An Introduction to Malliavin Calculus and its applications to Finance

Vlad Bally[1], Lucia Caramellino[2], Luana Lombardi[3]

May 24, 2010

[1]Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050, Université Paris-Est Marne-la-Vallée; mailto: <bally@univ-mlv.fr>

[2]Dipartimento di Matematica, Università di Roma-Tor Vergata; mailto: <caramell @mat.uniroma2.it>.

[3]Dipartimento di Matematica, Università di L'Aquila; mailto: <lombardi.luana @gmail.com>.

Contents

Preface

From the theoretical point of view, these notes follow the ones written by Vlad Bally [1]. In addition, examples of applications of Malliavin calculus coming from Finance are developed. This has been the main contribution of Luana Lombardi, who worked on these arguments in 2008 to achieve an internship required by her PhD project.

Lucia Caramellino

Chapter 1

Abstract Integration by Parts Formula

In this chapter we introduce in an abstract way the main tool of Malliavin calculus we are going to study, that is *integration by parts formulas*, and we stress some important consequences: the use for computing sensitivities, as well as for representing the density and the conditional expectation. For the sake of simplicity, we split such an introduction in two sections, giving the one-dimensional case and the multidimensional one.

1.1 The one dimensional case

Let $(\Omega, \mathscr{F}, \mathbb{P})$ denote a probability space and let $\mathbb{E}$ stand for the expectation under $\mathbb{P}$.
The sets $C_c^k(\mathbb{R}^d)$ and $C_b^k(\mathbb{R}^d)$ denote the space of functions $f : \mathbb{R}^d \to \mathbb{R}$ which are continuously differentiable up to order k, with compact support and with bounded derivatives respectively. When the functions are infinitely differentiable, we similarly set $C_c^\infty(\mathbb{R}^d)$ and $C_b^\infty(\mathbb{R}^d)$.

Definition 1.1.1. *Let $F, G : \Omega \to \mathbb{R}$ be integrable random variables. We say that the integration by parts formula* $\mathrm{IP}(F;G)$ *holds if there exists an integrable random variable $H(F;G)$ such that*

$$\mathrm{IP}(F;G): \qquad \mathbb{E}(\phi'(F)G) = \mathbb{E}(\phi(F)H(F;G)), \quad \forall \phi \in C_c^\infty(\mathbb{R}). \tag{1.1}$$

Moreover, we say that the integration by parts formula $\mathrm{IP}_k(F;G)$ *holds if there exists an integrable random variable $H_k(F;G)$ such that*

$$\mathrm{IP}_k(F;G): \qquad \mathbb{E}(\phi^{(k)}(F)G) = \mathbb{E}(\phi(F)H_k(F;G)), \quad \forall \phi \in C_c^\infty(\mathbb{R}) \tag{1.2}$$

Remark 1.1.2. • *By using standard regularization results (e.g. by mollifiers), the test functions $C_c^\infty(\mathbb{R})$ in $\mathrm{IP}_k(F;G)$ can be replaced by $C_c^k(\mathbb{R})$ or also by $C_b^\infty(\mathbb{R})$ and $C_b^k(\mathbb{R})$.*

• *Obviously,* $\mathrm{IP}_1(F;G)$ *means* $\mathrm{IP}(F;G)$ *and* $H(F;G) = H_1(F;G)$. *Moreover, if* $\mathrm{IP}(F;G)$ *and* $\mathrm{IP}(F;H(F;G))$ *hold, then* $\mathrm{IP}_2(F;G)$ *holds with* $H_2(F;G) =$

$H(F;H(F;G))$. A similar statement holds for higher order derivatives. As an example, in $\mathrm{IP}_k(F,;1)$ this leads us to define $H_k(F,1) \equiv H_k(F)$ by recurrence:

$$H_0(F) = 1, \quad H_k(F) = H(F;H_{k-1}(F)), \quad k \geq 1.$$

• *If $\mathrm{IP}(F,G)$ holds then $\mathbb{E}(H(F,G)) = 0$: take $G = 1$ in (1.1). Moreover, the weight $H(F;G)$ in $\mathrm{IP}(F;G)$ is not unique: for any random variable R such that $\mathbb{E}(\phi(F)R) = 0$ (that is, $\mathbb{E}(R\,|\,F) = 0$ a.s.) one may use $H(F;G) + R$ as well (in fact, what is unique is $\mathbb{E}(H(F,G)\,|\,F)$). In numerical methods this plays an important role because if one wants to compute $\mathbb{E}(\phi(F)H(F;G))$ using a Monte Carlo method then one would like to work with a weight which gives minimal variance (see e.g. Fournié et al. [9]). Note also that in order to perform a Monte Carlo algorithm one has to simulate F and $H(F;G)$. In some particular cases $H(F;G)$ may be computed directly, using some methods ad hoc. But Malliavin calculus gives a systematic access to the computation of this weight. Typically, in the applications F is the solution of some stochastic equation and $H(F;G)$ appears as an aggregate of differential operators (in Malliavin's sense) acting on F. These quantities are also related to some stochastic equations and so one may use some approximations of these equations in order to produce concrete algorithms.*

Let us give a simple example. Take $F = \Delta$ and $G = g(\Delta)$ where f, g are some differentiable functions and Δ is a centered gaussian random variable of variance σ. Then

$$\mathbb{E}(f'(\Delta)g(\Delta)) = \mathbb{E}\Big(f(\Delta)\Big[g(\Delta)\frac{\Delta}{\sigma} - g'(\Delta)\Big]\Big) \tag{1.3}$$

so $\mathrm{IP}(F;G)$ holds true with $H(F;G) = g(\Delta)\frac{\Delta}{\sigma} - g'(\Delta)$. This follows from a direct application of the standard integration by parts, but in the presence of the gaussian density $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{x^2}{2\sigma})$:

$$\begin{aligned}
\mathbb{E}(f'(\Delta)g(\Delta)) &= \int f'(x)g(x)p(x)dx \\
&= -\int f(x)(g'(x)p(x) + g(x)p'(x))dx \\
&= -\int f(x)\Big[g'(x) + g(x)\frac{p'(x)}{p(x)}\Big]p(x)dx \\
&= \mathbb{E}\Big(f(\Delta)\Big[g(\Delta)\tfrac{\Delta}{\sigma} - g'(\Delta)\Big]\Big)
\end{aligned}$$

Malliavin calculus produces the weights $H(F;G)$ for a large class of random variables - (1.3) represents the simplest example of this kind - but this is not the subject of this section. Here we give some consequences of the above property.

1.1.1 The sensitivity problem

In many applications one considers quantities of the form $\mathbb{E}(\phi(F^x))$ where F^x is a family of random variables indexed on a finite dimensional parameter x. A typical example is $F^x = X_t^x$ which is a diffusion process starting from x. In order to study the sensitivity of this quantity with respect to the parameter x, one has to prove that $x \mapsto \mathbb{E}(\phi(F^x))$ is differentiable and to evaluate the derivative.

There are two ways to tackle this problem: using a pathwise approach or an approach in law.
The pathwise approach supposes that $x \mapsto F^x(\omega)$ is differentiable for almost every ω (and this is the case for $x \mapsto X_t^x(\omega)$ for example) and ϕ is differentiable also. Then $\partial_x \mathbb{E}(\phi(F^x)) = \mathbb{E}(\phi'(F^x)\partial_x F^x)$. But this approach breaks down if ϕ is not differentiable. The second approach overcomes this difficulty using the smoothness of the density of the law of F^x. So, in this approach one assumes that $F^x \sim p^x(y)dy$ and $x \mapsto p^x(y)$ is differentiable for each y. Then $\partial_x \mathbb{E}(\phi(F^x)) = \int \phi(y)\partial_x p^x(y)dy = \int \phi(y)\partial_x \ln p^x(y) p^x(y)dy = \mathbb{E}(\phi(F^x)\partial_x \ln p^x(F))$. Sometimes engineers call $\partial_x \ln p^x(F)$ the *score function*. But of course this approach works when one knows the density of the law of F^x. The integration by parts formula $\mathrm{IP}(F^x, \partial_x F^x)$ permits to write down the equality

$$\partial_x \mathbb{E}(\phi(F^x)) = \mathbb{E}(\phi'(F^x)\partial_x F^x) = \mathbb{E}(\phi(F^x)H(F^x; \partial_x F^x))$$

without having to know the density of the law of F^x. It is worth remarking that the above equality holds true even if ϕ is not derivable because there are no derivatives in the first and last term - in fact one may use some regularization arguments and then pass to the limit. Therefore the quantity of interest is the weight $H(F^x; \partial_x F^x)$. Malliavin calculus is a machinery allowing to compute such quantities for a large class of random variables for which the density of the law is not known explicitly (for example, diffusion processes). This is the approach in Fournié *et al.* [8] and [9] to the computation of Greeks (sensitivities of the price of European and American options with respect to certain parameters) in Mathematical Finance problems.

1.1.2 The density of the law

Hereafter, the notation $\mathbf{1}_A(x)$ or $\mathbf{1}_{x \in A}$ stands for the indicator function, that is $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ if $x \notin A$.

Lemma 1.1.3. *Suppose that F satisfies $\mathrm{IP}(F;1)$. Then the law of F is absolutely continuous with respect to the Lebesgue measure and the density of the law is given by*

$$p(x) = \mathbb{E}(\mathbf{1}_{[x,\infty)}(F)H(F;1)). \tag{1.4}$$

Moreover, p is continuous and $p(x) \to 0$ as $|x| \to \infty$.

Proof. The formal argument is the following: since $\delta_0(y) = \partial_y \mathbf{1}_{[0,\infty)}(y)$ one uses $\mathrm{IP}(F;1)$, so that

$$\begin{aligned}\mathbb{E}(\delta_0(F-x)) &= \mathbb{E}(\partial_y \mathbf{1}_{[0,\infty)}(F-x)) \\ &= \mathbb{E}(\mathbf{1}_{[0,\infty)}(F-x)H_1(F;1)) = \mathbb{E}(\mathbf{1}_{[x,\infty)}(F)H(F;1)).\end{aligned}$$

In order to let this reasoning rigorous, one has to regularize the Dirac function. So we take a positive function $\phi \in C_c^\infty(\mathbb{R})$ with the support equal to $[-1,1]$ and such that $\int \phi(y)dy = 1$ and for each $\delta > 0$ we define $\phi_\delta(y) = \delta^{-1}\phi(y\delta^{-1})$. Moreover we define Φ_δ to be the primitive of ϕ_δ, i.e. $\Phi_\delta(y) = \int_{-\infty}^y \phi_\delta(z)dz$, and we construct some random variables θ_δ of law $\phi_\delta(y)dy$ which are independent of F. Since θ_δ weakly converges to 0 as $\delta \to 0$, for each $f \in C_c^\infty(\mathbb{R})$ we have

$$\mathbb{E}(f(F)) = \lim_{\delta \to 0} \mathbb{E}(f(F - \theta_\delta)). \tag{1.5}$$

Setting Λ as the law of F, we can write

$$\begin{aligned}\mathbb{E}(f(F-\theta_\delta)) &= \int\int f(u-v)\phi_\delta(v)dvd\Lambda(u) = \int\int f(z)\phi_\delta(u-z)dzd\Lambda(u)\\ &= \int f(z)\mathbb{E}(\phi_\delta(F-z))dz = \int f(z)\mathbb{E}(\Phi'_\delta(F-z))dz\\ &= \int f(z)\mathbb{E}(\Phi_\delta(F-z)H(F;1))dz.\end{aligned}$$

Now, Φ_δ is uniformly bounded in δ and $\Phi_\delta(y)\to\mathbf{1}_{[x,\infty)}(y)$, as $\delta\to 0$ for a.e. y. Then using Lebesgue dominated convergence theorem we pass to the limit in the above relationship and we obtain

$$\mathbb{E}(f(F)) = \int f(z)\mathbb{E}(\mathbf{1}_{[z,\infty)}(F)H(F;1))dz$$

for any $f\in C_c^\infty(\mathbb{R})$, so that $z\mapsto\mathbb{E}(\mathbf{1}_{[z,\infty)}(F)H(F;1))$ is the probability density function of F, which is also continuous. In fact, if $z_n\to z$ one has $\mathbf{1}_{[z_n,\infty)}(F)\to\mathbf{1}_{[z,\infty)}(F)$ a.s. So, by applying the Lebesgue dominated convergence theorem, one has $p(z_n)=\mathbb{E}(\mathbf{1}_{[z_n,\infty)}(F)H(F;1))\to\mathbb{E}(\mathbf{1}_{[z,\infty)}(F)H(F;1))=p(z)$, i.e. p is a continuous function. Finally, if $z\to+\infty$ then $\mathbf{1}_{[z,\infty)}(F)\to 0$ a.s. and then $p(z)\to 0$. If instead $z\to-\infty$, one uses the same argument but to the representation

$$p(x) = -\mathbb{E}(\mathbf{1}_{(-\infty,x)}(F)H(F;1)) \qquad (1.6)$$

which follows from the fact that $\mathbf{1}_{[x,+\infty)}=1-\mathbf{1}_{(-\infty,x)}$ and by recalling that $\mathbb{E}(H(F;1))=0$ (see Remark 1.1.2). $\square$

Remark 1.1.4. [Bounds] *Suppose that $H(F;1)$ is square integrable. Then, using Chebishev's inequality*

$$p(x)\le\sqrt{\mathbb{P}(F\ge x)}\,\|H(F;1)\|_2\,.$$

In particular $\lim_{x\to\infty}p(x)=0$ and the convergence rate is controlled by the tails of the law of F. For example if F has finite moments of order p this gives $p(x)\le C\,x^{-p/2}$. In significant examples, such as diffusion processes, the tails have even exponential rate. So the problem of the upper bounds for the density is rather simple (on the contrary, the problem of lower bounds is much more challenging). The above formula gives a control for $x\to\infty$. In order to obtain similar bounds for $x\to-\infty$ one has to employ formula (1.6).

We go now further and treat the problem of the derivatives of the density function.

Lemma 1.1.5. *Suppose that $\mathrm{IP}_i(F;1), i=1,\dots,k+1$ holds true. Then the density p is k times differentiable and*

$$p^{(i)}(x) = (-1)^i\mathbb{E}(\mathbf{1}_{(x,\infty)}(F)H_{i+1}(F;1)),\quad i=0,1,\dots,k. \qquad (1.7)$$

Proof. Let $i=1$. We define $\Psi_\delta(x)=\int_{-\infty}^x\Phi_\delta(y)dy$, so that $\Psi''_\delta=\phi_\delta$, and we come back to the proof of Lemma 1.1.3. By using $\mathrm{IP}_2(F,1)$ we have

$$\mathbb{E}(\phi_\delta(F-z)) = \mathbb{E}(\Psi''_\delta) = \mathbb{E}(\Psi_\delta(F-z)H_2(F;1)),$$

so that

$$\mathbb{E}(f(F-\theta_\delta) = \int f(z)\mathbb{E}(\Psi_\delta(F-z)H_2(F;1))dz.$$

Since $\lim_{\delta\to 0}\Psi_\delta(F-z) = (F-z)_+$ we obtain

$$\mathbb{E}(f(F)) = \int f(z)\mathbb{E}((F-z)_+H_2(F;1))dz$$

and so

$$p(z) = \mathbb{E}((F-z)_+H_2(F;1)).$$

The pleasant point in this new integral representation of the density is that $z \mapsto (F-z)_+$ is differentiable. Taking derivatives in the above formula gives

$$p'(z) = -\mathbb{E}(\mathbf{1}_{[z,\infty)}(F)H_2(F;1))$$

and the proof is completed for $i=1$. In order to deal with higher order derivatives, one uses more integration by parts in order to obtain

$$p(z) = \mathbb{E}(\eta_i(F-z)H_{i+1}(F;1))$$

where η_i is an i times differentiable function such that $\eta_i^{(i)}(x) = (-1)^i\mathbf{1}_{[0,\infty)}(x)$. $\square$

Remark 1.1.6. [Bounds] *The integral representation formula (1.7) permits to obtain upper bounds of the derivatives of the density p. In particular, suppose that F has finite moments of any order and that* $\mathrm{IP}_i(F;1)$ *holds true for every* $i\in\mathbb{N}$ *and* $H_i(F;1)$ *is square integrable. Then p is infinitely differentiable and* $\left|p^{(i)}(x)\right| \leq \sqrt{\mathbb{P}(F>x)}\,\|H_i(F;1)\|_2 \leq C\,x^{-q/2}$ *for every* $q\in\mathbb{N}$. *So* $p\in\mathscr{S}$, *the Schwartz space of rapidly decreasing functions.*

[Integration by parts & densities] *Lemma 1.1.5 shows that there is an intimate relationship (quasi equivalence) between the integration by parts formula and the existence of a "good" density of the law of F. In fact, suppose that* $F\sim p(x)dx$, *where p is differentiable and* $p'(F)$ *is integrable. Then, for every* $f\in C_c^\infty(\mathbb{R})$

$$\begin{aligned}
\mathbb{E}(f'(F)) &= \int f'(x)p(x)dx = -\int f(x)p'(x)dx \\
&= -\int f(x)\frac{p'(x)}{p(x)}\mathbf{1}_{(p>0)}(x)p(x)dx \\
&= -\mathbb{E}\Big(f(F)\frac{p'(F)}{p(F)}\mathbf{1}_{(p>0)}(F)\Big).
\end{aligned}$$

So $\mathrm{IP}(F,1)$ *holds with* $H(F;1) = -\frac{p'(F)}{p(F)}\mathbf{1}_{(p>0)}(F)\in L^1$ *(because* $p'(F)\in L^1(\Omega)$*). By iteration, we obtain the following chain of implications:*

$$\mathrm{IP}_{k+1}(F,1)\ \textit{holds true}$$
$$\Downarrow$$
$$p\ \textit{is}\ k\ \textit{times differentiable and}\ p^{(k)}(F)\in L^1(\Omega)$$
$$\Downarrow$$
$$\mathrm{IP}_k(F,1)\ \textit{holds true and}\ H_k(F;1) = (-1)^k\frac{p^{(k)}(F)}{p(F)}\mathbf{1}_{(p>0)}(F)\in L^1(\Omega).$$

1.1.3 Conditional expectations

The computation of conditional expectations is crucial for numerically solving certain non linear problems coming from dynamical programming algorithms. Several authors (see Fournié *et al.* [9], Lions and Regnier [15], Bally *et al.* [3], Kohatsu-Higa and Petterson [11], Bouchard *et al.* [6]) have employed formulas based on Malliavin calculus techniques in order to compute conditional expectations. In this section we give the abstract form of this formula.

Lemma 1.1.7. *Let F and G be real random variables such that* $\mathrm{IP}(F;1)$ *and* $\mathrm{IP}(F;G)$ *hold true. Then*

$$\mathbb{E}(G \mid F = x) = \frac{\mathbb{E}(\mathbf{1}_{[x,\infty)}(F)H(F;G))}{\mathbb{E}(\mathbf{1}_{[x,\infty)}(F)H(F;1))} \tag{1.8}$$

with the convention that the term in the right hand side is null when the denominator is null.

Proof. Let $\theta(x)$ stand for the term in the right hand side of the above equality. We have to check that for every $f \in C_c^\infty(\mathbb{R})$ one has $\mathbb{E}(f(F)G) = \mathbb{E}(f(F)\theta(F))$. Using the regularizing functions from the proof of Lemma 1.1.3 we write

$$\begin{aligned}
\mathbb{E}(\theta(F)f(F)) &= \int f(z)\theta(z)p(z)dz \\
&= \int f(z)\mathbb{E}(\mathbf{1}_{[0,\infty)}(F-z)H(F;G))dz \\
&= \lim_{\delta\to 0}\int f(z)\mathbb{E}(\Phi_\delta(F-z)H(F;G))dz \\
&= \lim_{\delta\to 0}\int f(z)\mathbb{E}(G\phi_\delta(F-z))dz \\
&= \mathbb{E}\Big(G\lim_{\delta\to 0}\int f(z)\phi_\delta(F-z)dz\Big) = \mathbb{E}(Gf(F))
\end{aligned}$$

and the proof is completed. □

1.2 The multidimensional case

In this section we deal with a d dimensional random variable $F = (F^1, \dots, F^d)$. The results concerning the density of the law and the conditional expectation are quite similar. Let us introduce some notations. For $i = 1, \dots, d$, we set $\partial_i \equiv \frac{\partial}{\partial_{x^i}}$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_k) \in \{1, \dots, d\}^k$, we denote $|\alpha| = k$ and $\partial_\alpha = \partial_{\alpha_1} \cdots \partial_{\alpha_k}$ with the convention that ∂_0 is just the identity. The integration by parts formula is now the following.

Definition 1.2.1. *Let $F : \Omega \to \mathbb{R}^d$ and $G : \Omega \to \mathbb{R}$ be integrable random variables. Let $\alpha \in \{1, \dots, d\}^k$, $k \in \mathbb{N}$, be a multi-index. We say that the integration by parts formula* $\mathrm{IP}_\alpha(F;G)$ *holds if there exists an integrable random variable $H_\alpha(F;G)$ such that*

$$\mathrm{IP}_\alpha(F;G): \qquad \mathbb{E}(\partial_\alpha\phi(F)G) = \mathbb{E}(\phi(F)H_\alpha(F;G)), \quad \forall\phi \in C_c^\infty(\mathbb{R}). \tag{1.9}$$

Again, for $|\alpha| = k$, the set $C_c^\infty(\mathbb{R}^d)$ can be replaced by $C_c^k(\mathbb{R}^d)$, $C_b^\infty(\mathbb{R}^d)$ or also $C_b^k(\mathbb{R}^d)$.

