$, we obtain that:

$$dv(t,S_t) = v'_t(t,S_t)dt + v'_y(t,S_t)dS_t + \frac{1}{2}\sigma^2 S_t^2 v''_{yy}(t,S_t)dt.$$

On the other hand, we have seen that the value of a self-financing portfolio satisfies the differential equation:

$$dv(t,S_t) = v(t,S_t)rdt + \beta(t,S_t)(dS_t - rS_t dt).$$

Hence, using the unicity of the representation of an Itô process, we have that:

$$\left\{ \begin{array}{l} v'_y(t,S_t) = \beta(t,S_t) \qquad \text{c.s.} \\ \frac{1}{2}\sigma^2 S_t^2 v''_{yy}(t,S_t) + rS_t v'_y(t,S_t) + v'_t(t,S_t) = v(t,S_t)r, \end{array} \right.$$

which shows the conclusion of the theorem. △

We define the risk premium as the quotient between the benefits mean and the viability of the option, that is, $\lambda = \frac{\mu - r}{\sigma}$. Consider the process $\tilde{B} = (\tilde{B}_t, t \in [0,T])$ defined as

$$\tilde{B}_t = B_t + \lambda t.$$

Consider the martingale

$$M_t = \exp(-\lambda B_t - \frac{1}{2}\lambda^2 t),$$

and note that $\mathrm{E}[M_t] = 1$. Then, by Girsanov's theorem the measure Q defined by $\frac{dQ}{d\mathrm{P}} = M_T$ is a probability measure, equivalent to P, and $\tilde{B}$ is a Brownian motion under Q.

The Black-Scholes model with respect to $\tilde{B}_t$ can be written as:

$$dS_t = S_t(\sigma dB_t + \mu dt) = S_t(\sigma d\tilde{B}_t + rdt), \tag{3.6}$$

and hence,

$$S_t = S_0 \exp(rt - \frac{\sigma^2}{2}t + \sigma\tilde{B}_t).$$

Consider the function

$$u(t, \tilde{B}_t) = e^{-rt}v(t, S_t),$$

where $v(t, S_t)$ is the value of the portfolio at time t that replicates the flow $h(S_T)$ of Theorem 3.4. We observe that

$$v(0,x) = u(0,x), \qquad u(T,y) = e^{-rT}h(x\exp((r - \frac{\sigma^2}{2})T + \sigma y)). \tag{3.7}$$

On the other hand, it holds that:

Lemma 3.5

$$du(t, \tilde{B}_t) = e^{-rt}\beta(t, S_t)S_t\sigma d\tilde{B}_t.$$

Proof. We will use the notation $\beta_t = \beta(t, S_t)$. We have that

$$du(t, \tilde{B}_t) = -re^{-rt}v(t, S_t)dt + e^{-rt}dv(t, S_t).$$

On the other hand, $v(t, S_t) = \alpha_t e^{rt} + \beta_t S_t$. Hence,

$$dv(t, S_t) = r\alpha_t e^{rt}dt + \beta_t dS_t.$$

Finally, using (3.6) we obtain the desired result. △

On the other hand, applying Itô's formula, we have that

$$du(t, \tilde{B}_t) = \frac{1}{2}u''_{yy}(t, \tilde{B}_t)dt + u'_t(t, \tilde{B}_t)dt + u'_y(t, \tilde{B}_t)d\tilde{B}_t. \tag{3.8}$$

Then, putting together (3.8) and Lemma 3.5, and using the unicity of the representation of the Itô process, we obtain that

$$\begin{cases} \frac{1}{2}u''_{yy}(t,y)dt + u'_t(t,y)dt = 0, \\ u'_y(t, \tilde{B}_t) = e^{-rt}\beta(t, S_t)S_t\sigma. \end{cases}$$

We observe that the first differential equation is the heat equation with final condition at T given by (3.7), whose solution is given by:

$$u(t,y) = e^{-rT}\int_{\mathbb{R}} h(x\exp((r-\frac{\sigma^2}{2})T + \sigma(y+z)))g(T-t,z)dz,\ 0\le t<T,$$

where $g(T-t,y)$ is the Gaussian density

$$g(T-t,z) = \frac{1}{\sqrt{2\pi(T-t)}}e^{-\frac{z^2}{2(T-t)}}.$$

Therefore, we deduce that the value of a replicating portfolio for the flow $h(S_T)$ under the Black-Scholes model is given by:

$$v(t,y) = e^{-r(T-t)}\int_{\mathbb{R}} h(xy\exp((r-\frac{\sigma^2}{2})(T-t)+\sigma z))g(T-t,z)dz,\ 0<t\le T,$$

$$v(0,x) = e^{-rT}\int_{\mathbb{R}} h(x\exp((r-\frac{\sigma^2}{2})T+\sigma z))g(T,z)dz.$$