Let us give a simple example which turns out to be central in Malliavin calculus. Take $F = f(\Delta^1, \ldots, \Delta^m)$ and $G = g(\Delta^1, \ldots, \Delta^m)$ where f, g are some differentiable functions and $\Delta^1, \ldots, \Delta^m$ are independent, centered gaussian random variables with variance $\sigma^1, \ldots, \sigma^m$ respectively.

We denote $\Delta = (\Delta^1, \ldots, \Delta^m)$. Then for each $i = 1, \ldots, m$

$$\mathbb{E}\Big(\frac{\partial f}{\partial x^i}(\Delta) g(\Delta)\Big) = \mathbb{E}\Big(f(\Delta)\Big[g(\Delta)\frac{\Delta^i}{\sigma^i} - \frac{\partial g}{\partial x^i}(\Delta)\Big]\Big), \tag{1.10}$$

as an immediate consequence of (1.3) and of the independence of $\Delta^1, \ldots, \Delta^m$. It then follows that $\mathrm{IP}_{\{i\}}(\Delta; g(\Delta))$ holds for every $i = 1, \ldots, d$.

We give now the result concerning the density of the law of F.

Proposition 1.2.2. *i) Suppose that* $\mathrm{IP}_{(1,2\ldots,d)}(F;1)$ *holds true. Then the density* p *of* F *exists and is given by*

$$p(x) = \mathbb{E}(\mathbf{1}_{I(x)}(F) H_{(1,2\ldots,d)}(F;1)) \tag{1.11}$$

where $I(x) = \prod_{i=1}^d [x^i, \infty)$. *In particular* p *is continuous.*
ii) Suppose that for every multi-index α, $\mathrm{IP}_\alpha(F;1)$ *holds true. Then* $\partial_\alpha p$ *exists and is given by*

$$\partial_\alpha p(x) = (-1)^{|\alpha|}\mathbb{E}(\mathbf{1}_{I(x)}(F) H_{(\alpha+1)}(F;1)) \tag{1.12}$$

where $(\alpha+1) =: (\alpha_1 + 1, \ldots, \alpha_d + 1)$. *Moreover, if* $H_\alpha(F;1) \in L^2(\Omega)$ *and* F *has finite moments of any order then* $p \in \mathscr{S}$, $\mathscr{S}$ *being the Schwartz space of the infinitely differentiable functions which decrease rapidly to infinity, together with all the derivatives.*

Proof. The formal argument for *i)* is based on $\delta_0(y) = \partial_{(1,\ldots,1)}\mathbf{1}_{I(0)}(y)$ and the integration by parts formula. In order to let it rigorous one has to regularize the Dirac function as in the proof of Lemma 1.1.3. In order to prove *ii)* one employs the same "pushing back Schwartz distribution" argument as in the proof of Lemma 1.1.5. Finally, in order to obtain bounds we write

$$|\partial_\alpha p(x)| \le \sqrt{\mathbb{P}(F^1 > x^1, \ldots, F^d > x^d)}\,\big\|H_{(\alpha+1)}(F;1)\big\|_2 .$$

If $x^1 > 0, \ldots, x^d > 0$, the Chebishev's inequality yields $|\partial_\alpha p(x)| \le C_q\,|x|^{-q}$ for every $q \in \mathbb{N}$. If the coordinates of x are not positive we have to use a variant of (1.12) which involves $(-\infty, x^i]$ instead of (x^i, ∞). $\square$

The result concerning the conditional expectation reads as follows.

Proposition 1.2.3. *Let* $F = (F^1, \ldots, F^d)$ *and* G *be two random variables such that either* $\mathrm{IP}_{(1,2\ldots,d)}(F;1)$ *and* $\mathrm{IP}_{(1,2\ldots,d)}(F;G)$ *hold true. Then*

$$\mathbb{E}(G \mid F = x) = \frac{\mathbb{E}(\mathbf{1}_{I(x)}(F) H_{(1,2\ldots,d)}(F;G))}{\mathbb{E}(\mathbf{1}_{I(x)}(F) H_{(1,2\ldots,d)}(F;1))} \tag{1.13}$$

with the convention that the term in the right hand side is null when the denominator is null.

Proof. The proof is the same as for Lemma 1.1.7, by using the regularization function $\overline{\phi}_\delta(x) = \prod_{i=1}^d \phi_\delta(x^i)$ and $\overline{\Phi}_\delta(x) = \prod_{i=1}^d \Phi_\delta(x^i)$ and the fact that $\partial^{(1,\ldots,1)}\overline{\Phi}_\delta(x) = \overline{\phi}_\delta(x)$. $\square$

Chapter 2

Brownian Malliavin calculus

2.1 The finite dimensional case

In this section we introduce the finite dimensional simple functionals and the finite dimensional simple process; we define the Malliavin derivative and the Skorohod integral for these finite dimensional objects and we derive their general important properties, as the duality formula, the chain rule, the Clark-Ocone formula and the integration by parts formula.
We will use here the space $C_p^k(\mathbb{R}^d)$ of the functions $f : \mathbb{R}^d \to \mathbb{R}$ whose derivatives up to order k exist, are continuous and with polynomial growth. Similarly we define $C_p^\infty(\mathbb{R}^d)$.

2.1.1 Main definitions and properties

Let $W = (W^1, \dots W^d)$ be a d dimensional Brownian Motion defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and we assume that the underlying filtration $\{\mathscr{F}_t\}_{t\in[0,1]}$ w.r.t. W is a Brownian motion, is the one generated by W and augmented by the $\mathbb{P}$-null sets. To simplify the notations, we suppose for the moment that $d = 1$, the multidimensional case to be deserved later in Section 2.3.
For each $n, k \in \mathbb{N}$ we denote $t_n^k = k2^{-n}$ and

$$\Delta_n^k = W(t_n^{k+1}) - W(t_n^k), \quad k = 0, \dots, 2^n - 1.$$

We denote $\Delta_n = (\Delta_n^0, \dots, \Delta_n^{2^n-1})$. Notice that Δ_n is a multidimensional Gaussian r.v., taking values in $\mathbb{R}^{2^n}$, with independent components: $\Delta_n \sim \mathcal{N}(0, 2^{-n} I_{2^n\times 2^n})$ (where $\mathcal{N}(m, \Gamma)$ denotes the Gaussian law with mean m and covariance matrix Γ and $I_{d\times d}$ the $d \times d$ identity matrix).

Definition 2.1.1. *A* ***simple functional*** *of order n is a random variable of the form $F = f(\Delta_n)$ where $f \in C_p^\infty(\mathbb{R}^{2^n})$. We denote the space S_n of the simple functionals of order n by*

$$S_n = \{F = f(\Delta_n) : f \in C_p^\infty(\mathbb{R}^{2^n})\}$$

and define the space of all simple functionals as

$$S = \bigcup_{n\in\mathbb{N}} S_n.$$

Remark 2.1.2. *1. $S_n \subset S_{n+1}$, in fact we have*

$$[t_n^k, t_n^{k+1}) = [t_{n+1}^{2k}, t_{n+1}^{2k+1}) \bigcup [t_{n+1}^{2k+1}, t_{n+1}^{2k+2}),$$

so that $F = f(\dots, \Delta_n^k, \dots) = f(\dots, \Delta_{n+1}^{2k} + \Delta_{n+1}^{2k+1}, \dots)$.

2. *$S \subset L^p(\Omega, \mathscr{F}_1, \mathbb{P})$ for all $p \geq 1$, as a consequence of the fact that f has polynomial growth and that any Gaussian r.v. has finite moment of any order.*

3. *S is a linear dense subset of $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$. There are several ways to show the validity of this assertion, we leave a possible proof in Appendix 2.6 (see next Proposition 2.6.4).*

Definition 2.1.3. *A process $U : [0,1] \times \Omega \to \mathbb{R}$ is called a* ***simple process*** *of order n if for any $k = 0, \dots, 2^n - 1$, there exists a process $U_k \in S_n$ such that*

$$U_t(\omega) = \sum_{k=0}^{2^n-1} U_k(\omega) \mathbf{1}_{[t_n^k, t_n^{k+1})}(t).$$

We denote by P_n the space of the simple processes of order n, i.e.

$$P_n = \Big\{ U : [0,1] \times \Omega \to \mathbb{R} : U_t(\omega) = \sum_{k=0}^{2^n-1} U_k(\omega) \mathbf{1}_{[t_n^k, t_n^{k+1})}(t); U_k \in S_n \Big\}$$

and the space of all simple processes is given by

$$P = \bigcup_{n \in \mathbb{N}} P_n.$$

Since $U_k \in S_n$, one has $U_k = u_k(\Delta_n^0, \dots, \Delta_n^{2^n-1})$, where $u_k \in C_p^\infty(\mathbb{R}^{2^n})$. Therefore, u_k depends on all the increments of the Brownian Motion, so that a simple process is generally not adapted. But, one has that U is adapted if and only if $U_k = u_k(\Delta_n^0, \dots, \Delta_n^{k-1})$ for any $k = 0, \dots, 2^n - 1$.

Remark 2.1.4. *1. $S_n \subset S_{n+1}$ implies that $P_n \subset P_{n+1}$.*

2. *For each fixed $\omega \in \Omega$, $t \mapsto U_t(\omega)$ is an element of $L^2([0,1], \mathscr{B}[0,1], dt)$, and in general belongs to $L^p([0,1], \mathscr{B}[0,1], dt)$ for any $p \geq 1$. Then, if $U, V \in P$ we can define the scalar product on this space by using the standard one on $L^2([0,1])$, that is*

$$\langle U, V \rangle = \int_0^1 U_s V_s ds.$$

Notice that $\langle U, V \rangle$ depends on ω and moreover, is an a.s. finite r.v.

3. *For the sake of simplicity, set $H_1 = L^2([0,1], \mathscr{B}[0,1], dt) = \{\varphi : [0,1] \to \mathbb{R}\,; \int_0^1 |\varphi_s|^2 ds < \infty\}$ and*

$$L^p(H_1) = \Big\{ U : \Omega \to H_1 : \mathbb{E}(\|U\|_{H_1}^p) = \mathbb{E}\Big(\Big[\int_0^1 |\, U_s \,|^2 \, ds \Big]^{\frac{p}{2}} \Big) < \infty \Big\}.$$

Then $P \subset L^p(H_1)$ for all $p \in \mathbb{N}$.

4. *P is a dense subset of $L^2(H_1) \equiv L^2(\Omega \times [0,1], \mathscr{F}_1 \times \mathscr{B}([0,1]), \mathbb{P} \times dt)$.*

2.1.2 Differential operators. First properties

We can now introduce the Malliavin derivative and its adjoint operator, the Skorohod integral.

Definition 2.1.5. *The **Malliavin derivative** of a r.v.* $F = f(\Delta_n) \in S_n$ *is the simple process* $\{D_t F\}_{t\in[0,1]} \in P_n$ *given by*

$$D_t F = \sum_{k=0}^{2^n-1} \frac{\partial f}{\partial x^k}(\Delta_n)\mathbf{1}_{[t_n^k, t_n^{k+1})}(t).$$

We recall that x^k represents the increment $\Delta_n^k = W_{t_n^{k+1}} - W_{t_n^k}$.
From the definition, we have that $D_t F = \frac{\partial F}{\partial \Delta_n^k}$ for $t \in [t_n^k, t_n^{k+1})$. If we denote $\Delta_n^t = \Delta_n^k$ when $t \in [t_n^k, t_n^{k+1})$, Δ_n^t represents the increment of W corresponding to t. Therefore, we can use the following notation:

$$D_t F = \frac{\partial F}{\partial \Delta_n^t}(\Delta_n) \equiv \frac{\partial f}{\partial \Delta_n^k}(\Delta_n^0, \Delta_n^1, \ldots, \Delta_n^{2^n-1}), \text{ as } t \in [t_n^k, t_n^{k+1}).$$

Notice that the definition is well posed, in the sense that the operator D does not depend on n. In fact, for $F \in S_n \subset S_{n+1}$ we have

$$\frac{\partial F}{\partial \Delta_n^k}(\Delta_n) = \frac{\partial F}{\partial \Delta_{n+1}^{2k}}(\Delta_{n+1}) = \frac{\partial F}{\partial \Delta_{n+1}^{2k+1}}(\Delta_{n+1}), \tag{2.1}$$

because $t \in [t_n^k, t_n^{k+1}) = [t_{n+1}^{2k}, t_{n+1}^{2k+1}) \cup [t_{n+1}^{2k+1}, t_{n+1}^{2k+2})$ and $F = f(\ldots, \Delta_n^k, \ldots) = f(\ldots, \Delta_{n+1}^{2k} + \Delta_{n+1}^{2k+1}, \ldots)$. Therefore, (2.1) allows to define

$$D : S = \bigcup_n S_n \to P = \bigcup_n P_n$$

as follows:

$$D_t F = \frac{\partial F}{\partial \Delta_n^t}(\Delta_n), \quad \text{as } t \in [0,1].$$

Definition 2.1.6. *The **Skorohod integral** is defined as the operator*

$$\delta : P \to S, \quad \delta(U) = \sum_{k=0}^{2^n-1} \left(u_k(\Delta_n)\Delta_n^k - \frac{\partial u_k}{\partial x^k}(\Delta_n)\frac{1}{2^n} \right)$$

where $U = \sum_{k=0}^{2^n-1} u_k(\Delta_n)\mathbf{1}_{[t_n^k, t_n^{k+1})} \in P_n \subset P$.

Note that the definition again does not depend on n and so it is correct.

Remark 2.1.7. ***(Skorohod integral vs Ito integral)*** *We have already noticed that a process* $U \in P_n$ *is* $\mathscr{F}_t$*-adapted if and only if* $u_k(\Delta_n)$ *does depend only on the variables* $\Delta_n^1, \ldots, \Delta_n^{k-1}$. *Consequently,* $\frac{\partial u_k}{\partial x^k} = 0$ *and in such a case,*

$$\delta(U) = \sum_{k=0}^{2^n-1} u_k(\Delta_n)\Delta_n^k = \int_0^1 U_s dW_s,$$

that is, $\delta(U)$ *coincide with the Ito integral w.r.t.* W. *This shows that the Skorohod integral aims to be an extension of the Ito integral over the set of non adapted processes.*

We can now prove the link between Malliavin derivatives and Skorohod integrals and investigate some immediate properties of such operators.

Proposition 2.1.8. *(i)* **[Duality]** *For any $F \in S$ and $U \in P$ one has*

$$\mathbb{E}(\langle DF, U\rangle) = \mathbb{E}(F\delta(U)).$$

(ii) **[Chain rule]** *Let $F = (F^1, \dots F^m)$ where $F^i \in S, i = 1, \dots m$ and $\Phi \in C_p^\infty(\mathbb{R}^m)$. Then $\Phi(F) \in S$ and*

$$D\Phi(F) = \sum_{i=1}^m \partial_{x^i}\Phi(F) DF^i.$$

(iii) **[Skorohod integral of a special product]** *Let $U \in P$ and $F \in S$. Then*

$$\delta(FU) = F\delta(U) - \langle DF, U\rangle.$$

Proof. *(i)* Let n denote an integer such that $F \in S_n$ and $U \in P_n$. Then,

$$\mathbb{E}(\langle DF, U\rangle) = \mathbb{E}\Big(\sum_{k=0}^{2^n-1} \frac{\partial f}{\partial x^k}(\Delta_n) u_k(\Delta_n) \times \frac{1}{2^n}\Big).$$

Δ_n is a vector of i.i.d. Gaussian r.v.'s with variance $h_n = 1/2^n$. Then, we can use (1.10) and we obtain

$$\mathbb{E}\Big(\frac{\partial f}{\partial x^k}(\Delta_n) u_k(\Delta_n)\Big) = \mathbb{E}\Big(f(\Delta_n)\Big[u_k(\Delta_n)\frac{\Delta_n^k}{h_n} - \frac{\partial u_k}{\partial x^k}(\Delta_n)\Big]\Big).$$

By replacing everything we obtain

$$\mathbb{E}(\langle DF, U\rangle) = \mathbb{E}\Big(f(\Delta_n)\sum_{k=0}^{2^n-1}\Big[u_k(\Delta_n)\Delta_n^k - \frac{\partial u_k}{\partial x^k}(\Delta_n)\frac{1}{2^n}\Big]\Big) = \mathbb{E}(F\delta(U)).$$

The proof of *(ii)* is straightforward.

(iii) Take $G \in S$. By using the duality formula and the chain rule, we have

$$\begin{aligned}
\mathbb{E}[G\delta(FU)] &= \mathbb{E}[\langle DG, FU\rangle] = \mathbb{E}[\langle FDG, U\rangle] \\
&= \mathbb{E}[\langle D(GF) - GDF, U\rangle] = \mathbb{E}[\langle D(GF), U\rangle] - \mathbb{E}[G\langle DF, U\rangle] \\
&= \mathbb{E}[GF\delta(U)] - \mathbb{E}[\langle DF, U\rangle].
\end{aligned}$$

Then, $\mathbb{E}[G\delta(FU)] = \mathbb{E}[G(F\delta(U) - \langle DF, U\rangle)]$ for any $G \in S$, and *(iii)* immediately follows. □

We are now ready to prove a first integration by parts formula in the Malliavin sense. For $F = (F^1, \dots, F^m)$, with $F^i \in S$ for any $i = 1, \dots, m$, set σ_F as the following $m \times m$ symmetric matrix:

$$\sigma_F^{ij} = \langle DF^i, DF^j\rangle = \int_0^1 D_tF^i\, D_tF^j\, dt, \quad i, j = 1, \dots, m.$$

σ_F is called the **Malliavin covariance matrix associated to** F. It is a positive definite matrix, because for any $\xi \in \mathbb{R}^m$ one has

$$\langle \sigma_F\xi, \xi\rangle = \sum_{i,j=1}^m \sigma_F^{ij}\xi^i\xi^j = \int_0^1 \sum_{i,j=1}^m D_tF^i\xi^i\, D_tF^j\xi^j dt = \int_0^1 \Big|\sum_{i=1}^m D_tF^i\xi^i\Big|^2 dt \geq 0.$$

Proposition 2.1.9. [MIbP formula] *Let $F = (F^1, \dots, F^m)$ and G be such that $F^1, \dots, F^m, G \in S$. Suppose that σ_F is invertible and let γ_F denote the inverse of σ_F. Suppose moreover that $\det \gamma_F \in S$. Then for every $\phi \in C_b^1(\mathbb{R}^m)$*

$$\mathbb{E}\Big(\frac{\partial \phi}{\partial x^i}(F)G\Big) = \mathbb{E}(\phi(F)H_i(F;G))$$

with

$$H_i(F;G) = \delta\Big(\sum_{j=1}^m \gamma_F^{ji} G DF^j\Big)$$

Proof. By using the chain rule, we can write

$$\langle D\phi(F), DF^j\rangle = \sum_{q=1}^m \frac{\partial \phi}{\partial x^q}(F)\langle DF^q, DF^j\rangle = \sum_{q=1}^m \frac{\partial \phi}{\partial x^q}(F)\sigma_F^{qj}, \quad j = 1, \dots, m.$$

Since σ_F is invertible with inverse matrix γ_F, we can write

$$\frac{\partial \phi}{\partial x^i}(F) = \sum_{j=1}^m \langle D\phi(F), DF^j\rangle \gamma_F^{ji}, \quad i = 1, \dots, m.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\Big(\frac{\partial \phi}{\partial x^i}(F)G\Big) &= \mathbb{E}\Big(\sum_{j=1}^m \langle D\phi(F), DF^j\rangle \gamma_F^{ji} G\Big) \\
&= \mathbb{E}\Big(\langle D\phi(F), \sum_{j=1}^m DF^j \gamma_F^{ji} G\rangle\Big) \\
&= \mathbb{E}\Big(\phi(F)\delta\Big(\sum_{j=1}^m DF^j \gamma_F^{ji} G\Big)\Big)
\end{aligned}$$

and the above steps make sense because all the r.v.'s and processes involved are, by hypothesis, in the right spaces. $\square$

2.2 The infinite dimensional case

The duality formula is the one to be used in order to show that the operators D and δ are closable and this last property allows one to extend them to the infinite dimensional case, that is for r.v.'s and processes non necessarily depending on the increments of the Brownian motion but depending on the whole path.
Let us start from the following facts. We have seen that

$$D : S \subset L^2(\Omega) \to P \subset L^2(H_1) \quad \text{and} \quad \delta : P \subset L^2(H_1) \to S \subset L^2(\Omega).$$

The operators δ and D are linear but unbounded, i.e. it does not exist a constant C such that for any $F \in S$ one has

$$\|DF\|^2_{L^2(H_1)} = \mathbb{E}\Big(\int_0^1 |D_s F|^2 ds\Big) \le C\|F\|^2_{L^2(\Omega)}.$$

Anyway, we can state the following property:

Lemma 2.2.1. *D and δ are both closable, that is*

i) *if $\{F_n\}_n \subset S$ is such that $\lim_n F_n = 0$ in $L^2(\Omega)$ and $\lim_n DF_n = U$ in $L^2(H_1)$ then $U = 0$;*

ii) *if $\{U_n\}_n \subset P$ is such that $\lim_n U_n = 0$ in $L^2(H_1)$ and $\lim_n \delta(U_n) = F$ in $L^2(\Omega)$ then $F = 0$.*

Proof. i) Take $\{F_n\}_n \subset S$ such that $\lim_n F_n = 0$ in $L^2(\Omega)$ and $\lim_n DF_n = U$ in $L^2(H_1)$. Since P is dense in $L^2(H_1)$, it is sufficient to prove that $\mathbb{E}(\langle U, V\rangle) = 0$ for any $V \in P$. In fact, if $V \in P$, by using the duality formula one has

$$\mathbb{E}(\langle U, V\rangle) = \lim_n \mathbb{E}(\langle DF_n, V\rangle) = \lim_n \mathbb{E}(F_n \delta(V)) = 0$$

The proof of **ii)** is similar. □

2.2.1 The set $Dom_p(D) = \mathbb{D}^{1,p}$

We first introduce a suitable set on which the Malliavin derivative D is well defined and then, extending the set S of the simple functionals.

Definition 2.2.2. *Let $p \in \mathbb{N}$. We say that $F \in Dom_p(D) = \mathbb{D}^{1,p}$ if there exists a sequence $\{F_n\}_n \subset S$ such that*

$$\lim_n F_n = F \text{ in } L^p(\Omega) \text{ and } \lim_n DF_n = U \text{ in } L^p(H_1) \text{ for some } U \in L^p(H_1).$$

In this case we define $DF = U = \lim_n DF_n$ in $L^p(H_1)$.

Since $\|\cdot\|_{p'} \le \|\cdot\|_p$ and $\|\cdot\|_{L^{p'}(H_1)} \le \|\cdot\|_{L^p(H_1)}$ for $p \ge p'$, we have $\mathbb{D}^{1,p} \subset \mathbb{D}^{1,p'}$. We put

$$\mathbb{D}^{1,\infty} = Dom_\infty D = \bigcap_{p\in\mathbb{N}} \mathbb{D}^{1,p}.$$

We observe that $\mathbb{D}^{1,2}$ does not depend on the sequence $F_n, n \in \mathbb{N}$ because D is closable, but is not an algebra. We note that $\mathbb{D}^{1,\infty}$ is an algebra and the definition of DF does not depend on p.
We define a norm $\|\cdot\|_{1,p}$ on $\mathbb{D}^{1,p}$ by

$$\|F\|_{1,p}^p = \|F\|_p^p + \|DF\|_{L^p(H_1)}^p \equiv \mathbb{E}(|F|^p) + \mathbb{E}\Big(\Big(\int_0^1 |D_t F|^2 dt\Big)^{p/2}\Big).$$

Notice that for $p = 2$, the norm $\|\cdot\|_{1,2}$ is the one resulting from the scalar product

$$\langle F, G\rangle_{1,2} = \mathbb{E}(FG) + \mathbb{E}\Big(\int_0^1 D_s F\, D_s G ds\Big).$$

Moreover, $\mathbb{D}^{1,2}$ is a Hilbert space.

Remark 2.2.3.
- *$F \in \overline{S}^{\|\cdot\|_{1,p}}$ if there exists $F_n \in S, n \in \mathbb{N}$ such that $F_n \to F$ in $L^p(\Omega)$ and $(F_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\|\cdot\|_{1,p}$;*
- *it then follows that $\mathbb{D}^{1,p} \equiv Dom_p(D) = \overline{S}^{\|\cdot\|_{1,p}}$;*

- *$Dom_p(D)$ is complete, i.e. every Cauchy sequence in $Dom_p(D)$ converges to an element of $Dom_p(D)$. Indeed consider a Cauchy sequence $(F_n)_{n\in\mathbb{N}}$ with respect to $\|\cdot\|_{1,p}$. This sequence is also a Cauchy one with respect to $\|\cdot\|_p$ and we know that L^p is complete, so there exists $F \in L^p(\Omega)$ such that $F_n \to F$ in $\|\cdot\|_p$. Since $F_n \in Dom_p(D)$ we may find a sequence of a simple functionals $F_n^{'}$ s.t. $\|F_n - F_n^{'}\|_{1,p} \le \frac{1}{n}$ so that $(F_n^{'})_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_{1,p}$ and $F_n^{'} \to F$ in $\|\cdot\|_p$. So $F \in Dom_p(D)$.*

2.2.2 The set $Dom_p(\delta)$

Again, we introduce a suitable set on which the Skorohod integral δ is well defined and then, extending the set P of the simple processes. We start similarly to Definition 2.2.2.

Definition 2.2.4. *Let $p \in \mathbb{N}$. We say that $U \in Dom_p(\delta)$ if there exists a sequence $U_n \in P, n \in \mathbb{N}$ such that*

$$\lim_n U_n = U \text{ in } L^p(H_1) \text{ and } \lim_n \delta(U_n) = F \text{ in } L^p(\Omega) \text{ for some } F \in L^p(\Omega).$$

In this case we define $\delta(U) = F = \lim_n \delta(U_n)$ in $L^p(\Omega)$.

On P, we consider the norm

$$\|U\|_{\delta,p} = \|U\|_{L^p(H_1)} + \|\delta(U)\|_p$$

and we have

$$Dom_p(\delta) = \overline{P}^{\|\cdot\|_{\delta,p}}.$$

2.2.3 Properties

Sometimes it is unpleasant to compute Malliavin derivatives or Skorohod integrals through limits. We necessarily need a criterion, for example as follows

Proposition 2.2.5. [Criterion]

(i) Let $F \in L^2(\Omega)$. Suppose that there exists a sequence $F_n \in \mathbb{D}^{1,2}$ s.t.

i $\lim_n F_n = F$ *in* $L^2(\Omega)$

ii $\sup_n \|F_n\|_{1,2} \le C < \infty$.

Then $F \in Dom_2(D)$ and $\|F\|_{1,2} \le C$. Moreover, if $\sup_n \|F_n\|_{1,p} \le C_p$ then $\|F\|_{1,p} \le C_p$.

(ii) Let $U \in L^2(H_1)$. Suppose that there exists a sequence $U_n \in Dom_2(\delta)$ s.t.

i $\lim_n U_n = U$ *in* $L^2(H_d)$

ii $\sup_n \|U_n\|_{\delta,2} \le C < \infty$.

Then $U \in Dom_2(\delta)$ and $\|U\|_{\delta,2} \le C$. Moreover if $\sup_n \|U_n\|_{\delta,p} \le C_p$ then $\|U\|_{\delta,p} \le C_p$.

Proof. (i) Any bounded set in a Hilbert space is relatively compact, so we may find $F^{'} \in \mathbb{D}^{1,2}$ s.t. $F_n \to F^{'}$ weakly. We use Mazur's lemma[1]:for each $n \in \mathbb{N}$ there exists k_n and $\lambda_k^n \geq 0, k = n, \ldots, k_n$, s.t. $\sum_{k=n}^{k_n} \lambda_k^n = 1$ and $\overline{F}_n := \sum_{k=n}^{k_n} \lambda_k^n F_k \to F^{'}$ strongly with respect to $\| \cdot \|_{1,2}$ and, in particular in $L^2(\Omega)$. Notice that

$$\|F - \overline{F}_n\|_2 = \| \sum_{k=n}^{k_n} \lambda_k^n (F - F_k)\|_2 \leq \sum_{k=n}^{k_n} \lambda_k \|F - F_k\|_2 \leq \sup_{k \geq n} \|F - F_k\|_2 \longrightarrow 0.$$

It follow that $F^{'} = F$ and so $F \in \mathbb{D}^{1,2}$. We also have

$$\|F\|_{1,2} = \lim_n \|\overline{F}_n\|_{1,2} \leq \lim_n \sum_{k=n}^{k_n} \lambda_k^n \|F_n\|_{1,2} \leq C.$$

Let us now prove the assertion concerning the p-norm. Passing to a subsequence we may assume that $|\overline{F}_n - F| \to 0$ a.s. Since $\sup_n \|\overline{F}_n\|_{1,p} \leq C_p$ we may use uniformly integrability in order to derive $\overline{F}_n \to F$ with respect to $\| \cdot \|_{1,p'}$ for $p' < p$. Then $\|F\|_{1,p'} \leq \sup_n \|\overline{F}_n\|_{1,p'} \leq \sup_n \|\overline{F}_n\|_{1,p} \leq C_p$. And finally, $\|F\|_{1,p} \leq \sup_{p'<p} \|\overline{F}\|_{1,p'} \leq C_p$.
Similar arguments give (ii). $\square$

We have seen in the finite dimensional framework that the Malliavin integration by parts formula can be achieved once some properties are verified, in particular the duality relationship, the chain rule and, for practical purposes, the Skorohod integral of a special product. In other words, if Proposition 2.1.8 continues to hold. The answer is positive, and in fact one has

Proposition 2.2.6. (i) **[Duality]** *For $F \in Dom_2(D)$ and $U \in Dom_2(\delta)$,*

$$\mathbb{E}(\langle DF, U\rangle) = \mathbb{E}(F\delta(U)).$$

(ii) **[Chain rule]** *Let $F = (F^1, \ldots F^m)$ where $F^i \in \mathbb{D}^{1,2}, i = 1, \ldots m$ and $\Phi \in C_b^1(\mathbb{R}^m)$. Then $\Phi(F) \in \mathbb{D}^{1,2}$ and*

$$D\Phi(F) = \sum_{i=1}^{m} \partial_{x^i} \Phi(F) DF^i.$$

If $F^i \in \mathbb{D}^{1,\infty}$ then the conclusion is true for $\phi \in C_p^1(\mathbb{R}^m)$.

(iii) **[Skorohod integral of a special product]** *Let $u \in Dom_2(\delta)$ and $F \in \mathbb{D}^{1,2}$ such that $Fu \in Dom_2(\delta)$. Then*

$$\delta(FU) = F\delta(U) - \langle DF, U\rangle.$$

[1]**Mazur's lemma.** Let $(X, \|\cdot\|)$ denote a Banach space and $\{u_n\}_n \subset X$ such that $u_n \to u$ weakly (that is, $f(u_n) \to f(u)$ for each continuous linear functional f). Then there exists a function $N : \mathbb{N} \to \mathbb{N}$ and for any $n \in \mathbb{N}$ some numbers $\{\alpha(n)_k\, ;\, k = 1, \ldots, N(n)\}$ such that $\alpha(n)_k > 0$ for any $k = 1, \ldots, N(n)$, $\sum_{k=1}^{N(n)} \alpha(n)_k = 1$ and such that the convex combination $v_n = \sum_{k=1}^{N(n)} \alpha(n)_k u_k$ strongly converges to u_0, i.e. $\|v_n - u_0\| \to 0$ as $n \to \infty$.

Proof. (i) For $F \in Dom_2(D)$ and $U \in Dom_2(\delta)$, take $\{F_n\}_n \subset S$ and $\{U_n\}_n \subset P$ such that, as $n \to \infty$,

$$F_n \to F, \delta(U_n) \to \delta(U) \text{ in } L^2(\Omega) \text{ and } DF_n \to DF, U_n \to U \text{ in } L^2(H_1).$$

By applying the duality relationship between S and P (Proposition 2.1.8),

$$\mathbb{E}(\langle DF, U\rangle) = \lim_{n\to\infty} \mathbb{E}(\langle DF_n, U_n\rangle) = \lim_{n\to\infty} \mathbb{E}(F_n\delta(U_n)) = \mathbb{E}(F\delta(U)).$$

(ii) Let us first prove that if $F^k \in S$ for any $k = 1, \ldots, m$ and $\Phi \in C_b^1(\mathbb{R}^m)$ then $\Phi(F) \in \mathbb{D}^{1,2}$ and the chain rule holds. In fact, let $\{\Phi_n\}_n \subset C_b^\infty(\mathbb{R}^m) \subset C_p^\infty(\mathbb{R}^m)$ denote a sequence such that $\|\Phi_n - \Phi\|_\infty \to 0$ and $\|\nabla\Phi_n - \nabla\Phi\|_\infty \to 0$ as $n \to \infty$. Since $\Phi_n(F) \in S$, the chain rule holds by Proposition 2.1.8. Now, $\|\Phi_n(F) - \Phi(F)\|_2 \le \|\Phi_n - \Phi\|_\infty \to 0$ and for each k one has

$$\left\|\partial_{x^k}\Phi_n(F)\, DF^k - \partial_{x^k}\Phi(F)\, DF_k\right\|_{L^2(H_1)} \le \|\nabla\Phi_n - \nabla\Phi\|_\infty \|D^kF\|_{L^2(H_1)} \to 0$$

and this gives the statement.

Suppose now that $F^k \in \mathbb{D}^{1,2}$ for any $k = 1, \ldots, m$ and $\Phi \in C_b^1(\mathbb{R}^m)$. We then take $\{F_n^k\}_n \subset S$ such that $\|F_n^k - F^k\|_{1,2} \to 0$. Since Φ has bounded derivatives we immediately obtain $\|\Phi(F_n) - \Phi(F)\|_2 \to 0$. Moreover, from the first part of the proof we know that $D\Phi(F_n) = \sum_{k=1}^m \partial_{x^k}\Phi(F_n)DF_n^k$. Then, we have to prove that for each k,

$$\|\partial_{x^k}\Phi(F_n)DF_n^k - \partial_{x^k}\Phi(F)DF^k\|_{L^2(H_1)} \to 0.$$

We can write

$$\|\partial_{x^k}\Phi(F_n)DF_n^k - \partial_{x^k}\Phi(F)DF^k\|_{L^2(H_1)} \le a_n + b_n$$

where

$$\begin{aligned} a_n &= \left\|\partial_{x^k}\Phi(F_n)\big(DF_n^k - DF^k\big)\right\|_{L^2(H_1)} \\ b_n &= \left\|\big(\partial_{x^k}\Phi(F_n) - \partial_{x^k}\Phi(F)\big)DF^k\right\|_{L^2(H_1)} \end{aligned}$$

Concerning a_n, since $\partial_{x^k}\Phi$ is bounded, one has

$$a_n \le const\|DF_n^k - DF^k\|_{L^2(H_1)} \to 0.$$

As for b_n, first notice that

$$b_n^2 = \mathbb{E}\Big(|\partial_{x^k}\Phi(F_n) - \partial_{x^k}\Phi(F)|^2 \int_0^1 |D_tF^k|^2 dt\Big)$$

Now, if we pass to any subsequence s.t. $F_n \to F$ a.s. and use Lebesgue's theorem, we immediately obtain $b_n^2 = \mathbb{E}(|\partial_{x^k}\Phi(F_n) - \partial_{x^k}\Phi(F)|^2 \int_0^1 |D_sF^k|^2) \to 0$.

(iii) Let $G \in S$. Using the duality formula we can write

$$\begin{aligned} \mathbb{E}[G\delta(FU)] &= \mathbb{E}[\textstyle\int_0^1 D_sG \times F \times U_s ds] \\ &= \mathbb{E}[\textstyle\int_0^1 (D_s(FG) - GD_sF) \times U_s ds] \\ &= \mathbb{E}[GF\delta(U)] - \mathbb{E}[G\textstyle\int_0^1 D_sF \times U_s ds]. \end{aligned}$$

This relation is true for all $G \in S$, so we have the thesis. $\square$

Remark 2.2.7. *Notice that if $F^i \in \mathbb{D}^{1,\infty}$, $i = 1, \ldots, m$, then we can use Holder's inequality (in particular, to show that $b_n \to 0$ as $n \to \infty$ in the above proof of (ii) in Proposition 2.2.6) and then we get that the chain rule holds also for $\Phi \in C^1_p(\mathbb{R}^m)$.*

Actually, the chain rule holds also in other situation, for example under the requirement that Φ is only Lipschitz continuous (see e.g. Nualart [18], Proposition 1.2.3, p. 30).

Example 2.2.8. Let $F \in \mathbb{D}^{1,\infty}$ be such that $e^F \in L^p$ for any p. Then $e^F \in \mathbb{D}^{1,2}$ and

$$De^F = e^F DF.$$

In fact, let $\{\psi_n\}_{n\geq 1} \subset C^\infty_c(\mathbb{R})$ be a sequence such that $\psi_n(x) = 1$ if $|x| \leq n$, $\psi_n(x) = 0$ if $|x| > n+1$, $0 \leq \psi_n \leq 1$ for any x and $\sup_n \sup_x |\psi'_n(x)| < \infty$. Set now $G_n = \psi_n(F)e^F$. Notice that $G_n = \Psi_n(F)$ with $\Psi_n(x) = \psi_n(x)e^x \in C^\infty_c(\mathbb{R})$, so that $G_n \in \mathbb{D}^{1,2}$ and the chain rule holds:

$$DG_n = \Psi'_n(F)DF = e^F DF\big(\psi'_n(F) + \psi_n(F)\big).$$

Then, it is sufficient to prove that $G_n \to e^F$ in $L^2(\Omega)$ and $DG_n \to e^F DF$ in $L^2(H_1)$. In fact, we have

$$\|G_n - e^F\|_2^2 = \mathbb{E}\big(e^{2F}|\psi_n(F) - 1|^2\big).$$

But, $e^{2F}|\psi_n(F) - 1|^2 \to 0$ a.s. and $e^{2F}|\psi_n(F) - 1|^2 \leq 2e^{2F} \in L^1$, so that by Lebesgue's dominated convergence theorem one has $\|G_n - e^F\|_2^2 \to 0$. As for the second statement, by Hölder's inequality we have

$$\begin{aligned}\|G_n - e^F DF\|^2_{L^2(H_1)} &= \mathbb{E}\Big(\int_0^1 e^{2F}|D_sF|^2\,|\psi'_n(F) + \psi_n(F) - 1|^2\,ds\Big)\\ &\leq \mathbb{E}\big(e^{2pF}\,|\psi'_n(F) + \psi_n(F) - 1|^{2p}\,ds\big)^{1/p}\|DF\|^2_{L^{2q}(H_1)}\end{aligned}$$

where $p, q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$. By using arguments similar to the ones developed above, one has $\mathbb{E}\big(e^{2pF}\,|\psi'_n(F) + \psi_n(F) - 1|^{2p}\,ds\big) \to 0$, and the statement holds.

2.2.4 Examples

We give here some leading examples.

Example 2.2.9. [Brownian motion] Take $F = W_t$, as $t \in [0,1]$. Then $F \in Dom_2(D)$ and

$$D_sW_t = \mathbf{1}_{s\leq t}.$$

In fact, we can write ($\lfloor\cdot\rfloor$ denoting the integer part)

$$W_t = \sum_{k=0}^{\lfloor 2^n t\rfloor}(W_{t_n^{k+1}} - W_{t_n^k}) + W_t - W_{\frac{\lfloor 2^n t\rfloor}{2^n}}.$$

Now, since

i $F_n := \sum_{k=0}^{\lfloor 2^n t\rfloor}(W_{t_n^{k+1}} - W_{t_n^k}) \to W_t$ in $L^2(\Omega)$ as $n \to \infty$,

ii $F_n \in \mathbb{D}^{1,2}$ and $D_s F_n = \mathbf{1}_{s \leq \frac{\lfloor 2^n t \rfloor}{2^n}} \to \mathbf{1}_{s \leq t} = U$ in $L^2(H_1)$ as $n \to \infty$,

it immediately follows that $D_s W_t$ exists and is equal to $\mathbf{1}_{s \leq t}$.

Example 2.2.10. [Ito integral of square integrable functions] Let $\phi \in L^2([0,1])$ and set $W(\phi) := \int_0^1 \phi_r dW_r$. Then, $W(\phi) \in \mathbb{D}^{1,2}$ and

$$D_s W(\phi) = \phi(s).$$

The proof is a consequence of the following steps.

step 1 Let ϕ be a step function on the dyadic intervals, i.e.

$$\phi(s) = \sum_{k=0}^{2^n - 1} \phi_k \mathbf{1}_{[t_n^k, t_n^{k+1})}(s).$$

Then $W(\phi) = \sum_{k=0}^{2^n-1} \phi_k \Delta_n^k$ is a simple functional and we compute directly the derivative: $D_s W(\phi) = \phi_{k(s)} = \phi(s)$.

step 2 Let $\phi \in L^2(0,1)$ be a continuous function. Then, there exists a sequence $\{\phi_n\}_n$ of step functions such that $\phi_n \to \phi$ in $L^2(0,1)$ as $n \to \infty$. Now, step 1 ensures us that

$$D_s W(\phi_n) = \phi_n(s)$$

Since $\phi_n \to \phi$ in $L^2(0,1)$, the statement immediately follows.

step 3 The generalization to general functions ϕ belonging to $L^2(0,1)$ follows from the fact that the set of the continuous functions on $(0,1)$ is a dense subset of $L^2(0,1)$.

Example 2.2.11. For $\phi_\ell \in L^2(0,1)$, $\ell = 1, \ldots, m$, and for $\Phi \in C_p^1(\mathbb{R}^m)$, set

$$F = \Phi\Big(\int_0^1 \phi_1(s) dW_s, \ldots, \int_0^1 \phi_m(s) dW_s \Big).$$

Then $F \in \mathbb{D}^{1,2}$ and

$$D_s F = \sum_{k=1}^m \partial_{x^k} \Phi\Big(\int_0^1 \phi_1(s) dW_s, \ldots, \int_0^1 \phi_m(s) dW_s \Big) \phi_k(s).$$

The proof is an immediate consequence of Example 2.2.10 and the chain rule.