Then, writing this formulas in terms of the measure Q, we observe that if $\{\mathcal{F}_t, t\in[0,T]\}$ is the natural filtration of the Brownian motion, the price of the self-financing portfolio that replicates the payoff $h(S_T)$ is given by

$$V_t = v(t,S_t) = e^{-r(T-t)}\mathrm{E}_{\mathrm{Q}}[h(S_T)|\mathcal{F}_t], \tag{3.9}$$

and the initial price is

$$V_0 = v(0,x) = e^{-rT}\mathrm{E}_{\mathrm{Q}}[h(S_T)]. \tag{3.10}$$

Furthermore, derivating (3.9) it yields that

$$v'_y(t,y) = e^{-r(T-t)}\int_{\mathbb{R}} x\exp((r-\frac{\sigma^2}{2})(T-t)+\sigma z)h'(xy\exp((r-\frac{\sigma^2}{2})(T-t)+\sigma z))g(T-t,z)dz.$$

Therefore,

$$\beta_t = \beta(t,S_t) = \frac{e^{-r(T-t)}}{S_t}\mathrm{E}_{\mathrm{Q}}[h(S_T)S_T|\mathcal{F}_t] \tag{3.11}$$

3.4 Sensibility with respect to the parameters: the greeks

A greek is defined as the derivative of the price of an option with respect to any o its parameters of the model. Hence, the greeks measure the stability of the option under variations of the parameters.

We observe that the price of an option V_0 of strike K and maturity T depends on five parameters (x,r,σ,T,K), where x is the premium, r is the interest rates, and σ the volatility. The greeks are then the partial derivatives of V_0 with respect to these parameters.

The greeks that are most used are:

Delta: $\Delta = \frac{\partial V_0}{\partial x}$.
Gamma: $\Gamma = \frac{\partial^2 V_0}{\partial x^2}$.
Vega: $\vartheta = \frac{\partial V_0}{\partial \sigma}$.

We will next make the explicit computation of these greeks under the Black-Scholes model. For this, we will use the integration by parts formula of the Malliavin calculus. By formula (3.10), the price of an option of final flow $h(S_T)$ at time $t = 0$ under the Black-Scholes moedl is given by

$$V_0 = e^{-rT}\mathrm{E}_{\mathrm{Q}}[h(S_T)].$$

Then, if the hypotheses of Proposition 2.1 are satisfied and α is one of the parameters (x, r, σ, T, K), we have that

$$\frac{\partial V_0}{\partial \alpha} = e^{-rT}\mathrm{E}_{\mathrm{Q}}\left[h'(S_T)\frac{\partial S_T}{\partial \alpha}\right] = e^{-rT}\mathrm{E}_{\mathrm{Q}}\left[h(S_T)H(S_T, \frac{\partial S_T}{\partial \alpha})\right].$$

Assume that $h \in \mathcal{C}^1$ with bounded derivatives.

Computation of Δ: We have that

$$\Delta = \frac{\partial V_0}{\partial x} = e^{-rT}\mathrm{E}_{\mathrm{Q}}\left[h'(S_T)\frac{\partial S_T}{\partial x}\right] = \frac{e^{-rT}}{S_0}\mathrm{E}_{\mathrm{Q}}[h'(S_T)S_T].$$

We now apply Proposition 2.1 with $F = S_T$, $G = S_T$, and $u = 1$ to obtain that

$$\Delta = \frac{e^{-rT}}{S_0\sigma T}\mathrm{E}_{\mathrm{Q}}[h(S_T)B_T].$$

Computation of Γ: We write

$$\Gamma = \frac{\partial^2 V_0}{\partial x^2} = e^{-rT}\mathrm{E}_{\mathrm{Q}}\left[h''(S_T)\left(\frac{\partial S_T}{\partial x}\right)^2\right] = \frac{e^{-rT}}{S_0^2}\mathrm{E}_{\mathrm{Q}}[h''(S_T)S_T^2].$$

If we assume that $h' \in \mathcal{C}^1$ with bounded derivatives, we can appeal to Proposition 2.1 with $f = h'$, $F = S_T$, $G = S_T^2$, and $u = 1$ to obtain that

$$\Gamma = \frac{e^{-rT}}{S_0^2}\mathrm{E}_{\mathrm{Q}}[h'(S_T)S_T(\frac{B_T}{\sigma T} - 1)].$$

Finally, appealing again to Proposition 2.1 with $F = S_T$, $G = S_T(\frac{B_T}{\sigma T} - 1)$, and $u = 1$, we conclude that

$$\Gamma = \frac{e^{-rT}}{S_0^2\sigma T}\mathrm{E}_{\mathrm{Q}}[h(S_T)(\frac{B_T^2}{\sigma T} - \frac{1}{\sigma} - B_T)].$$

Computation of ϑ: We write

$$\vartheta = \frac{\partial V_0}{\partial \sigma} = e^{-rT}\mathrm{E}_{\mathrm{Q}}\left[h'(S_T)\frac{\partial S_T}{\partial \sigma}\right] = e^{-rT}\mathrm{E}_{\mathrm{Q}}[h'(S_T)S_T(B_T - \sigma T)].$$

Applying Proposition 2.1 with $F = S_T$, $G = S_T(B_T - \sigma T)$, and $u = 1$ it yields that

$$\vartheta = e^{-rT}\mathrm{E}_{\mathrm{Q}}[h(S_T)(\frac{B_T^2}{\sigma T} - \frac{1}{\sigma} - B_T)].$$

In general, h will not be derivable, and in this case, using an approximation argument one can prove that these formulas are valid for h a continuous function with jump discontinuities and linear growth. The Monte Carlo methods allows to obtain numerical simulations of these derivatives (see [KM98]).