Remark 2.2.12. *Example 2.2.11 is particularly important if one is interested in studying the link with the definition of Malliavin derivatives as done in many texts, as for example the widely well-known one by Nualart [18]. There, the set of simple processes $\mathscr{S}$ is given by the random variables F of the form*

$$F = f\Big(\int_0^1 \phi_1(s) dW_s, \ldots, \int_0^1 \phi_n(s) dW_s \Big)$$

where $n \in \mathbb{N}$, $f \in C_c^\infty(\mathbb{R}^n)$ and $\phi_i \in H_1 = L^2([0,1]\mathscr{B}([0,1]), dt)$ Then, for F as above, the Malliavin derivative is defined as

$$D_t F = \sum_{k=1}^n \partial_{x^k} f\Big(\int_0^1 \phi_1(s) dW_s, \ldots, \int_0^1 \phi^n(s) dW_s \Big) \phi_k(t). \tag{2.2}$$

Furthermore, on $\mathscr{S}$ one sets

$$\|F\|_{1,2}^2 = \|F\|_{L^2(\Omega)}^2 + \|DF\|_{L^2(\Omega\times[0,1])}^2$$

and defines $\mathbb{D}^{1,2} = \overline{\mathscr{S}}^{\|\cdot\|_{1,2}}$. Now, Exercise 2.3.7 allows one to prove that this definition of Malliavin derivative agrees with the one already presented in these notes.

Remark 2.2.13. *Consider a smooth functional of the form*

$$F = f(W_{t_1}, \dots, W_{t_n})$$

with $f \in C_p^\infty$ and $0 < t_1 < \cdots < t_n \leq 1$, so that

$$D_t F = \sum_{i=1}^n \partial_{x^i} f(W_{t_1}, \dots, W_{t_n}) \mathbf{1}_{t \leq t_i}$$

Then, for $h \in H_1 = L^2([0,1], \mathscr{B}([0,1]), dt)$ one has

$$\begin{aligned}
\langle DF, h\rangle &= \int_0^1 \sum_{i=1}^n \partial_{x^i} f(W_{t_1}, \dots, W_{t_n}) \mathbf{1}_{t \leq t_i} h_t dt \\
&= \sum_{i=1}^n \partial_{x^i} f(W_{t_1}, \dots, W_{t_n}) \int_0^{t_i} h_t dt \\
&= \lim_{\varepsilon \to 0} \frac{f(W_{t_1} + \varepsilon \int_0^{t_1} h_t dt, \dots, W_{t_n} + \varepsilon \int_0^{t_1} h_t dt) - f(W_{t_1}, \dots, W_{t_n})}{\varepsilon}
\end{aligned}$$

Therefore, for any $h \in H_1$ one gets

$$\langle DF, h\rangle = \frac{d}{d\varepsilon} F\Big(\omega + \varepsilon \int_0^{\cdot} h(t) dt\Big)\Big|_{\varepsilon=0}$$

that is, for such F's the Malliavin derivative DF is linked to the directional derivative of F in the directions of the Cameron Martin space $\mathscr{H}_1 = \{\varphi \in C([0,1], \mathbb{R}) : \varphi_t = \int_0^t h_s ds, \text{ for } h \in L^2([0,1])\}$.

Example 2.2.14. [Lebesgue and Ito integrals] Let U denote an adapted process such that $\mathbb{E}(\int_0^1 |U_r|^2 dr) < \infty$. Set

$$I_0(U) = \int_0^1 U_r dr \text{ and } I_1(U) = \int_0^1 U_r dW_r.$$

We assume that for each fixed $r \in [0,1]$, $U_r \in \mathbb{D}^{1,2}$ and

i $\sup_{r \leq 1} \|U_r\|_{1,2} < \infty$;
ii setting $\tau_n(r) = \lfloor r2^n \rfloor / 2^n$ and $U_r^n = U_{\tau_n(r)}$, then

$$\int_0^1 \|U_r - U_r^n\|_{1,2}^2 dr = \int_0^1 \mathbb{E}\Big(|U_r - U_{\tau_n(r)}|^2 + \int_0^1 |D_s U_r - D_s U_{\tau_n(r)}|^2 ds\Big) dr \to 0$$

as $n \to \infty$.

Then, $I_i(U) \in \mathbb{D}^{1,2}$ for $i = 0, 1$ and one has:

$$D_s I_0(U) = D_s \int_0^1 U_r dr = \int_s^1 D_s U_r dr \tag{2.3}$$

and

$$D_s I_1(U) = D_s \int_0^1 U_r dW_r = U_s + \int_s^1 D_s U_r dW_r \tag{2.4}$$

In fact, suppose first $i = 1$. Then,

$$I_1(U^n) = \sum_{k=0}^{2^n-1} U_{k/2^n} \Delta_n^k.$$

Therefore,

$$D_s I_1(U^n) = \sum_{k=0}^{2^n-1} D_s \Big(U_{k/2^n} \Delta_n^k \Big) = U_{\lfloor 2^n s \rfloor / 2^n} + \sum_{k=\lfloor 2^n s \rfloor}^{2^n-1} D_s U_{k/2^n} \Delta_n^k$$

and notice that

$$D_s I_1(U^n) \to U_s + \int_s^1 D_s U_r dW_r \text{ in } L^2(\Omega) \text{ as } n \to \infty$$

because of **i**. Now, by **ii**, we have $I_1(U^n) \to I_1(U)$ in $L^2(\Omega)$. Using **i**, we obtain $\sup_n \|I_1(U^n)\|_{1,2} < \infty$. Then we can use the criterion in Proposition 2.2.5 in order to get $I_1(U) \in \mathbb{D}^{1,2}$. Now, since we know that $I_1(U) \in \mathbb{D}^{1,2}$, we have $DI_1(U) = \lim_{n\to\infty} DI_1(U^n)$ in $L^2(\Omega)$, and (2.4) is proved. Concerning (2.3), one can proceed in a similar way.

Example 2.2.15. We show here the Malliavin differentiability of the maximum of a Brownian motion. Let us put $M = \sup_{s\leq 1} W_s$ (we test the time interval $[0,1]$ but nothing changes for more general intervals) and we show that $D_t M = I_{[0,\tau]}(t)$, where τ is the a.s. unique point at which W attains its maximum.
For any $n \in \mathbb{N}$, we put $M_n = \max_{k=0,\dots,2^n} W_{k/2^n}$. Notice that $M_n \to M$ a.s. and[2] $|M_n - M|^2 \leq 4M^2 \in L^1(\Omega)$, so that by the Lebesgue dominated convergence theorem one has $M_n \to M$ in $L^2(\Omega)$. Thus, it remains to show that $M_n \in \mathbb{D}^{1,2}$ and $D_t M_n \to 1_{[0,\tau]}(t)$ in $L^2([0,1] \times \Omega)$.
By setting $\phi_n : \mathbb{R}^{2^n+1} \to \mathbb{R}$, $\phi_n(x) = \max(x^0, \dots, x^{2^n})$, then obviously $M_n = \phi_n(W_0, W_{1/2^n}, \dots, W_1)$. The function ϕ_n is not a C_p^1 function, so the chain rule in Proposition 2.2.6 cannot be immediately applied. However, ϕ_n is a Lipschitz continuous function and its partial derivatives exist a.e., so smoothing arguments allow to state the validity of the chain rule (see e.g. Nualart [18], Proposition 1.2.3, p. 30): $M_n = \phi_n(W_0, W_{1/2^n}, \dots, W_1) \in \mathbb{D}^{1,2}$ and

$$\begin{aligned} D_t M_n &= \sum_{k=0}^{2^n} \frac{\partial \phi_n}{\partial x^k}(W_0, W_{1/2^n}, \dots, W_1)\, D_t W_{k/2^n} \\ &= \sum_{k=0}^{2^n} \frac{\partial \phi_n}{\partial x^k}(W_0, W_{1/2^n}, \dots, W_1)\, 1_{t<k/2^n}. \end{aligned}$$

[2]Recall the reflecting principle for a Brownian motion: for any $x > 0$, one has $\mathbb{P}(\sup_{t\leq T} W_t > x) = 2\mathbb{P}(W_T > x)$. For $T = 1$ one gets $\mathbb{P}(M > x) = 2\mathbb{P}(W_1 > x)$ and then M has a probability density function given by $f_M(x) = \sqrt{2/\pi}\, \exp(-x^2/2)\, \mathbf{1}_{x>0}$, which tells us that $M \in L^p$ for any p.

We set $A_0 = \{\phi_n(x) = x^0\}$ and, as $k = 1, \dots, 2^n$, $A_k = \{\phi_n(x) \neq x^0, \dots, \phi_n(x) \neq x^{k-1}, \phi_n(x) = x^k\}$. Then, $\partial_{x^k}\phi_n(x) = 1_{A_k}(x)$ a.e., so that we can write

$$D_t M_n = \sum_{k=0}^{2^n} \mathbf{1}_{(W_0, W_{1/2^n}, \dots, W_1) \in A_k} 1_{t<k/2^n} = \mathbf{1}_{[0,\tau_n]}(t)$$

where τ_n denotes the a.s. unique point among the $k/2^n$'s such that $M_n = W_{\tau_n}$. Straightforward computations allow to see that

$$\mathbb{E}\Big(\int_0^1 |D_t M_n - \mathbf{1}_{[0,\tau]}(t)|^2\,dt\Big) = \mathbb{E}(|\tau_n - \tau|).$$

Now, $\tau_n \to \tau$ a.s. because W has continuous paths - notice that this proves the a.s. uniqueness of τ - and $|\tau_n - \tau| \leq 2$, so $\mathbb{E}(|\tau_n - \tau|) \to 0$, which in turn implies that $D_t M_n \to 1_{[0,\tau]}(t)$ in $L^2([0,1] \times \Omega)$. Then,

$$D_t M = \mathbf{1}_{[0,\tau]}(t).$$

Example 2.2.16. We compute here the Skorohod integral of the Brownian bridge process on $[0,1]$, which corresponds in some sense to a Brownian motion forced to be in two fixed points x and y at time 0 and 1 respectively. There are several ways to introduce such a process; for example, the Brownian bridge can be seen as

$$u(t) = x + t(y-x) + W_t - tW_1,$$

where B is a one dimensional Brownian motion. Then, by recalling that Skorohod and Ito integrals coincide on adapted processes, one has

$$\delta(u) = xW_1 + (y-x)\int_0^1 t\,dW_t + \int_0^1 W_t\,dW_t - \delta(v\,W_1),$$

where $v(t) = t$. By using (iii) of Proposition 2.2.6, $\delta(v\,W_1) = W_1\int_0^1 t\,dW_t - \int_0^1 D_t W_1\,t\,dt = W_1\int_0^1 t\,dW_t - \frac{1}{2}$. Moreover, by Ito's formula applied to $f(W_t) = W_t^2$ and to $g(t, W_t) = tW_t$ one gets $\int_0^1 W_t\,dW_t = \frac{1}{2}(W_1^2 - 1)$ and $\int_0^1 t\,dW_t = W_1 - \int_0^1 W_t\,dt$ respectively. Then

$$\delta(u) = yW_1 + (W_1 + x - y)\int_0^1 W_t\,dt - \frac{1}{2}W_1^2.$$

2.2.5 The Clark-Ocone formula

We recall the martingale representation formula: if $F \in L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ then there exists a real valued and $\mathscr{F}_t$-adapted process $\phi \in L^2(\Omega \times [0,1], \mathscr{F}_1 \times \mathscr{B}([0,1]), \mathbb{P} \times dt)$ such that $F = \mathbb{E}(F) + \int_0^1 \phi_s dW_s$. When the random variable F is Malliavin derivable, one can write down explicitly the process ϕ. In fact, one has

Proposition 2.2.17. [Clark-Ocone formula] *If $F \in \mathbb{D}^{1,2}$ then*

$$F = \mathbb{E}(F) + \int_0^1 \mathbb{E}(D_t F \,|\, \mathscr{F}_t) dW_t.$$

Proof. Without loss of generality we can assume that $\mathbb{E}(F) = 0$ (otherwise, we work with $F - \mathbb{E}(F)$), so that by the Brownian martingale representation theorem one has $F = \int_0^1 \phi_s dW_s$ for some $\mathscr{F}_t$-adapted process in $L^2(\Omega \times [0,1])$. Let us set $P_{\rm ad}$ the subset of the simple processes P which are $\mathscr{F}_t$-adapted. For $U \in P_{\rm ad}$ one has $\delta(U) = \int_0^1 U_s dW_s$, so that

$$\mathbb{E}(F\delta(U)) = \mathbb{E}\Big(\int_0^1 \phi_s dW_s \int_0^1 U_s dW_s\Big) = \mathbb{E}\Big(\int_0^1 \phi_s U_s ds\Big).$$

On the other hand, by the duality one has

$$\begin{aligned} \mathbb{E}(F\delta(U)) &= \mathbb{E}(\langle DF, U\rangle) = \mathbb{E}\Big(\int_0^1 D_sFU_s ds\Big) \\ &= \mathbb{E}\Big(\int_0^1 \mathbb{E}(D_sF \,|\, \mathscr{F}_s)U_s ds\Big). \end{aligned}$$

It then follows that

$$\langle U, \phi - \mathbb{E}(D.F \,|\, \mathscr{F}.)\rangle_{L^2(\Omega\times[0,1])} = \mathbb{E}\Big(\int_0^1 U_s\Big(\phi_s - \mathbb{E}(D_sF \,|\, \mathscr{F}_s)\Big)ds\Big) = 0$$

for any $U \in P_{\rm ad}$. The statement now follows by noticing that the closure of $P_{\rm ad}$ w.r.t. the norm in $L^2(\Omega \times [0,1])$ is given by all the $\mathscr{F}_t$-adapted processes belonging to $L^2(\Omega \times [0,1])$. $\square$

Corollary 2.2.18. *1. If $F \in \mathbb{D}^{1,2}$ then F is a.s. constant if and only if $DF = 0$.*
2. If $A \in \mathscr{F}_1$ then $\mathbf{1}_A \in \mathbb{D}^{1,2}$ if and only if either $\mathbb{P}(A) = 1$ or $\mathbb{P}(A) = 0$. As a consequence, $\mathbb{D}^{1,2}$ is strictly included in $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$.

Proof. The proof of 1. is immediate from the Clark-Ocone formula. As for 2., if $\mathbf{1}_A \in \mathbb{D}^{1,2}$ then by the chain rule we get $D\mathbf{1}_A = D(\mathbf{1}_A^2) = 2\mathbf{1}_A D\mathbf{1}_A$. Now, if $D\mathbf{1}_A \neq 0$ then $1 = 2\mathbf{1}_A$ which is impossibile. Then, $D\mathbf{1}_A = 0$, that is $\mathbf{1}_A = const$ which is true if either $\mathbb{P}(A) = 1$ or $\mathbb{P}(A) = 0$. The converse is immediate. $\square$

As an example, tale $A = \{W_t > 0\}$ and $F = \mathbf{1}_A$. Then $F \in L^2(\Omega, \mathscr{F}_1, \mathbb{P})$, because $\mathbb{E}(F^2) = \mathbb{P}(W_t > 0) = 1/2$ while $\mathbf{1}_A \notin \mathbb{D}^{1,2}$, so that $\mathbb{D}^{1,2}$ is actually strictly included in $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$.

2.2.6 The set $Dom_p(L)$

We introduce here the Ornstein-Uhlembeck operator L. On the class of simple functionals S one has

$$L : S \to S, \quad LF = -\delta(DF).$$

The following duality relationship holds:

$$\mathbb{E}(FLG) = -\mathbb{E}(\langle DF, DG\rangle) = \mathbb{E}(LFG).$$

Similar arguments give that L is closable, so that one can give the following

Definition 2.2.19. *$F \in Dom(L) \equiv Dom_2(L)$ if there exists a sequence of simple functionals $\{F_n\}_n$ such that $F_n \to F$ in $L^2(\Omega)$ and $LF_n \to G$ in $L^2(\Omega)$, for some $G \in L^2(\Omega)$. We then we define $LF := G = \lim_n LF_n$. If the above convergence holds in $L^p(R)$, $p \geq 2$ we say that $F \in Dom_p(L)$. We put $Dom_\infty(L) = \cap_{p\geq 2} Dom_p(L)$*

Obviously, for $F \in Dom(L)$ one again has $LF = -\delta(DF)$. Moreover, on S we may define the norm

$$\|F\|_{L,p} = \|F\|_p + \|LF\|_p$$

so that $Dom_p(L) = \overline{S}^{\|\cdot\|_{L,p}}$. The following chain rule holds:

Proposition 2.2.20. *Let $F = (F^1, \dots F^m)$ where $F^i \in Dom_\infty(L), i = 1, \dots m$ and $\Phi \in C_p^\infty(\mathbb{R}^m)$. Then $\Phi(F) \in Dom_\infty L$ and*

$$L\Phi(F) = \sum_{i=1}^m \partial_{x^i}\Phi(F) LF^i + \sum_{i,j=1}^m \partial_{x^i}\partial_{x^j}\Phi(F)\langle DF^i, DF^j\rangle.$$

The proof is left as an exercise.

Remark 2.2.21. *Consider $m \geq 1$ paths $\phi^1, \dots, \phi^m$ in H_1 and set $F^i = W(\phi^i)$ $= \int_0^1 \phi_s^i dW_s$. Such r.v.'s play a crucial role in Malliavin calculus (see also next Appendix 2.6) and in this special context, they allow to give a rough interpretation of the denomination "Ornstein-Uhlembeck operator" given to $L = -\delta(D)$. But for a deeper motivation, we refer to the interesting initial part of the book of Sanz-Solé [19].*
Set $a^{ij} = \langle \phi^i, \phi^j\rangle = \int_0^1 \phi_s^i \phi_s^j\, ds$ and notice that this is a symmetric, non negative definite $m \times m$ matrix, so that it has a square root σ (that is, σ is a $m \times m$ matrix such that $\sigma\sigma^* = a$). Now, for $F^i = W(\phi^i)$ one has $DF^i = \phi^i$. Therefore, $LF^i = -\delta(\phi^i) = -W(\phi^i) = -F^i$ and $\langle DF^i, DF^j\rangle = \langle \phi^i, \phi^j\rangle = a^{ij}$. Then for any $f \in C_p^\infty(\mathbb{R}^m)$, Proposition 2.2.20 gives

$$Lf(F) = -\sum_{i=1}^m F^i \partial_{x^i} f(F) + \sum_{i,j=1}^m a^{ij}\partial^2_{x^i x^j} f(F)$$

Now, the analogous operator on $\mathbb{R}^m$, that is

$$\mathscr{L} f(x) = -\sum_{i=1}^m x^i \partial_{x^i} f(x) + \sum_{i,j=1}^m a^{ij}\partial^2_{x^i x^j} f(x)$$

is the infinitesimal generator of the diffusion process X on $\mathbb{R}^m$ evolving as

$$dX_t = -\, X_t dt + \sqrt{2}\,\sigma\, dW_t$$

which is an Ornstein-Uhlembeck process.

2.2.7 The integration by parts formula

An important consequence of the duality formula is the integration by parts formula.

Definition 2.2.22. *Let $F = (F^1, \dots, F^m)$ with $F^i \in \mathbb{D}^{1,2}$. The* ***Malliavin covariance matrix of*** *F is defined as the symmetric positive definite matrix given by*

$$\sigma_F^{ij} = \langle DF^i, DF^j \rangle = \int_0^1 D_s F^i \, D_s F^j ds.$$

We introduce the non-degeneracy assumption:

$$\text{(N-D)} \quad \mathbb{E}((\det \sigma_F)^{-p}) < \infty, \forall p \in \mathbb{N}. \tag{2.5}$$

If (N-D) is holds then σ_F is almost surely invertible and we denote $\gamma_F = \sigma_F^{-1}$. The integration by parts formula reads as follows:

Theorem 2.2.23. [MIbP formula] *Let $F = (F^1, \dots, F^m)$ with $F^i \in \mathbb{D}^{1,\infty}$ and $G \in \mathbb{D}^{1,\infty}$. Suppose also that $\sigma_F^{i,j} \in \mathbb{D}^{1,\infty}$, (N-D) holds for F and that $DF_i \in \bigcap_{p\in\mathbb{N}} Dom_p(\delta)$, $i = 1, \dots m$. Then for every $\phi \in C_p^1(\mathbb{R}^m)$ we have*

$$\mathbb{E}(\partial_i \phi(F)G) = \mathbb{E}(\phi(F)H_i(F,G)), \quad i = 1, \dots, m \tag{2.6}$$

where

$$H_i(F,G) = \sum_{j=1}^m \delta(G\gamma_F^{ij} DF^j) = -\sum_{j=1}^m \Big(G\gamma_F^{i,j} LF^j + \langle D(G\gamma_F^{i,j}), DF^j \rangle \Big). \tag{2.7}$$

Proof. First, let us notice that the second equality in (2.7) follows from the Skorohod integral of a special product property (see (*iii*) of Proposition 2.2.6). Using the chain rule we can write that

$$D_s \phi(F) = \nabla \phi(F) D_s F.$$

Then,

$$\langle D\phi(F), DF^i \rangle_{H_1} = \Big(\sigma_F \nabla \phi(F) \Big)^i,$$

which yields

$$\partial_i \phi(F) = \langle D\phi(F), (\gamma_F DF)^i \rangle.$$

By using the duality formula, one gets

$$\mathbb{E}(\partial_i \phi(F)G) = \mathbb{E}(\langle D\phi(F), G(\gamma_F DF)^i \rangle) = \mathbb{E}(\phi(F)\delta(G(\gamma_F DF)^i)$$

and the statement holds. $\square$

2.3 Multidimensional Brownian motion

In this section we deal with a d-dimensional Brownian motion $W = (W^1, \dots W^d)$ defined on a complete probability space $(\Omega, \mathcal{F}, P)$, where $\mathcal{F} = \{\mathcal{F}_t\}_{t\in[0,1]}$ is the one generated by W and augmented by the $\mathbb{P}$-null sets.
The definitions of Malliavin derivative and Skorohod integral, as well as the resulting properties, can be extended as in the standard calculus. It is easy to describe what are the main ideas. For example, we have seen that the Malliavin derivative is given by

$$D_t F = \frac{\partial F}{\partial \Delta W_t}$$

where the above derivative has to be intended "in some sense". Now, since we have now a d-dimensional Brownian motion, and then d independent Brownian motions, such derivative becomes now "a gradient" since in principle it can be done w.r.t. all the d directions:

$$D_tF = (D^1_tF, \dots, D^d_tF), \quad D^i_tF = \frac{\partial F}{\partial \Delta W^i_t}, \quad i = 1, \dots, d.$$

Now, concerning the Skorohod integral, it will be again the adjoint operator. Since the principal tool is the duality relationship, that is $\mathbb{E}(\langle DF, U\rangle) = \mathbb{E}(F\delta(U))$, it is clear that the domain of the operator δ is necessarily based on processes taking values on $\mathbb{R}^d$. And moreover, for adapted processes the Skorohod and the Ito integral will agree: for an adapted process $U_t = (U^1_t, \dots, U^d_t)$ with the usual properties giving the Ito integrability,

$$\delta(U) = \int_0^1 \sum_{i=1}^d U^i_t dW^i_t.$$

But let us start by introducing the notations.
For $n, k \in \mathbb{N}$, we denote $t^k_n = k2^{-n}$ and

$$\Delta^{k,i}_n = W^i(t^{k+1}_n) - W^i(t^k_n), \quad k = 0, \dots, 2^n - 1 \text{ and } i = 1, \dots d.$$

We set now

$$\Delta^k_n = (\Delta^{k,1}_n, \dots, \Delta^{k,d}_n)^*, \quad k = 0, \dots, 2^n - 1.$$

(the symbol * denoting the transpose). Let us recall that, as i, k vary, the r.v.'s $\Delta^{k,i}_n$ are i.i.d. and $\Delta^{k,i}_n \sim \mathcal{N}(0, \frac{1}{2^n})$. Therefore, $\Delta_n = (\Delta^0_n, \dots, \Delta^{2^n-1}_n) \in \mathbb{R}^{d\times 2^n}$ is a $d \times 2^n$ matrix.

Now, a **simple functional of order** n is a random variable of the form $F = f(\Delta_n)$ where $f \in C^\infty_p(\mathbb{R}^{d\times 2^n})$. The space of the simple functionals of order n is

$$S_n = \{F = f(\Delta_n) : f \in C^\infty_p(\mathbb{R}^{d\times 2^n})\}.$$

We set $S = \bigcup_n S_n$ as the set of all the simple functionals.
A process $U : [0, 1] \times \Omega \to \mathbb{R}^d$ is called a **simple process of order** n if $U_t = (U^1_t, \dots, U^d_t)$ with

$$U^i_t(\omega) = \sum_{k=0}^{2^n-1} U^i_k \mathbf{1}_{[t^k_n, t^{k+1}_n)}(t), \quad U^i_k \in S_n, k = 0, \dots, 2^n - 1, i = 1, \dots, d.$$

It is worth to notice that U_t is a r.v. taking values on $\mathbb{R}^d$. Recall that the requirement $U^i_k \in S_n$ allows to write the ith component U^i of a simple process of order n as

$$U^i_t(\omega) = \sum_{k=0}^{2^n-1} u^i_k(\Delta_n)\mathbf{1}_{[t^k_n, t^{k+1}_n)}(t), \quad u^i_k \in C^\infty_p(\mathbb{R}^{d\times 2^n}), k = 0, \dots, 2^n - 1.$$

as $i = 1, \dots, d$. Again, a simple process of order n is adapted if and only if

$$u^i_k(\Delta_n) \equiv u^i_k(\Delta^0_n, \dots, \Delta^{2^n-1}_n) = u^i_k(\Delta^0_n, \dots, \Delta^{k-1}_n)$$

for any k and i.
We set P_n^d as the set of the simple processes of order n and $P^d = \bigcup_n P_n^d$ as the set of all the simple processes.
For each fixed $\omega \in \Omega$, $t \mapsto U_t$ is an element of $L^2([0,1], \mathcal{B}[0,1], dt, \mathbb{R}^d) = \{\varphi : [0,1] \to \mathbb{R}^d \,:\, \varphi \text{ is Borel measurable and } \int_0^1 |\varphi(s)|^2 ds < \infty\} := H_d$. Then, on P^d we can define the scalar product by using the usual one on L^2: for $U, V \in P^d$,

$$\langle U, V\rangle = \int_0^1 \sum_{i=1}^d U_s^i \times V_s^i ds.$$

Notice the resulting value is a r.v.
Now, let us denote

$$L^p(H_d) = \Big\{ U : \Omega \to H_d \,:\, \mathbb{E}[\|U\|_{H_d}^p] = \mathbb{E}\Big(\Big(\int_0^1 \sum_{i=1}^d \mid U_s^i \mid^2 ds\Big)^{\frac{p}{2}}\Big) < \infty \Big\}.$$

Then, $P^d \subset L^p(H_d)$ for all $p \in \mathbb{N}$.

Definition 2.3.1. *The* ***Malliavin derivative*** *of a variable $F = f(\Delta_n) \in S_n$ is the simple process $\{D_t F\}_{t\in[0,1]} \in P_n^d$ given by*

$$D_t F = (D_t^1 F, \dots D_t^d F),$$

where

$$D_t^i F = \sum_{k=0}^{2^n - 1} \frac{\partial f}{\partial x^{k,i}}(\Delta_n) \mathbf{1}_{[t_n^k, t_n^{k+1})}(t), \quad i = 1, \dots, d.$$

Notice that D_t^i is the Malliavin derivative described in the previous section if one considers the Brownian motion W^i. In some sense, in order to define D_t^i one has to freeze all the random sources expect for the ith one. That is why D_t^i is often called as the Malliavin derivative in the ith direction of the Brownian motion.

Definition 2.3.2. *The* ***Skorohod integral*** *is defined as the operator*

$$\delta \,:\, P^d \to S, \quad \delta(U) = \sum_{i=1}^d \delta^i(U^i)$$

where, as $i = 1, \dots, d$, for $U_t^i = \sum_{k=0}^{2^n-1} u_k^i(\Delta_n) \mathbf{1}_{[t_n^k, t_n^{k+1})}(t)$,

$$\delta^i(U^i) = \sum_{k=0}^{2^n-1} \Big(u_k^i(\Delta_n)\Delta_n^{k,i} - \frac{\partial u_k^i}{\partial x^{k,i}}(\Delta_n)\frac{1}{2^n}\Big).$$

Again, $\delta^i(U^i)$ agrees with the one-dimensional definition of the Skorohod integral: simply, work on the ith Brownian motion W^i, or equivalently, on the ith direction of the Brownian motion W.
Notice also that whenever U is adapted, $\frac{\partial u_k}{\partial x^{k,i}} = 0$, for any i, so that

$$\delta(U) = \sum_{i=1}^d \sum_{k=0}^{2^n-1} u_k(\Delta_n)\Delta_n^{k,i} = \int_0^1 \sum_{i=1}^d U_s^i dW_s^i,$$

that is the Skorohod integral coincides with the Ito one.

Similarly to what developed in Section 2.1.2, one has the same result as in Proposition 2.1.8, i.e.

Proposition 2.3.3. *(i)* **[Duality]** *For any $F \in S$ and $U \in P$,*

$$\mathbb{E}(\langle DF, U \rangle) = \mathbb{E}(F\delta(U))$$

(ii) **[Chain rule]** *Let $F = (F^1, \dots F^m)$ where $F^i \in S, i = 1, \dots m$ and $\Phi \in C_b^1(\mathbb{R}^m)$. Then $\Phi(F) \in S$ and*

$$D^i \Phi(F) = \sum_{\ell=1}^m \partial_{x^\ell} \Phi(F) D^i F^\ell, \quad i = 1, \dots, d.$$

(iii) **[Skorohod integral of a special product]** *For $U \in P^d$ and $F \in S$,*

$$\delta(FU) = F\delta(U) - \langle DF, U \rangle.$$

The proofs are identical to the ones of Proposition 2.1.8. In particular, the duality relationship allow to extend the operators in the infinite dimensional case. In fact, by developing the same arguments as in Section 2.2, one can immediately prove that the operators D and δ are closable. Then,

$$D : \mathbb{D}^{1,2} \subset L^2(\Omega) \to L^2(H_d) \quad \text{and} \quad \delta : Dom_2(\delta) \subset L^2(H_d) \to L^2(\Omega).$$

All properties in Proposition 2.3.3 can be extended and read as follows.

Proposition 2.3.4. *(i)* **[Duality]** *For any $F \in \mathbb{D}^{1,2}$ and $U \in Dom_2(\delta)$,*

$$\mathbb{E}(\langle DF, U \rangle) = \mathbb{E}(F\delta(U))$$

(ii) **[Chain rule]** *Let $F = (F^1, \dots F^m)$ where $F^i \in \mathbb{D}^{1,2}, i = 1, \dots m$ and $\Phi \in C_b^1(\mathbb{R}^m)$. Then $\Phi(F) \in S$ and*

$$D^i \Phi(F) = \sum_{\ell=1}^m \partial_{x^\ell} \Phi(F) D^i F^\ell, \quad i = 1, \dots, d.$$

(iii) **[Skorohod integral of a special product]** *For $U \in Dom_2(\delta)$ and $F \in \mathbb{D}^{1,2}$ such that $FU \in Dom_2(\delta)$,*

$$\delta(FU) = F\delta(U) - \langle DF, U \rangle.$$

Again, the proof follows by density arguments similar to the ones developed in Proposition 2.2.6.

Concerning the examples discussed in Section 2.2.4, let us see what happens in the multidimensional case (the proofs are similar, so we omit them).

Example 2.3.5. [Brownian motion - see Example 2.2.9] Take $F = W_t^i$, as $t \in [0, 1]$. Then $F \in Dom_2(D)$ and

$$D_s^j W_t^i = \mathbf{1}_{i=j} \mathbf{1}_{s \le t}.$$

Example 2.3.6. [Ito integral of square integrable functions - see Example 2.2.10] Let $\phi \in L^2([0,1])$ and set $W^j(\phi) := \int_0^1 \phi_r dW_r^j$. Then, $W^j(\phi) \in \mathbb{D}^{1,2}$ and

$$D_s^i W^j(\phi) = \begin{cases} \phi(s) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Example 2.3.7. [See Example 2.2.11] For $\phi_\ell^j \in L^2(0,1)$, $\ell = 1, \ldots, m$ and $j = 1, \ldots, d$, and for $\Phi \in C_p^1(\mathbb{R}^m)$, set

$$F = \Phi\Big(\sum_{j=1}^d \int_0^1 \phi_1^j(s) dW_s^j, \ldots, \sum_{j=1}^d \int_0^1 \phi_m^j(s) dW_s^j\Big).$$

Then $F \in \mathbb{D}^{1,2}$ and

$$D_s^i F = \sum_{k=1}^m \partial_{x^k} \Phi\Big(\sum_{j=1}^d \int_0^1 \phi_1^j(s) dW_s^j, \ldots, \sum_{j=1}^d \int_0^1 \phi_m^j(s) dW_s^j\Big)\, \phi_k^i(s).$$

Example 2.3.8. [Ito integrals - see Example 2.2.14] Let U denote an adapted process such that $\mathbb{E}(\int_0^1 |U_r|^2 dr) < \infty$. Set

$$I_0(U) = \int_0^1 U_r dr \text{ and for } i = 1, \ldots, d,\ I_i(U) = \int_0^1 U_r dW_r^i.$$

We assume that for each fixed $r \in [0,1]$, $U_r \in \mathbb{D}^{1,2}$ and

i $\sup_{r \le 1} \|U_r\|_{1,2} < \infty$;
ii setting $\tau_n(r) = \lfloor r2^n \rfloor / 2^n$ and $U_r^n = U_{\tau_n(r)}$, then

$$\int_0^1 \|U_r - U_r^n\|_{1,2}^2 dr = \int_0^1 \mathbb{E}\Big(|U_r - U_{\tau_n(r)}|^2 + \int_0^1 |D_s U_r - D_s U_{\tau_n(r)}|^2 ds\Big) dr \to 0$$

as $n \to \infty$.

Then, $I_i(U) \in \mathbb{D}^{1,2}$ for any $i = 0, 1, \ldots, d$ and one has:

$$D_s^j I_0(U) = D_s^j \int_0^1 U_r dr = \int_s^1 D_s^j U_r dr \tag{2.8}$$

and as $i = 1, \ldots, d$,

$$D_s^j I_i(U) = D_s^j \int_0^1 U_r dW_r^i = \begin{cases} U_s + \displaystyle\int_s^1 D_s^i U_r dW_r^i & \text{if } i = j \\ \displaystyle\int_s^1 D_s^j U_r dW_r^i & \text{if } i \neq j \end{cases} \tag{2.9}$$

As for the Ornstein-Uhlembeck operator L, on the class of simple functionals S one has

$$L : S \to S, \quad LF = -\delta(DF) = -\sum_{i=1}^d \delta_i(D^i F),$$

so that

$$\mathbb{E}(FLG) = -\mathbb{E}(\langle DF, DG \rangle) = \mathbb{E}(LFG).$$

Then, one proves that L is closable, so that $Dom_p L = \overline{S}^{\|\cdot\|_{L,p}}$, where for $F \in S$,

$$\|F\|_{L,p} = \|F\|_p + \|LF\|_p.$$

Again, $Dom_\infty L = \cap_{p\geq 2} Dom_p L$ and the chain rule holds for L, that is

Proposition 2.3.9. *Let $F = (F^1, \dots F^m)$ where $F^i \in Dom_\infty L, i = 1, \dots m$ and $\Phi \in C_p^\infty(\mathbb{R}^m)$. Then $\Phi(F) \in Dom_\infty L$ and*

$$L\Phi(F) = \sum_{i=1}^m \partial_{x^i}\Phi(F) LF^i + \sum_{i,j=1}^m \partial_{x^i}\partial_{x^j}\Phi(F)\langle DF^i, DF^j\rangle.$$

As for the Clark-Ocone formula, one gets the same results, that is

Proposition 2.3.10. *1.* **[Clark-Ocone formula]** *If $F \in \mathbb{D}^{1,2}$ then*

$$F = \mathbb{E}(F) + \sum_{k=1}^d \int_0^1 \mathbb{E}(D_t^k F \mid \mathscr{F}_t) dW_t^k.$$

2. If $F \in \mathbb{D}^{1,2}$ then F is a.s. constant if and only if $DF = 0$.
3. If $A \in \mathscr{F}_1$ then $\mathbf{1}_A \in \mathbb{D}^{1,2}$ if and only if either $\mathbb{P}(A) = 1$ or $\mathbb{P}(A) = 0$.

Let us now discuss the MIbP formula. Let us start by introducing the Malliavin covariance matrix and the non-degeneracy assumption (N-D) in the multidimensional case.

Definition 2.3.11. *Let $F = (F^1, \dots, F^m)$ with $F^\ell \in \mathbb{D}^{1,2}$. The* ***Malliavin covariance matrix*** *is the symmetric positive definite matrix defined by*[3]

$$\sigma_F^{\ell j} = \langle DF^\ell, DF^j\rangle = \int_0^1 \sum_{i=1}^d D_s^i F^\ell \times D_s^i F^j ds.$$

We say that σ_F fulfils the non degeneracy assumption if

$$\mathbb{E}((\det \sigma_F)^{-p}) < \infty, \forall p \in \mathbb{N}. \tag{2.10}$$

If (2.10) is true then σ_F is almost surely invertible. We denote $\gamma_F = \sigma_F^{-1}$. Then,

Theorem 2.3.12. [MIbP formula] *Let $F = (F^1, \dots, F^m)$ with $F^\ell \in \mathbb{D}^{1,\infty}$ and $G \in \mathbb{D}^{1,\infty}$. Suppose also that $\sigma_{F^{i,j}} \in \mathbb{D}^{1,\infty}$, $D^j F^\ell \bigcap_{p\in\mathbb{N}} Dom_p(\delta), j = 1, \dots d$ and the non degeneracy condition (2.5) holds for F. Then for every $\phi \in C_b^1(\mathbb{R}^m)$ we have*

$$\mathbb{E}(\partial_i \phi(F) G) = \mathbb{E}(\phi(F) H_i(F, G)), \quad i = 1, \dots, m \tag{2.11}$$

where

$$H_i(F, G) = \sum_{j=1}^m \delta(G\gamma_F^{ij} DF^j) = -\sum_{j=1}^m \Big(G\gamma_F^{i,j} LF^j + \langle D(G\gamma_F^{i,j}), DF^j\rangle\Big). \tag{2.12}$$

[3] For any $\xi \in \mathbb{R}^m$ one has

$$\langle \sigma_F \xi, \xi\rangle = \sum_{\ell,j=1}^m \sigma_F^{\ell j}\xi^\ell \xi^j = \int_0^1 \sum_{\ell,j=1}^m \sum_{i=1}^d D_t^i F^\ell \xi^\ell \, D_t^i F^j \xi^j dt = \sum_{i=1}^d \int_0^1 \Big|\sum_{\ell=1}^m D_t^i F^\ell \xi^\ell\Big|^2 dt \geq 0$$

so that σ_F is actually a non negative definite matrix.

We also have the following easy generalization. Consider $\xi^i = (\xi i1, \dots, \xi^{im})$ with $\xi^i \in Dom_\infty(\delta)$ for any $i = 1, \dots, m$. For $F = (F^1, \dots, F^m)$, and set

$$\sigma^{ij}_{F,\xi} = \langle DF^i, \xi^j \rangle = \sum_{k=1}^{d} \int_0^1 D^k_s F^i \xi^j(s) ds$$

Then, one can easily adapt the proof of the Malliavin integration by parts formula in order to get the following

Proposition 2.3.13. [Generalized MIbP formula] *Let $F = (F^1, \dots, F^m)$ with $F^i \in \mathbb{D}^{1,\infty}$ and let $\xi^i = (\xi^{i1}, \dots, \xi^{im})$ with $\xi^i \in Dom_\infty(\delta)$ for any $i = 1, \dots, m$. Suppose that $\sigma^{ij}_{F,\xi} \in \mathbb{D}^{1,\infty}$ for any $i, j = 1, \dots, m$ and $\mathbb{E}(|\det \sigma_{F,\xi}|^{-p}) < \infty$ for any p. Then for every $G \in \mathbb{D}^{1,\infty}$ and $\phi \in C^1_p(\mathbb{R}^m)$ we have*

$$\mathbb{E}(\partial_\ell \phi(F) G) = \mathbb{E}(\phi(F) H^\xi_\ell(F, G)), \quad \ell = 1, \dots, m \tag{2.13}$$

where

$$H^\xi_\ell(F, G) = \sum_{j=1}^{m} \delta(G \gamma^{\ell j}_{F,\xi} DF^j) = -\sum_{j=1}^{m} \Big(G \gamma^{i,j}_{F,\xi} L F^j + \langle D(G \gamma^{i,j}_{F\xi}), DF^j \rangle \Big) \tag{2.14}$$

being $\gamma_{F,\xi} = \sigma^{-1}_{F,\xi}$.