3.5 Application of the Clark-Ocone formula in hedging

In Section 3.3 we found explicit formulas for the hedging portfolio that replicates the flow $h(S_T)$ in the Black-Scholes model.

In this section we will show how the Clark-Ocone formula can be applied to find explicit formulas for the hedging portfolio that replicates a general payoff in the Black-Scholes model. Suppose that $H \geq 0$ is an $\mathcal{F}_T$-measurable random variale such that $\mathrm{E}_Q[H^2] < \infty$. The random variable H represents the payoff of some derivative. The Itô integral representation (1.25) implies that any square integrable payoff with respect to the probability Q can be replicable, and this shows that the Black-Scholes market is complete. Indeed, by (1.25), there exists a square integrable adapted process u such that

$$e^{-rT}H = \mathrm{E}_Q[e^{-rT}H] + \int_0^T u_s dB_s.$$

Using (3.1), if we write $\tilde{V}_t = e^{-rt}V_t$ for the discounted value of the self-financing portfolio and $\tilde{S} = e^{-rt}S_t$, we get that $\tilde{V}_t - \beta_t \tilde{S}_t = \alpha_t$, and hence

$$\tilde{V}_t = x + \int_0^t \beta_s d\tilde{S}_s = x + \int_0^t \beta_s \sigma \tilde{S}_s d\tilde{B}_s.$$

Thus, it suffices to take

$$\beta_t = \frac{u_t}{\sigma \tilde{S}_t} = e^{rt}\frac{u_t}{\sigma S_t}.$$

Then, applying the Clark-Ocone formula (1.26), we get that

$$\beta_t = \frac{e^{-r(T-t)}}{\sigma S_t}\mathrm{E}_Q[D_t H|\mathcal{F}_t].$$

In the particular case where $H = h(S_T)$, then

$$\beta_t = \frac{e^{-r(T-t)}}{\sigma S_t}\mathrm{E}_Q[h'(S_T)\sigma S_T|\mathcal{F}_t],$$

which coincides with the expression obtained in (3.11).

3.6 Exercises

Exercise 3.6 *Prove Lemma 3.3 using the Gaussian calculus.*

Exercise 3.7 *Compute the price at time $t \in [0, T[$ of an European call of strike K assuming the Black-Scholes model.*

Exercise 3.8 *Compute the greeks Δ, Γ and ϑ for an European call of strike K assuming the Black-Scholes model. Compare the obtained result with the result that one obtains if we compute the greeks derivating directly the expression obtained in Exercise* 3.7.

A Appendix

Let $\{M(t), t \in [0,T]\}$ be a continuous local martingale with repect to an increasing family of σ-fileds $\{\mathcal{F}_t, t \geq 0\}$.

The following inequalities are Doob's maximal inequality and Burkhölder's inequality, respectively, which play a fundamental role in the stochastic calculus:

Proposition A.1 *For any $p > 1$, we have*

$$\mathrm{E}\left[\sup_{0\leq t\leq T} |M(t)|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathrm{E}[|M(T)|^p],$$

and this inequality also holds if we replace $|M(t)|$ by any continuous nonegative submartingale.

Proposition A.2 *For any $p > 1$, there exists constants $c_1(p)$ and $c_2(p)$ such that*

$$c_1(p)\mathrm{E}\left[\langle M(t)\rangle_T^{p/2}\right] \leq \mathrm{E}\left[\sup_{0\leq t\leq T} |M(t)|^p\right] \leq c_2(p)\mathrm{E}\left[\langle M(t)\rangle_T^{p/2}\right].$$

References

[EK03] El Karoui, N. (2003), *Couverture des risques dans les marchés financés*, Cours de DEA, http://www.cmap.polytechnique.fr/∼elkaroui/.

[KM98] Kohatsu-Higa, A. and Montero, M. (2004), Malliavin calculus in finance, *In: Handbook of computational and numerical methods in finance*, Birkhäuser Boston, 111-174.

[K06] Kuo, H. (2006), *Introduction to Stochastic Integration*, Springer.

[LL06] Lamberton, D. and Lapeyre, B. (1996), *Introduction to stochastic calculus applied to finance*, Chapman and Hall.

[M91] Malliavin, P. (1991), *Stochastic analysis*, Springer-Verlag.

[MT06] Malliavin, P. and Thalmaier (2006), *Stochastic calculus of variations in mathematical finance*, Springer-Verlag.

[N98] Nualart, D. (1998), Analysis on Wiener space and anticipating stochastic calculus, *Ecole d'Eté de Probabilités de Saint-Flour XXV, Lect. Notes in Math.* **1690**, Springer-Verlag, 123-227.

[N06] Nualart, D. (2006), *The Malliavin calculus and related topics, Second Edition*, Springer-Verlag.

[N09] Nualart, D. (2009), *Malliavin calculus and its Applications*, CBMS Notes, Americal Mathematical Society.