2.4 Higher order derivatives and integration by parts formulas

The higher order derivatives are defined in the same way as the first order derivatives: to begin one defines them on the simple functionals and then pass to the limit in order to obtain an extension. For $F \in S_n$ we define

$$D^{(i,j)}_{t_1 t_2} F = D^i_{t_1} D^j_{t_2} F = \frac{\partial^2}{\partial \Delta^{t_1,i}_n \partial \Delta^{t_2,j}_n} F$$

where $\Delta^{t,\ell}_n = \Delta^{k,\ell}_n$ for any $t \in [t^k_n, t^{k+1}_n)$. It is easy to see that the definition does not depend on n. It is clear that now

$$D^{(i,j)} : S \to L^p([0,1]^2, \mathscr{B}([0,1]^2), \mathrm{Leb}_2)$$

where Leb_n denotes the Lebesgue measure on $\mathbb{R}^n$. Moreover we have the following duality relation: for $U_1, U_2 \in P$ we have

$$\mathbb{E}\Big(\int_0^1 \int_0^1 D^{(i,j)}_{t_1 t_2} F U_1(t_1) U_2(t_2) dt_1 dt_2 \Big) = \mathbb{E}\Big(F \int_0^1 \Big(\int_0^1 U_1(t_1) U_2(t_2) \hat{d}W^i_{t_1} \Big) \hat{d}W^j_{t_2} \Big)$$

where in the above formula $\hat{d}W_t$ denotes the Skorohod integral. We do not give a more explicit expression of the above double integral. But recall that U_1 and U_2 are simple processes, then it is clear that the above random variable is in any L^p. Using the above formula one can check that $D^{(i,j)}$ is closable. Then one defines the domain of the second order derivative and the extension of this operator as usual. The notation is rather heavy so we prefer to give directly the space of second order differentiable functionals by using the Sobolev norms. We define on S the norm

$$\begin{aligned}\|F\|_{2,p}^p &= \|F\|_p^p + \mathbb{E}\Big(\Big(\int_0^1 \sum_{i=1}^d |D_t^i F|^2 dt\Big)^{p/2}\Big) + \\ &+ \mathbb{E}\Big(\Big(\int_{[0,1]^2} \sum_{i,j=1}^d |D_{t_1 t_2}^{(i,j)} F|^2 dt_1 dt_2\Big)^{p/2}\Big)\end{aligned}$$

and we put

$$\mathbb{D}^{2,p} = \overline{S}^{\|\cdot\|_{2,p}} \quad \text{and} \quad \mathbb{D}^{2,\infty} = \cap_p \mathbb{D}^{2,p}.$$

In order to define higher order derivatives we proceed similarly. We consider a multi-index $\alpha = (\alpha_1, \ldots, \alpha_k) \in \{1, \ldots, d\}^k$ and we denote $|\alpha| = k$. Then, for $F \in S$ we define

$$D_{t_1,\ldots,t_k}^{\alpha} F = D_{t_1}^{\alpha_1} \cdots D_{t_k}^{\alpha_k} F = \frac{\partial^k}{\partial \Delta_n^{t_1,\alpha_1} \cdots \partial \Delta_n^{t_k,\alpha_k}} F$$

and one has

$$D^\alpha \,:\, S \to L^p([0,1]^k, \mathscr{B}([0,1]^k), \mathrm{Leb}_k).$$

We use a duality argument in order to check that D^α is closable and we define the extension of the operator. Finally we define on S the norm

$$\|F\|_{k,p}^p = \|F\|_p^p + \sum_{r=1}^k \mathbb{E}\Big(\Big(\int_{[0,1]^r} \sum_{|\alpha|=r} |D_{t_1\cdots t_r}^{\alpha} F|^2 dt_1 \cdots dt_r\Big)^{p/2}\Big).$$

and we set

$$\mathbb{D}^{k,p} = \overline{S}^{\|\cdot\|_{k,p}}, \quad \mathbb{D}^{k,\infty} = \cap_p \mathbb{D}^{k,p} \quad \text{and} \quad \mathbb{D}^{\infty} = \cap_k \mathbb{D}^{k,\infty}.$$

The space $\mathbb{D}^\infty$ is the "good" space where to work because one is able to iterate the integration by parts formulas. It represents the analogues of C^∞ in the standard analysis. Moreover, $\mathbb{D}^\infty$ is an algebra.

Example 2.4.1. [Refined Clark-Ocone formula] *If $F \in \mathbb{D}^{2,2}$ then one has*

$$F = \underbrace{\mathbb{E}(F)}_{I} + \underbrace{\int_0^1 \mathbb{E}(D_s F) dW_s}_{II} + \underbrace{\int_0^1 \Big(\int_0^s \mathbb{E}(D_r D_s F \,|\, \mathscr{F}_r) dW_r\Big) dW_s}_{III}$$

which tells us that any $F \in \mathbb{D}^{2,2}$ can be split as the sum of three terms: a constant (term I), a Gaussian random variable (term II - notice that $\mathbb{E}(D_s F)$ is a deterministic square integrable function) and an iterated Ito integral (term III). Let us prove the above formula. By the Clark-Ocone formula one has

$$\begin{aligned}F &= \mathbb{E}(F) + \int_0^1 \mathbb{E}(D_s F \,|\, \mathscr{F}_s) dW_s \\ &= \mathbb{E}(F) + \int_0^1 \mathbb{E}(D_s F) dW_s + \int_0^1 \Big(\mathbb{E}(D_s F \,|\, \mathscr{F}_s) - \mathbb{E}(D_s F)\Big) dW_s\end{aligned}$$

Now, for each s, $D_sF \in \mathbb{D}^{1,2}$, so that the Clark-Ocone formula gives

$$D_sF = \mathbb{E}(D_sF) + \int_0^1 \mathbb{E}(D_rD_sF \,|\, \mathscr{F}_r)dW_r.$$

Therefore,

$$\begin{aligned}\mathbb{E}(D_sF \,|\, \mathscr{F}_s) - \mathbb{E}(D_sF) &= \mathbb{E}\Big(\int_0^1 \mathbb{E}(D_rD_sF \,|\, \mathscr{F}_r)dW_r \,\Big|\, \mathscr{F}_s\Big)\\ &= \int_0^s \mathbb{E}(D_rD_sF \,|\, \mathscr{F}_r)dW_r.\end{aligned}$$

By inserting above, the statement holds. If $F \in \mathbb{D}^\infty$, in principle such a procedure might be iterated infinitely many times and therefore F could be represented as the infinite sum of iterated Ito integrals (of any order) of deterministic functions. This actually holds for any $F \in L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ and is strictly connected to the Wiener chaos expansion (for details, see e.g. Nualart [18] or Sanz-Solé [19]).

A recurrence procedure based on the previous integration by parts formulas gives (recall that σ_F denotes the Malliavin covariance matrix and γ_F its inverse):

Theorem 2.4.2. *Suppose that $F^1, \dots, F^d \in \mathbb{D}^\infty$ and $G \in \mathbb{D}^\infty$. Suppose also that $\sigma_F^{i,j} \in \mathbb{D}^\infty$ and the non degeneracy condition (2.5) holds for F. Then for every multi-index α one has*

$$\mathbb{E}(\partial_\alpha f(F)G) = \mathbb{E}(f(F)H_\alpha(F,G)) \tag{2.15}$$

where for $|\alpha| = 1$, i.e. $\alpha = \{i\}$ as $i = 1, \dots, d$, $H_\alpha(F,G) = H_i(F,G)$ is given by

$$H_i(F,G) = \sum_{j=1}^m \delta(G\gamma_F^{ij}DF^j) = -\sum_{j=1}^m \Big(G\gamma_F^{i,j}LF^j + \langle D(G\gamma_F^{i,j}), DF^j\rangle\Big).$$

and for $|\alpha| = k > 1$

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

Moreover, $H_\alpha(F;G) \in \cap_{p\in\mathbb{N}} L^p$. In particular, for any k and p, there exist q and a constant C depending on k, p and d such that

$$\|H_\alpha(F, G\psi)\|_p \le C_\alpha \|G\|_{k,q}(1 + \|F\|_{2,q})^{\ell_{p,d}}\Big(1 + \mathbb{E}\big(|\det\sigma_F|^{-q}\big)^{1/q}\Big)$$

where α is such that $|\alpha| = k$.

The proof is omitted. In particular, the last inequality comes form Hölder's inequality and Meyer's inequality (for details, see e.g. Nualart [18] or Sanz-Solé [19]).

2.5 Diffusion processes

Let X denote the diffusion process solution to

$$X_t^i = x^i + \int_0^t b^i(X_s)ds + \sum_{j=1}^d \int_0^t \sigma_j^i(X_s)dW_s^j, \quad i = 1, \dots, m. \tag{2.16}$$

Assumption 2.5.1. *Suppose that for $i = 1, \ldots, m$ and $j = 1, \ldots, d$, σ^i_j, b^i have sub-linear growth, belong to $C^2(\mathbb{R}^m)$ and the first and second derivatives are bounded.*

Notice that under Assumption 2.5.1 one has:

- $x \mapsto b(x)$ and $x \mapsto \sigma(x)$ have sub-linear growth and are Lipschitz continuous on the compact sets;
- setting, for $(x, z) \in \mathbb{R}^m \times \mathbb{R}^m$, $i = 1, \ldots, m$ and $j = 1, \ldots, d$,

$$\bar{b}^i(x, z) = \sum_{k=1}^m \partial_{x^k} b^i(x) z^k \text{ and } \bar{\sigma}^i_j(x, z) = \sum_{k=1}^m \partial_{x^k} \sigma^i_j(x) z^k,$$

  also $(x, z) \mapsto \bar{b}(x, z)$ and $(x, z) \mapsto \bar{\sigma}(x, z)$ have sub-linear growth and are Lipschitz continuous on the compact sets.

We have then the following result.

Theorem 2.5.2. *Let Assumption 2.5.1 hold. Then, for $i = 1, \ldots, m$, $X^i_t \in \mathbb{D}^{1,p}$ for any p. Moreover, for a fixed $\ell = 1, \ldots, d$ and $s > 0$, the Malliavin derivative process $D^\ell_s X^i_t$, is equal to zero if $t < s$ and if $t \geq s$, it is the solution to*

$$\begin{aligned} D^\ell_s X^i_t = \quad & \sigma^i_\ell(X_s) + \int_s^t \sum_{k=1}^m \partial_{x^k} b^i(X_r) D^\ell_s X^k_r dr \\ & + \sum_{j=1}^d \int_s^t \sum_{k=1}^m \partial_{x^k} \sigma^i_j(X_r) D^\ell_s X^k_r dW^j_r. \end{aligned} \tag{2.17}$$

Proof. We sketch the main steps of the proof, which is a natural development of the one seen in Example 2.3.8 (and in fact, in some sense it is given by formula (2.9)). Let $n \in \mathbb{N}$ and $\bar{X}^n \equiv \bar{X}$ be the Euler scheme of step 2^{-n}, defined by

$$\bar{X}^i(t^{k+1}_n) = \bar{X}^i(t^k_n) + b^i(\bar{X}(t^k_n)) \frac{1}{2^n} + \sigma^i_j(\bar{X}(t^k_n)) \Delta^{k,j}_n \tag{2.18}$$

and $\bar{X}(0) = x$. We also interpolate on $[t^k_n, t^{k+1}_n)$ by keeping the coefficients $\sigma^i_j(\bar{X}(t^k_n))$ and $b^i(\bar{X}(t^k_n))$ to be constant but we allow the Brownian motion and the time to move. This means that $\bar{X}(t)$ solves the SDE

$$\bar{X}(t) = x + \int_0^t b(\bar{X}(\tau_s)) ds + \sum_{j=1}^d \int_0^t \sigma_j(\bar{X}(\tau_s)) dW^j_s$$

where $\tau_s = t^k_n$ for $s \in [t^k_n, t^{k+1}_n)$. Notice that $\bar{X}(t) \in \mathbb{D}^{1,2}$. In fact, for $t \in [t^k_n, t^{k+1}_n)$ one has

$$\begin{aligned} \bar{X}^i(t) \quad = \quad & x^i + \sum_{h=0}^{k-1} b^i(\bar{X}(t^h_n)) \frac{1}{2^n} + b^i(\bar{X}(t^k_n)) \big(t - t^k_n\big) \\ & + \sum_{h=0}^{k-1} \sum_{j=1}^d \sigma^i_j(\bar{X}(t^h_n)) \Delta^{h,j}_n + \sum_{j=1}^d \sigma^i_j(\bar{X}(t^k_n)) \big(W^j_t - W^j_{t^k_n}\big) \end{aligned}$$

So, $\bar{X}_t$ is a quite smooth functional of the increments of the Brownian motion on the dyadic intervals and increments of the type $W^j_t - W^j_{t^k_n}$, for $t \in [t^k_n, t^{k+1}_n)$. By the chain rule, one has

$$\begin{aligned} D^\ell_s \bar{X}^i(t) &= x^i + \sum_{h=0}^{k-1} D^\ell_s b^i(\bar{X}(t^h_n))\frac{1}{2^n} + D^\ell_s b^i(\bar{X}(t^k_n))\big(t - t^k_n\big) \\ &+ \sum_{h=0}^{k-1}\sum_{j=1}^{d} D^\ell_s\Big(\sigma^i_j(\bar{X}(t^h_n))\Delta^{h,j}_n\Big) + \sum_{j=1}^{d} D^\ell_s\Big(\sigma^i_j(\bar{X}(t^t_n))\big(W^j_t - W^j_{t^k_n}\big)\Big) \end{aligned}$$

It immediately follows that $D^\ell_s \bar{X}^i(t) = 0$ if $s > t$. If instead $s \le t$, one has

$$\begin{aligned} D^\ell_s b^i(\bar{X}(t^h_n)) &= \sum_{q=1}^{m} \partial_{x^q} b^i(\bar{X}(t^h_n)) D^\ell_s \bar{X}^q(t^h_n) \\ D^\ell_s\Big(\sigma^i_j(\bar{X}(t^h_n))\Delta^{h,j}_n\Big) &= \sum_{q=1}^{m} \partial_{x^q}\sigma^i_j(\bar{X}(t^h_n)) D^\ell_s \bar{X}^q(t^h_n)\Delta^{h,j}_n + \sigma^i_j(\bar{X}(t^h_n)) D^\ell_s \Delta^{h,j}_n \end{aligned}$$

and similarly for the term in which the increment $W_t - W_{t^k_n}$ appears. Since $D^\ell_s \bar{X}^q(t^h_n) = 0$ for any h such that $t^h_n < s$ and $D^\ell_s \Delta^{h,j}_n = \mathbf{1}_{s \in [t^h_n, t^{h+1}_n)} \mathbf{1}_{\ell = j}$, we can resume by writing

$$\begin{aligned} D^\ell_s \bar{X}^i(t) &= \sigma^i_\ell(\bar{X}(\tau_s)) + \int_0^t \sum_{q=1}^{m} \partial_{x^q} b^i(\bar{X}(\tau_r)) D^\ell_s \bar{X}^q(\tau_r) dr \\ &+ \sum_{j=1}^{d} \int_s^t \sum_{q=1}^{m} \partial_{x^q}\sigma^i_j(\bar{X}(\tau_r)) D^\ell_s \bar{X}^q(\tau_r) dW^j_r, \end{aligned}$$

So, in order to prove that $X_t \in \mathbb{D}^{1,p}$, we have to prove that $\bar{X}(t)$ converges in $L^p(\Omega)$ to $X(t)$ - and this is a standard result concerning the Euler scheme approximation (see e.g. Kloeden and Platen [10]) - and that $D\bar{X}(t)$ converges in $L^p(H_d)$ to some limit, and this will be DX_t.
Assume now that s is fixed and let $Q_s(t), t \ge s$ be the solution of the $d \times m$-dimensional SDE

$$\begin{aligned} Q^{\ell,i}_s(t) &= \sigma^i_\ell(X_s) + \int_0^t \sum_{q=1}^{m} \partial_{x^q} b^i(X_r) Q^{\ell,q}_s(r) dr \\ &+ \sum_{j=1}^{d} \int_s^t \sum_{q=1}^{m} \partial_{x^q}\sigma^i_j(X_r) Q^{\ell,q}_s(r) dW^j_r, \end{aligned}$$

Notice that the solution $Q^{\ell,i}_s(t)$ exists: in fact, the [very] multi-dimensional process $(X(t), Q_s(t))$ as $t \ge s$ is a diffusion process solving a SDE whose drift and diffusion coefficients satisfy the usual properties allowing to get existence and uniqueness of the solution (in fact, by Assumption 2.5.1 we get the Lipschitz continuity on compact sets and the sublinear growth for the [very] multidimensional drift and diffusion coefficient associated to the pair $(X(t), Q_s(t))$). Then $D^\ell_s \bar{X}^i(t), t \ge s$ is the Euler scheme for $Q^{\ell,i}_s(t), t \ge s$ and so standard arguments give $\left\|D_s\bar{X}(t) - Q_s(t)\right\|_p \leq C_p 2^{-n/2}, \forall p > 1$. A quick inspection of the arguments leading to this inequality shows that C_p does not depend on s. Define now $Q_s(t) = D_s\bar{X}(t) = 0$ for $t \leq s$. Then, we obtain

$$\begin{aligned}\mathbb{E}\Big(\Big|\int_0^1 |Q_s(t) - D_s\bar{X}(t)|^2 ds\Big|^{p/2}\Big) &= \mathbb{E}\Big(\Big|\int_0^t |Q_s(t) - D_s\bar{X}(t)|^2 ds\Big|^{p/2}\Big)\\ &\leq \frac{C_p}{2^{n/2}} \to 0 \text{ as } n \to \infty.\end{aligned}$$

Therefore, $X_t \in \mathbb{D}^{1,p}$ and $D_sX(t) = Q_s(t)$. Recall that for a fixed t, the path $s \mapsto D_sX(t)$ is an element of $L^2([0,1])$ and so is determined ds-a.e. But we have here a precise version $Q_s(t)$ such that $t \mapsto Q_s(t)$, for $t \geq s$, is continuous and solves a SDE. So, from now on we will refer to the Malliavin derivative of $X(t)$ as to the solution of (2.17) as $s \leq t$ and $D_sX(t) = 0$ for $s > t$. $\square$

We can represent the Malliavin derivative also in another way. Let us first recall the following important result.

Theorem 2.5.3. *Let Assumption 2.5.1 holds. Then the dependence of the diffusion X on the initial datum x is C^1 and setting $Y_t = \partial_x X_t$, that is*

$$Y_t^{ij} = \frac{\partial X_t^i}{\partial x^j}, i, j = 1, \ldots, m,$$

then Y is the solution to the following SDE

$$Y_t = I + \int_0^t \partial b(X_s) Y_s ds + \sum_{j=1}^d \int_0^t \partial\sigma_j(X_s) Y_s dW_s^j,$$

where ∂b and $\partial\sigma_j$ denote the $m \times m$ matrix fields defined as

$$(\partial b)^{ik} = \partial_k b^i \quad \text{and} \quad (\partial\sigma_j)^{ik} = \partial_k \sigma_j^i, \quad i, k = 1, \ldots, m, j = 1 \ldots, d$$

respectively. Moreover, the inverse matrix valued process $Z_t = Y_t^{-1}$ exists and satisfies

$$Z_t = I - \int_0^t Z_s\Big(\partial b(X_s) - \sum_{j=1}^d (\partial_j\sigma)^2\Big) ds - \sum_{j=1}^d \int_0^t Z_s \partial\sigma_j(X_s) dW_s^j,$$

The proof is omitted. But, let us remark that it follows a technique similar to the one used in Theorem 2.5.2: take the Euler scheme for X, prove that it is C^1 in the initial datum, find the SDE associated to $\bar{Y} = \partial\bar{X}$ and then pass to the limit.

We can now state the following result:

Proposition 2.5.4. *Let Assumption 2.5.1 hold and let $Y_t = \partial X_t$. Then, one has*

$$D_s X_t = Y_t Y_s^{-1} \sigma(X_s)\, \mathbf{1}_{t>s}.$$

Proof. For $t \geq s$ set $Q_t = Y_t Z_s \sigma(X_s)$. By Ito's formula,

$$Q_t = \sigma(X_s) + \int_s^t \partial b(X_r) Q_r dr + \sum_{j=1}^d \int_0^t \partial\sigma_j(X_r) Q_r dW_r^j,$$

which is, from (2.17), the same equation satisfied by D_sX_t. $\square$

Example 2.5.5. [Geometric Brownian motion-Black&Scholes model] Let X solve the SDE

$$dX_t^i = \mu^i X_t^i dt + \sum_{j=1}^d \sigma_j^i X_t^i dW_t^j,\ X_0^i = x^i,\quad i = 1, \ldots, m. \tag{2.19}$$

Let us stress that μ and σ are supposed to be constant. A suitable use of the Ito's formula give the exact solution:

$$X_t^i = x^i \exp\Big\{\Big(\mu^i - \frac{1}{2}\sum_{j=1}^d (\sigma_j^i)^2\Big)t + \sum_{j=1}^d \sigma_j^i W_t^j\Big\},\quad i = 1, \ldots, m. \tag{2.20}$$

Then, following Example 2.2.8, one can apply the chain rule, so that

$$D_s^j X_t^i = \sigma_j^i X_t^i \,\mathbf{1}_{\{t>s\}}. \tag{2.21}$$

But we can arrive to the same result by using Proposition 2.5.4: Assumption 2.5.1 is trivially fulfilled and (2.20) gives

$$\partial X_t = Y_t = \mathrm{diag}\Big[\frac{X_t^1}{x^1}, \ldots, \frac{X_t^m}{x^m}\Big] \text{ and then } Y_s^{-1} = \mathrm{diag}\Big[\frac{x^1}{X_s^1}, \ldots, \frac{x^m}{X_s^m}\Big]$$

then, as $t > s$,

$$D_s^j X_t^i = \sum_{k=1}^m (Y_t Y_s^{-1})_{ik} (\sigma(X_s))_{kj} = \frac{X_t^i}{x^i}\,\frac{x^i}{X_s^i}\,\sigma_j^i X_s^i = \sigma_j^i\, X_t^i.$$

2.6 Appendix. Wiener chaos decomposition

For the sake of simplicity of notations, we assume here that W is a one-dimensional Brownian motion. For the general case, see e.g. Nualart [18] or Sanz-Solé [19]. Recall that $\mathscr{F}_t$ is the σ-algebra generated by the Brownian motion up to time t and completed with the $\mathbb{P}$-null sets.
For $n \in \mathbb{N}$, let H_n denote the nth Hermite polynomial, that is

$$H_0(x) = 1, \quad H_n(x) = \frac{(-1)^n}{n!} e^{\frac{x^2}{2}} \frac{d^n}{dx^n}\Big(e^{\frac{-x^2}{2}}\Big) \text{ if } n \geq 1. \tag{2.22}$$

Such polynomials give the power series expansion in t of the function $F(x,t) = \exp(tx - \frac{t^2}{2})$:

$$F(x,t) = e^{tx - \frac{t^2}{2}} = e^{\frac{x^2}{2} - \frac{(x-t)^2}{2}} = e^{\frac{x^2}{2}} \sum_{n=0}^\infty \frac{t^n}{n!}\frac{d^n}{dt^n}\Big(e^{-\frac{(x-t)^2}{2}}\Big)\Big|_{t=0} = \sum_{n=0}^\infty H_n(x) t^n.$$

As an immediate consequence, for $n \geq 1$ one has:

$$H_n'(x) = H_{n-1}(x) \tag{2.23}$$
$$(n+1)H_{n+1}(x) = xH_n(x) - H_{n-1}(x) \tag{2.24}$$
$$H_n(-x) = (-1)^n H_n(x) \tag{2.25}$$

Indeed, (2.23) and (2.24) follow from $\partial_x F = tF$ and $\partial_t F = (x-t)F$, while (2.25) follow from $F(-x,t) = F(x,-t)$.
The link between Hermite polynomials and Gaussian random variables is given by the following

Lemma 2.6.1. *Let (X,Y) denote a Gaussian r.v. on $\mathbb{R}^2$, with $\mathbb{E}(X)=\mathbb{E}(Y)=0$ and $\mathrm{Var}(X)=\mathrm{Var}(Y)=1$. Then*

$$\mathbb{E}(H_n(X)H_m(Y)) = \begin{cases} 0 & \text{if } n\neq m \\ \dfrac{1}{n!}\, e^{\mathbb{E}(XY)} & \text{if } n=m \end{cases}$$

As a consequence, $\{H_n\}_n$ defines a sequence of orthogonal polynomials in $L^2(\mathbb{R}, \mathscr{B}(\mathbb{R}),\mu_1)$, where μ_1 denotes the standard Gaussian measure on $\mathbb{R}$.

Proof. Notice first that the covariance matrix Γ of the r.v. (X,Y) is given by $\Gamma_{11}=\Gamma_{22}=1$ and $\Gamma_{12}=\Gamma_{21}=\mathbb{E}(XY)$. Now, for any $s,t\in\mathbb{R}$, one has

$$\mathbb{E}(F(X,s)F(Y,t)) = \mathbb{E}\Big(\exp\Big(sX-\frac{s^2}{2}\Big)\exp\Big(tY-\frac{t^2}{2}\Big)\Big) = e^{st\,\mathbb{E}(XY)}.$$

Taking the $(n+m)$th derivative $\partial_s^n\partial_t^m$ for $s=t=0$, on the l.h.s. we get

$$\begin{aligned}\partial_s^n\partial_t^m\mathbb{E}(F(X,s)F(Y,t))\Big|_{s=t=0} &= \mathbb{E}\Big(\partial_s^n F(X,s)\Big|_{s=0}\partial_t^m F(Y,t)\Big|_{t=0}\Big)\\ &= n!m!\,\mathbb{E}(H_n(X)H_m(Y))\end{aligned}$$

On the r.h.s. we easily have $\partial_s^n\partial_t^m e^{st\,\mathbb{E}(XY)}|_{s=t=0}=0$ if $n\neq m$, otherwise $\partial_s^n\partial_t^n e^{st\,\mathbb{E}(XY)}|_{s=t=0}=n!\,e^{\mathbb{E}(XY)}$, and the statement follows. Finally, if $X=Y$, one gets

$$\int_{\mathbb{R}} H_n(x)H_m(x)\mu_1(dx) = \frac{1}{n!}\,\mathbf{1}_{n=m}$$

where μ_1 denotes the standard Gaussian measure on $\mathbb{R}$, so that $\{H_n\}_n$ defines a sequence of orthogonal polynomials in $L^2(\mathbb{R},\mathscr{B}(\mathbb{R}),\mu_1)$. $\square$

For $\phi\in H_1$ we set $W(\phi)=\int_0^1\phi_t dW_t$ and we define $\mathrm{span}\{e^{W(\phi)}\,;\,\phi\in H_1\}$ the subspace of $L^2(\Omega,\mathscr{F}_1,\mathbb{P})$ given by the r.v.'s which are linear combinations of r.v.'s of the form $e^{W(\phi)}$. Then,

Lemma 2.6.2. $\mathrm{span}\{e^{W(\phi)}\,;\,\phi\in H_1\}$ *is a dense subspace of $L^2(\Omega,\mathscr{F}_1,\mathbb{P})$.*

Proof. It is sufficient to prove that if $X\in L^2(\Omega,\mathscr{F}_1,\mathbb{P})$ is such that $\mathbb{E}(Xe^{W(\phi)})=0$ for any $\phi\in H_1$ then $X=0$ a.s. In fact, since $\phi\mapsto W(\phi)$ is linear, one has also

$$\mathbb{E}(Xe^{\sum_{i=1}^m\lambda_i W(\phi_i)})=0$$

for any $m\geq 1$, $\lambda_1,\ldots,\lambda_m\in\mathbb{R}$ and $\phi_1,\ldots\phi_m\in H_1$. Therefore, the Laplace transform of the random vector $(W(\phi_1),\ldots,W(\phi_m))$ done w.r.t. the signed measure $\mathbb{P}_X(A)=\mathbb{E}(X\mathbf{1}_A)$, $A\in\mathscr{F}_1$, is null. Then, $(W(\phi_1),\ldots,W(\phi_m))$ has a null law under $\mathbb{P}_X$, that is

$$\mathbb{E}(X\mathbf{1}_{(W(\phi_1),\ldots,W(\phi_m))\in B})=0 \text{ for any } B\in\mathbb{R}^m.$$

Consequently (recall that $(W_{t_1},\ldots,W_{t_m})=(W(\phi_1),\ldots,W(\phi_m))$ when $\phi_i(t)=\mathbf{1}_{t\leq t_i}$), $\mathbb{E}(X\mathbf{1}_A)=0$ for any $A\in\mathscr{F}_1$, that is $X=0$ a.s. $\square$

We set now

$$\mathscr{H}_0=\mathbb{R},\quad \mathscr{H}_n=\mathrm{span}\{H_n(W(\phi))\,;\,\phi\in H_1\} \text{ if } n\geq 1. \tag{2.26}$$

The set $\mathscr{H}_n$ is called the nth Wiener chaos. By Lemma 2.6.1

$$\mathbb{E}(H_n(W(\phi))H_m(W(\psi))) = 0 \quad \text{if } n \neq m$$

for any $\phi, \psi \in H_1$, so that $\mathscr{H}_n$ and $\mathscr{H}_m$ are orthogonal subspaces of $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ if $n \neq m$. Moreover, the following important result holds

Theorem 2.6.3. [Wiener chaos decomposition] *The space $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ can be decomposed into the infinite orthogonal sum of the subspaces $\mathscr{H}_n$:*

$$L^2(\Omega, \mathscr{F}_1, \mathbb{P}) = \bigoplus_{n=0}^{\infty} \mathscr{H}_n.$$

For $F \in L^2(\Omega, \mathscr{F}_1, \mathbb{P})$, the representation $F = \sum_{n=0}^{\infty} F_n$ with $F_n \in \mathscr{H}_n$ is called the Wiener chaos decomposition of F.

Proof. It is sufficient to prove that if X is orthogonal to each $\mathscr{H}_n$ then $X = 0$ a.s. In particular, one has that $\mathbb{E}(XH_n(W(\phi))) = 0$ for any n and $\phi \in H_1$. Without loss of generality, we may assume that $\|\phi\|_{H_1} = 1$. Now, x^n can be seen as a linear combination of $H_k(x)$ for $k \leq n$ and then we have $\mathbb{E}(X(W(\phi))^n) = 0$ for each n, so that

$$\mathbb{E}(Xe^{W(\phi)}) = 0$$

for any $\phi \in H_1$. By Lemma 2.6.2, we immediately obtain that X is orthogonal to $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$, and then $X = 0$ a.s. $\square$

As a consequence, we obtain

Proposition 2.6.4. *The set S of the simple functional is a dense subset in $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$.*

Proof. In view of the Wiener chaos decomposition, it is sufficient to prove that any r.v. in $\mathscr{H}_n$ can be approximated in $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ by a sequence of r.v.'s in S, for any n. This in turn follows by proving that any r.v.'s of the type $H_n(W(\phi))$ is the $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$-limit of r.v.'s in S. As a consequence, it is sufficient to prove that all the r.v.'s of the form $(W(\phi))^n$, with $n \geq 1$ and $\phi \in H_1$, can be approximated in $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ by a sequence in S. Now, let us prove this last assertion.

Since $\phi \in H_1$, there exists a sequence piecewise constant functions $\{\phi_N\}_N$ of the type $\phi_N(t) = \sum_{k=0}^{2^N-1} c_{k,N} \mathbf{1}_{t \in [t_N^k, t_N^{k+1})}$ such that $\phi_N \to \phi$ in H_1. One has

$$W(\phi_N) = \sum_{k=0}^{2^N-1} c_{k,N} \Delta_N^k$$

so that $(W(\phi_N))^n \in S$ for any N. We show now that $(W(\phi_N))^n \to (W(\phi))^n$ in $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ as $N \to \infty$. For this, we use the estimate[4]

$$|y^n - x^n| \leq n|y - x|(|x| + |y|)^{n-1}.$$

[4]For any differentiable function f one has $f(y) - f(x) = (y - x)\int_0^1 f'(x + t(y - x))dt$. Taking $f(\xi) = \xi^n$, one gets $|y^n - x^n| \leq |y - x| \int_0^1 n|x + t(y - x)|^{n-1}dt$. By recalling that $|x + t(y - x)| \leq |x| + |y|$ for $t \in [0, 1]$, one obtains the desired inequality.

So, we obtain

$$\begin{aligned}\mathbb{E}\Big(\big((W(\phi))^n-(W(\phi_N))^n\big)^2\Big) &\leq n^2\mathbb{E}\Big(|W(\phi)-W(\phi_N)|^2\times \\ &\qquad \times(|W(\phi)|+|W(\phi_N)|)^{2(n-1)}\Big) \\ &\leq n^2\sqrt{\mathbb{E}\big(|W(\phi-\phi_N)|^4\big)}\times \\ &\quad \times\sqrt{\mathbb{E}\big((|W(\phi)|+|W(\phi_N)|)^{4(n-1)}\big)} \\ &\leq C_n\sqrt{\mathbb{E}\big(|W(\phi-\phi_N)|^4\big)}\times \\ &\quad \times\sqrt{\mathbb{E}\big(|W(\phi)|^{4(n-1)}\big)+\mathbb{E}\big(|W(\phi_N)|^{4(n-1)}\big)}\end{aligned}$$

where C_n denotes a suitable positive constant. Now, for $\psi\in H_1$ one has $\mathbb{E}(|W(\psi)|^k)=d_k\|\psi\|_{H_1}^k$ where d_k denotes the moment of order k of a standard Gaussian r.v. Therefore, by setting again C_n as a suitable positive constant, we get

$$\mathbb{E}\Big(\big((W(\phi))^n-(W(\phi_N))^n\big)^2\Big)\leq C_n\|\phi-\phi_N\|_{H_1}^2\times\Big(\|\phi\|_{H_1}^{4(n-1)}+\|\phi_N\|_{H_1}^{4(n-1)}\Big)^{1/2}$$

which converges to 0 as $N\to\infty$, and the statement holds. □

Chapter 3

Applications to Finance

We describe here only some cases of use of Malliavin calculus techniques in Finance, giving representation formulas for the replicating portfolio, for the sensitivities and the conditional expectation. The last two ones are used in practice to set up Monte Carlo methods: for the computation of the Greeks and to build up Monte Carlo algorithms for the pricing of American style options. Let us stress that we reduce here our discussion to the Gaussian Malliavin calculus. In fact, we do not consider other sources for the noise. Nevertheless, it is worth to say that many recent research papers in Finance allow the presence of jumps in the model for the underlying asset prices and, in particular, jump-diffusions are taken into account. In such a case, one could develop also a Malliavin calculus in the direction of the jump times and/or the jump amplitudes of the compounded Poisson process modelling the noise coming from the jumps. For details, we refer to the papers [2] and [4] by Bally, Bavouzet-Morel and Messaoud.

3.1 The Clark-Ocone formula and the replicating portfolio

Let $\{S_t\}_t$ denote the underlying asset price process, evolving following

$$dS_t^i = \mu^i(S_t)S_t^i dt + \sum_{j=1}^{d} \sigma_j^i(S_t)S_t^i dW_t^j, \quad i = 1, \dots, d \tag{3.1}$$

with $S_0 = x$. If the drift term μ and the volatility matrix field σ are assumed to be bounded and Lipschitz continuous, the sde (3.1) admits a unique solution.

From a financial point of view, we assume that the short interest rate process is constant, that is $r_t = r > 0$, σ is invertible and furthermore, the matrix $a(x) = \sigma\sigma^*(x)$ is uniformly elliptic. Under these assumptions, there exists a unique equivalent martingale measure $\mathbb{P}^*$, under which the discounted asset price process is a martingale. Therefore, we can assume to work directly under $\mathbb{P}^*$, so that the underlying asset price process S_t and the associated discounted one $\tilde{S}_t = e^{-rt} S_t$ evolve as

$$dS_t^i = rS_t^i dt + \sum_{j=1}^d \sigma_j^i(S_t) S_t^i dW_t^j, \quad i = 1, \dots, d \tag{3.2}$$

$$d\tilde{S}_t^i = \sum_{j=1}^d \sigma_j^i(S_t) \tilde{S}_t^i dW_t^j, \quad i = 1, \dots, d \tag{3.3}$$

respectively. By Ito's formula, as $i = 1, \dots, d$ one has

$$\begin{aligned} S_t^i &= x^i \exp\Big(\int_0^t \big(r - \frac{1}{2}\sum_{j=1}^d \sigma_j^i(S_u)^2\big) du + \int_0^t \sum_{j=1}^d \sigma_j^i(S_u) dW_u^j\Big), \\ \tilde{S}_t^i &= x^i \exp\Big(-\int_0^t \frac{1}{2}\sum_{j=1}^d \sigma_j^i(S_u)^2 du + \int_0^t \sum_{j=1}^d \sigma_j^i(S_u) dW_u^j\Big). \end{aligned}$$

Such formulas show that if $x \in \mathbb{R}_+^d$ then S_t and $\tilde{S}_t$ take values in $\mathbb{R}_+^d$.
Let (H, T) denote a European option, that is: T is the maturity time and H is a non negative $\mathscr{F}_T$-measurable random variable representing the payoff of the contingent claim. A replicating portfolio for such an option is given by a process

$$V_t = \phi_t^0 e^{rt} + \sum_{i=1}^d \phi_t^i S_t^i$$

such that:

- [technical assumptions] $\phi^0, \phi^1, \dots, \phi^d$ are adapted processes such that $\phi^0 \in L^2([0, T])$ a.s. and $\phi^i \in L^2([0, 1] \times \Omega)$ for any $i = 1, \dots d$;
- [self-financing] $dV_t = r\phi_t^0 e^{rt} dt + \sum_{i=1}^d \phi_t^i dS_t^i$, $t < T$
- [admissibility] $V_t \geq 0$ a.s. for a.e. $t \leq T$;
- [replicating] $V_T = H$ a.s.

Now, if H is square integrable[1] a replicating portfolio for (H, T) exists and is given by

$$V_t = \mathbb{E}(e^{-r(T-t)} H \,|\, \mathscr{F}_t).$$

The value V_t is also the (non arbitrage) price of the option (H, T) as seen at time t. This closes the problem of the price, but what about the hedging problem? That is, how to find the shares $\phi^0, \phi^1, \dots, \phi^d$ to invest in order to replicate the option? The answer is the following.
First, the discounted portfolio $\tilde{V}_t = e^{-rt} V_t$ satisfies the sde

$$d\tilde{V}_t = \sum_{i=1}^d \phi_t^i d\tilde{S}_t^i = \sum_{j=1}^d \sum_{i=1}^d \phi_t^i \sigma_j^i(S_t) \tilde{S}_t^i dW_t^j. \tag{3.4}$$

[1]We stress that we are working under the risk neutral measure, so the square integrability and the further expectation giving the price have to be intended under such a measure.

Moreover, in the developments giving the non arbitrage price of the option, one has that

$$\tilde{V}_t = \mathbb{E}(e^{-rT} H \mid \mathscr{F}_t)$$

is a square integrable Brownian martingale, so that it can be represented as

$$\tilde{V}_t = V_0 + \int_0^t \sum_{j=1}^d \Phi_s^j \, dW_s^j \tag{3.5}$$

where Φ^i is an adapted process such that $\Phi^i \in L^2([0,1] \times \Omega)$, $i = 1, \dots, d$. Therefore, (3.4) and (3.5) give

$$\Phi_t^j = \sum_{i=1}^d \phi_t^i \sigma_j^i(S_t) \tilde{S}_t^i \quad j = 1, \dots, d$$

so that

$$\phi_t^i = \frac{1}{\tilde{S}_t^i} \sum_{j=1}^d \Phi_t^j \sigma_{ji}^{-1}(S_t) \quad i = 1, \dots, d. \tag{3.6}$$

Notice that once $\phi^1, \dots, \phi^d$ and V are known then also ϕ^0 is known:

$$\phi_t^0 = \tilde{V}_t - \sum_{i=1}^d \phi_t^i \tilde{S}_t^i.$$

So, the only unpleasant point is that (3.6) gives the replicating strategy $\phi^1, \dots, \phi^d$ in terms of the processes $\Phi^1, \dots, \Phi^d$ which in turn are given by the representation theorem for Brownian martingales only in an abstract form. But, if the payoff satisfies some regularity properties in Malliavin sense then the Clark Ocone formula allows to conclude. In fact,

Proposition 3.1.1. *If $H \in \mathbb{D}^{1,2}$ then*

$$\phi_t^i = \frac{e^{-r(T-t)}}{S_t^i} \sum_{j=1}^d \mathbb{E}\big(D_t^j H \mid \mathscr{F}_t\big) \sigma_{ji}^{-1}(S_t) \quad i = 1, \dots, d$$

Proof. By the Clark-Ocone formula one has

$$e^{-rT} H = \mathbb{E}(e^{-rT} H) + e^{-rT} \int_0^T \sum_{j=1}^d \mathbb{E}\big(D_s^j H \mid \mathscr{F}_s\big) dW_s^j$$

so that

$$\tilde{V}_t = \mathbb{E}\big(e^{-rT} H \mid \mathscr{F}_t\big) = \mathbb{E}(e^{-rT} H) + e^{-rT} \int_0^t \sum_{j=1}^d \mathbb{E}\big(D_s^j H \mid \mathscr{F}_s\big) dW_s^j.$$

Then, in (3.5) one has

$$\Phi_t^j = e^{-rT} \mathbb{E}\big(D_t^j H \mid \mathscr{F}_t\big), \quad j = 1, \dots, d, \quad t \in [0, T]$$

and (3.6) gives the result. □

As an example, suppose that $H = \Psi(S_T)$. This is a case in which the option price can be written in terms of a price-function. In fact, by using the Markov property, one has

$$V_t = P(t, S_t) \quad \text{where} \quad P(t,\xi) = e^{-r(T-t)}\mathbb{E}\big(\Psi(S_T^{t,\xi})\big).$$

Now, if one requires that $\sigma \in C_b^2$ then $S \in \mathbb{D}^{1,\infty}$ and if moreover $\Psi \in C_p^1(\mathbb{R}^d)$ then $H = \Psi(S_T) \in \mathbb{D}^{1,2}$. So, in this case one has

$$D_t^j H = \sum_{k=1}^d \partial_k \Psi(S_T) D_t^j S_T^k.$$

Now, from Proposition 2.5.4 one has $D_t S_T = Y_T Z_t \hat\sigma(S_t)$, in which $Y = \partial S$ denotes the first variation process of S, $Z = Y^{-1}$ and $\hat\sigma$ denote the diffusion coefficient associated to S, that is $\hat\sigma_j^i(x) = \sigma_j^i(x)x^i$. Now, setting $\Lambda(x) = \mathrm{diag}[x^1, \dots, x^d]$ one has $\hat\sigma(x) = \Lambda(x)\sigma(x)$, so that

$$D_t S_T = Y_T Z_t \Lambda(S_t)\sigma(S_t), \quad t \le T$$

and then

$$D_t^j H = \sum_{k=1}^d \partial_k \Psi(S_T)\big(Y_T Z_t \Lambda(S_t)\sigma(S_t)\big)^{kj} = \big(\nabla\Psi(S_T) Y_T Z_t \Lambda(S_t)\sigma(S_t)\big)^j.$$

By Proposition 3.1.1, as $i = 1, \dots, d$ the replicating strategy is then given by

$$\phi_t^i = \frac{e^{-r(T-t)}}{S_t^i} \sum_{j=1}^d \mathbb{E}\Big(\big(\nabla\Psi(S_T) Y_T Z_t \Lambda(S_t)\sigma(S_t)\big)^j \,|\, \mathscr{F}_t\Big)\sigma_{ji}^{-1}(S_t)$$

Now, we have

$$\begin{aligned}
\phi_t^i &= \frac{e^{-r(T-t)}}{S_t^i}\mathbb{E}\Big(\big(\nabla\Psi(S_T) Y_T Z_t \Lambda(S_t)\sigma(S_t)\sigma^{-1}(S_t)\big)^i \,|\, \mathscr{F}_t\Big) \\
&= \frac{e^{-r(T-t)}}{S_t^i}\mathbb{E}\Big(\big(\nabla\Psi(S_T) Y_T Z_t \Lambda(S_t)\big)^i \,|\, \mathscr{F}_t\Big) \\
&= \frac{e^{-r(T-t)}}{S_t^i}\sum_{q=1}^d \mathbb{E}\Big(\big(\nabla\Psi(S_T) Y_T\big)^q \,|\, \mathscr{F}_t\Big) Z_t^{qi} S_t^i
\end{aligned}$$

and finally

$$\phi_t^i = e^{-r(T-t)}\mathbb{E}\Big(\big(\nabla\Psi(S_T) Y_T Z_t\big)^i \,|\, \mathscr{F}_t\Big) \tag{3.7}$$

But we can go further in the interpretation of the strategy. In fact, notice that

$$Y_T = \partial S_T \equiv \partial_x S_T^{0,x} = \partial_x S_T^{t,S_t^{0,x}} = \partial_\xi S_T^{t,\xi}|_{\xi = S_t^{0,x}} \cdot \partial_x S_t^{0,x} = \partial_\xi S_T^{t,\xi}|_{\xi = S_t^{0,x}} \cdot Y_t$$

so that

$$Y_T Z_t = \partial_\xi S_T^{t,\xi}|_{\xi = S_t^{0,x}}.$$

Therefore, by using the Markov property in (3.7), we get

$$\begin{aligned}
\phi_t^i &= e^{-r(T-t)}\mathbb{E}\Big(\big(\nabla\Psi(S_T^{t,\xi})\partial_\xi S_T^{t,\xi}\big)^i|_{\xi=S_t^{0,x}} \mid \mathscr{F}_t\Big) \\
&= e^{-r(T-t)}\mathbb{E}\Big(\big(\nabla\Psi(S_T^{t,\xi})\partial_\xi S_T^{t,\xi}\big)^i\Big)|_{\xi=S_t^{0,x}} \\
&= e^{-r(T-t)}\partial_{\xi^i}\mathbb{E}\big(\Psi(S_T^{t,\xi})\big)|_{\xi=S_t^{0,x}}
\end{aligned}$$

that is

$$\phi_t^i = \Delta^i(t, S_t) \quad \text{where} \quad \Delta^i(t,\xi) = \partial_{\xi^i} P(t,\xi)$$

$P(t,\xi)$ being the price-function associated to the option $H = \Psi(S_T)$.
By resuming, the Clark Ocone formula gives the replicating strategy in terms of Malliavin derivatives. If the payoff function Ψ is smooth enough then the replicating strategy is given by the Delta function, a fact which is known also for payoff functions which are not smooth (e.g. for digital type options).

3.2 Sensitivity computation

Let us first give a brief informal introduction to the concept of financial Greek. Suppose to have a financial asset S_t and that its dynamic under the risk neutral measure follows

$$dS_t = rdt + \sigma(S_t)S_t dW_t, \quad S_0 = x, \tag{3.8}$$

where W is an one dimensional Brownian Motion and for simplicity we assume that the spot rate r is constant.
Consider a European option whose payoff depends only on the final value of the underling asset price, that is it is a function of S_T. The price of such an option of maturity T and payoff function ϕ is given by

$$\Pi = \mathbb{E}(e^{-rT}\phi(S_T)). \tag{3.9}$$

A Greek is the derivative of the option price with respect to a prescribed parameter, i.e. it is a measure of the sensibility of the option price with respect to such a parameter.
Greeks are very important in Mathematical Finance because they could be used to measure the stability of the option (see e.g. the Vega Greek, that is the derivative of the option price with respect to the volatility σ) or to describe the replicating portfolio (see e.g. the Δ, i.e. the derivative of the option price with respect to x, that is the initial value of S).
In general, if we denote with α the parameter we are interested in, one aims to compute

$$\partial_\alpha \Pi(e^{-rT}\phi(S_T)). \tag{3.10}$$

If we can use the rule of interchange between expectation and differentiation, we would have

$$\partial_\alpha \Pi = \partial_\alpha \mathbb{E}(e^{-rT}\phi(S_T)) = \mathbb{E}(e^{-rT}\phi'(S_T)\partial_\alpha S_T). \tag{3.11}$$

(in which we have supposed $\alpha \neq r$, otherwise also the derivative of e^{-rT} has to be taken into account). When this expression does not have a closed form formula, but ϕ is differentiable, Monte Carlo simulations or finite difference methods can be used in order to approximate the Greek directly using (3.11). Unfortunately in various cases the payoff function is singular and in this case the finite difference method do not work very well. Then one can resort to the use of the "Malliavin integration by parts formula". In this case we obtain

$$\partial_\alpha \Pi = \mathbb{E}(e^{-rT}\phi^{'}(S_T)\partial_\alpha S_T) = \mathbb{E}(e^{-rT}\phi(S_T)H(S_T, \partial_\alpha S_T)), \tag{3.12}$$

Moreover, if the MIbP formula holds then the equality holds for any smooth function so that if the function ϕ is not differentiable but can be suitably approximated by differentiable functions, then the equality between the left and the right hand side in (3.12) continues to hold, so that one has

$$\partial_\alpha \Pi = \mathbb{E}(e^{-rT}\phi(S_T)\Theta_\alpha), \tag{3.13}$$

with $\Theta_\alpha = H(S_T, \partial_\alpha S_T)$. Now, if the weight Θ_α can be written or at least approximated in a good way and in particular, can be simulated, one can perform a Monte Carlo method, giving

$$\partial_\alpha \Pi \simeq \frac{e^{-rT}}{Q}\sum_{q=1}^{Q}\phi(S_T^q)\Theta_\alpha^q),$$

where $\{S_T^q\}_q$ and $\{\Theta_\alpha^q\}_q$ denote independent replications of S_T and Θ_α respectively and Q has to be chosen sufficiently large.

3.2.1 The delta

Let X here denote a diffusion process, solving to

$$X_t^i = x^i + \int_0^t b^i(X_s)ds + \sum_{j=1}^{d}\int_0^t \sigma_j^i(X_s)dW_s^j, \quad i = 1, \ldots, d. \tag{3.14}$$

Notice that we are supposing $d = m$, i.e. X and W take value both on $\mathbb{R}^d$. What we are going to state can be generalized also if such a condition does not hold. But, for the sake of simplification, we avoid such a complication.

Proposition 3.2.1. *Let Assumption 2.5.1 hold and, in addition, suppose that the diffusion coefficient σ is invertible and that $\mathbb{E}(\int_0^T |\sigma^{-1}(X_s)Y_s|^{2+\delta}ds) < \infty$ for some $\delta > 0$, in which Y denotes, as usual, the first variation process ($Y_t^{ki} = \partial_{x^k} X_t^i$). Let $G \in \mathbb{D}^{1,\infty}$ be a r.v. which does not depend on x. Then for any measurable function ϕ with polynomial growth one has*

$$\partial_{x^i}\mathbb{E}(\phi(X_T)G) = \mathbb{E}(\phi(X_T)\Theta_i^G), \quad i = 1, \ldots, d,$$

$$\Theta_i^G = \frac{1}{T}\sum_{\ell=1}^{d}\Big(G\int_0^T (\sigma^{-1}(X_s)Y_s)^{\ell i}\, dW_s^\ell - \int_0^T D_s^\ell G(\sigma^{-1}(X_s)Y_s)^{\ell i}\, ds\Big). \tag{3.15}$$

Proof. Suppose first that $\phi \in C_b^1$, the general case to be deserved later. Then, we can pass the derivative inside the expectation, so that

$$\partial_{x^i}\mathbb{E}(\phi(X_T)G) = \mathbb{E}\Big(\sum_{k=1}^{d}\partial_{x^k}\phi(X_T)\,\partial_{x^i}X_t^k G\Big) = \mathbb{E}\Big(\sum_{k=1}^{d}\partial_{x^k}\phi(X_T)\, Y_T^{ki}G\Big)$$

By using Proposition 2.5.4, $Y_T = D_s X_T \sigma^{-1}(X_s) Y_s$ for any $s < T$, and then

$$\begin{aligned}\sum_{k=1}^{d} \partial_{x^k}\phi(X_T)\, Y_T^{ki} &= \sum_{k=1}^{d} \partial_{x^k}\phi(X_T)\,(D_s X_T \sigma^{-1}(X_s) Y_s)^{ki} \\ &= \sum_{k=1}^{d} \partial_{x^k}\phi(X_T) \sum_{\ell=1}^{d} D_s^\ell X_T^k (\sigma^{-1}(X_s) Y_s)^{\ell i} \\ &= \sum_{\ell=1}^{d}\sum_{k=1}^{d} \partial_{x^k}\phi(X_T) D_s^\ell\, X_T^k (\sigma^{-1}(X_s) Y_s)^{\ell i} \\ &= \sum_{\ell=1}^{d} D_s^\ell \phi(X_T) (\sigma^{-1}(X_s) Y_s)^{\ell i}\end{aligned}$$

in which we have used the chain rule. Therefore,

$$\begin{aligned}\sum_{k=1}^{d} \partial_{x^k}\phi(X_T)\, Y_T^{ki} &= \frac{1}{T}\int_0^T \sum_{k=1}^{d} \partial_{x^k}\phi(X_T)\, Y_T^{ki}\, ds \\ &= \frac{1}{T}\int_0^T \sum_{\ell=1}^{d} D_s^\ell \phi(X_T)(\sigma^{-1}(X_s) Y_s)^{\ell i}\, ds.\end{aligned}$$

Now, by applying the duality we get

$$\mathbb{E}\Big(\sum_{k=1}^{d} \partial_{x^k}\phi(X_T)\, Y_T^{ki} G\Big) = \mathbb{E}(\phi(X_T)\Theta_i^G)$$

where

$$\Theta_i^G = \sum_{\ell=1}^{d} \delta^\ell\Big(\frac{1}{T}(\sigma^{-1}(X_\cdot)Y_\cdot)_{\ell i} G\Big)$$

Finally, $\sigma^{-1}(X_s)Y_s$ is adapted, so that

$$\Theta_i^G = \frac{1}{T}\sum_{\ell=1}^{d}\Big(G\int_0^T (\sigma^{-1}(X_s)Y_s)^{\ell i} dW_s^\ell - \int_0^T D_s^\ell G (\sigma^{-1}(X_s)Y_s)^{\ell i} ds\Big).$$

If ϕ is not in the class C_b^1, the statement follows by using standard density arguments: one can regularize ϕ with some suitable mollifier and by using density arguments, the statement follows. □

As a consequence, in the Black and Scholes model we obtain

Corollary 3.2.2. *Suppose $b^i(x) = \mu^i x^i$ and $\sigma_i^j(x) = \sigma_{ij} x^i$, $i, j = 1, \dots, d$, with σ invertible. Then the weights Θ_i's in Theorem 3.2.1 are given by*

$$\Theta_i^G = \frac{1}{T x^i}\sum_{\ell=1}^{d} (\sigma^{-1})_{\ell i}\Big(G W_T^\ell - \int_0^T D_s^\ell G\, ds\Big), \quad i = 1, \dots d.$$

Proof. Setting $\Gamma(x) = \mathrm{diag}[x^1, \dots, x^d]$, one has $\sigma(x) = \Gamma(x)\,\sigma$, where σ is the volatility matrix, so that $\sigma^{-1}(X_s) = \sigma^{-1}\Gamma^{-1}(X_s)$. Moreover, since $Y_s^{ij} = \partial_{x^j} X_s^i$, we can write shortly $Y_s = \Gamma(X_s)\,\Gamma^{-1}(x)$. Therefore

$$\sigma^{-1}(X_s)Y_s = \sigma^{-1}\Gamma^{-1}(X_s)\,\Gamma(X)\,\Gamma(x) = \sigma^{-1}\Gamma^{-1}(x),$$

so that

$$\begin{aligned} \Theta_i^G &= \frac{1}{T}\sum_{\ell=1}^d \int_0^T \delta^\ell\Big((\sigma^{-1}(X_s)Y_s)_{\ell i} G\Big) = \frac{1}{Tx^i}\sum_{\ell=1}^d (\sigma^{-1})_{\ell i}\delta^\ell(G) \\ &= \frac{1}{Tx^i}\sum_{\ell=1}^d (\sigma^{-1})_{\ell i}\Big(GW_T^\ell - \int_0^1 D_s^\ell G\,ds\Big) \end{aligned}$$

and the statement holds. $\square$

As a consequence, one has

Proposition 3.2.3. *Suppose $b^i(x) = \mu^i x^i$ and $\sigma_i^j(x) = \sigma_{ij}x^i$, $i,j = 1,\dots,d$, with σ invertible. Then for any ϕ with polynomial growth, for $i,j = 1,\dots,d$ one has*

$$\partial_{x^i}\mathbb{E}(\phi(X_T)) = \mathbb{E}(\phi(X_T)\Lambda_i^\Delta) \quad \partial^2_{x^ix^j}\mathbb{E}(\phi(X_T)) = \mathbb{E}(\phi(X_T)\Lambda_{ij}^\Gamma)$$

where

$$\Lambda_i^\Delta = \frac{1}{Tx^i}\sum_{\ell=1}^d (\sigma^{-1})_{\ell i} W_T^\ell \tag{3.16}$$

$$\Lambda_{ij}^\Gamma = \frac{1}{T}\Lambda_i^\Delta\Lambda_j^\Delta - \frac{1}{Tx^ix^j}\sum_{\ell=1}^d (\sigma^{-1})_{\ell i}(\sigma^{-1})_{\ell j} - \mathbf{1}_{i=j}\frac{1}{x^i}\Lambda_i^\Delta \tag{3.17}$$

Proof. The proof of (3.16) is immediate from Corollary 3.2.2 applied to $G = 1$. As for the gamma, notice that

$$\partial^2_{x^ix^j}\mathbb{E}(\phi(X_T)) = \partial^2_{x^j}\mathbb{E}(\phi(X_T)\Lambda_i^\Delta) = \partial_{x^j}\mathbb{E}(\phi(X_T)\frac{1}{x^i}\,x^i\Lambda_i^\Delta)$$

Now, $G_i = x^i\Lambda_i^\Delta$ is independent of x, so that

$$\partial^2_{x^ix^j}\mathbb{E}(\phi(X_T)) = \frac{1}{x^i}\partial_{x^j}\mathbb{E}(\phi(X_T)\,G_i) - \mathbf{1}_{i=j}\frac{1}{x_i^2}\mathbb{E}(\phi(X_T)\,G_i)$$

and in particular,

$$\Lambda_{ij}^\Gamma = \Theta_j^{G_i} - \mathbf{1}_{i=j}\frac{1}{x^i}\Lambda_i^\Delta.$$

By applying again Corollary 3.2.2, one immediately obtains (3.17). $\square$

Let us write explicitly the weights allowing to represent the delta and gamma Greeks in dimension $d = 1,2$ for the Black and Scholes model.

- Dimension $d = 1$**.** Here, one immediately obtains

$$\Lambda^\Delta = \frac{W_T}{x\,T\,\sigma} \quad \text{and} \quad \Lambda^\Gamma = (\Lambda^\Delta)^2 - \frac{1}{Tx^2\sigma^2} - \frac{\Lambda^\Delta}{x}.$$

- Dimension $d = 2$. In dimension 2, the volatility matrix is usually written as a sub-diagonal one:

$$\sigma = \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2 \rho & \sigma_2 \sqrt{1-\rho^2} \end{pmatrix}$$

where σ_i stands for the volatility of the ith asset, $i = 1, 2$, and $\rho \in (-1, 1)$ gives the correlation between the Gaussian noises. Then,

$$\sigma^{-1} = \frac{1}{\sigma_1 \sigma_2 \sqrt{1-\rho^2}} \begin{pmatrix} \sigma_2 \sqrt{1-\rho^2} & 0 \\ -\sigma_2 \rho & \sigma_1 \end{pmatrix}$$

so that

$$\Lambda_1^\Delta = \frac{W_T^1 \sqrt{1-\rho^2} - W_T^2 \rho}{x^1 T \sigma_1} \quad \text{and} \quad \Lambda_2^\Delta = \frac{W_T^2}{x^2 T \sigma_2 \sqrt{1-\rho^2}}.$$

As for the gamma weight, one has

$$\begin{aligned} \Lambda_{ii}^\Gamma &= \frac{1}{T}(\Lambda_i^\Delta)^2 - \frac{1}{T(x^i)^2 \sigma_i^2 \sqrt{1-\rho^2}} - \frac{1}{x^i} \Lambda_i^\Delta, \quad i = 1, 2 \\ \Lambda_{12}^\Gamma = \Lambda_{21}^\Gamma &= \frac{1}{T} \Lambda_1^\Delta \Lambda_2^\Delta + \frac{\rho}{T x^1 x^2 \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \end{aligned}$$

For example, for a digital change option, we have

$$\text{price} = e^{-rt} \mathbb{E}^* (\mathbf{1}_{S_T^1 > S_T^2})$$

and the delta and gamma Greeks are given by

$$\begin{aligned} \Delta_i &= \partial_{x^i} \left(e^{-rt} \mathbb{E}^* (\mathbf{1}_{S_T^1 > S_T^2}) \right) = e^{-rt} \mathbb{E}^* (\mathbf{1}_{S_T^1 > S_T^2} \Lambda_1^\Delta) \quad i = 1, 2 \\ \Gamma_{ij} &= \partial^2_{x^i x^j} \left(e^{-rt} \mathbb{E}^* (\mathbf{1}_{S_T^1 > S_T^2}) \right) = e^{-rt} \mathbb{E}^* (\mathbf{1}_{S_T^1 > S_T^2} \Lambda_{ij}^\Gamma) \quad i, j = 1, 2 \end{aligned}$$

where $\mathbb{E}^*$ denotes the expectation under the risk neutral measure (i.e., the one under which (S^1, S^2) evolves following the Black and Scholes model with $\mu^1 = \mu^2 = r$).

3.2.2 Some other examples

We discuss here some examples giving sensitivities for the price of suitable European options.

Delta for Asian options in the Black&Scholes model ($d = 1$)

Assume that S follows the Black Scholes dynamics:

$$dS_t = rS_t dt + \sigma S_t dW_t, \quad S_0 = x.$$

Here, we compute the delta weight for Asian type options, that is options whose payoff is a function ϕ of the time mean over $[0, T]$: $\frac{1}{T} \int_0^T S_u du$. For simplicity, set $F = \frac{1}{T} \int_0^T S_u du$. One has

$$\partial_x \mathbb{E}(\phi(F)) = \mathbb{E}(\phi'(F) \partial_x F) \tag{3.18}$$

where

$$\partial_x F = \frac{1}{T}\int_0^T \partial_x S_u \, du = \frac{1}{T}\int_0^T \frac{S_u}{x}\, du.$$

Now, one has

$$D_u F = \frac{1}{T} D_u \int_0^T S_t dt = \frac{1}{T}\int_u^T D_u S_t dt = \frac{\sigma}{T}\int_u^T S_t dt$$

and

$$\begin{aligned}
\Big(\int_0^T S_u du\Big)^2 &= \int_0^T \frac{d}{dt}\Big(\Big(\int_0^t S_u du\Big)^2\Big) dt = \int_0^T 2\Big(\int_0^t S_u du\Big) S_t dt \\
&= 2\int_0^T S_u \int_u^T S_t dt du = \frac{2T}{\sigma}\int_0^T S_u D_u F du.
\end{aligned}$$

Therefore

$$\partial_x F = \frac{1}{Tx}\int_0^T S_u du = \frac{2}{\sigma x \int_0^T S_u du}\int_0^T S_u D_u F du$$

and

$$\phi'(F)\partial_x F = \frac{2}{\sigma x}\int_0^T \frac{S_u}{\int_0^T S_t dt}\phi'(F) D_u F du = \frac{2}{\sigma x}\int_0^T \frac{S_u}{\int_0^T S_t dt} D_u \phi(F) du$$

By inserting in (3.18) and by applying the duality one gets

$$\mathbb{E}\big(\phi'(F)\partial_x F\big) = \mathbb{E}\Big(\int_0^T \frac{2S_u}{\sigma x \int_0^T S_t dt} D_u \phi(F) du\Big) = \mathbb{E}\big(\phi(F)\Lambda^{\Delta}\big)$$

where

$$\Lambda^{\Delta} = \delta\Big(\frac{2S.}{\sigma x \int_0^T S_t dt}\Big).$$

By using the property for product, we get

$$\begin{aligned}
\Lambda^{\Delta} &= \frac{2}{\sigma x}\Big[\frac{1}{\int_0^T S_t dt}\int_0^T S_t dW_t - \int_0^T S_t D_t\Big(\frac{1}{\int_0^T S_t dt}\Big) dt\Big] \\
&= \frac{2}{\sigma x}\Big[\frac{1}{\int_0^T S_t dt}\int_0^T S_t dW_t + \int_0^T S_t \frac{D_t \int_0^T S_u du}{(\int_0^T S_t dt)^2} dt\Big] \\
&= \frac{2}{\sigma x}\Big[\frac{1}{\int_0^T S_t dt}\int_0^T S_t dW_t + \int_0^T S_t \frac{\sigma \int_t^T S_u du}{(\int_0^T S_t dt)^2} dt\Big]
\end{aligned}$$

Consider the last addendum. One has:

$$\int_0^T S_t \frac{\sigma \int_t^T S_u du}{(\int_0^T S_t dt)^2} dt = \frac{\sigma}{(\int_0^T S_t dt)^2}\int_0^T -\frac{1}{2}\frac{d}{dt}\Big(\Big(\int_t^T S_u du\Big)^2\Big) dt = \frac{\sigma}{2}.$$

Therefore

$$\Lambda^{\Delta} = \frac{2}{\sigma x}\Big[\frac{1}{\int_0^T S_t dt}\int_0^T S_t dW_t + \frac{\sigma}{2}\Big]$$

Moreover, from $dS_t = rS_t dt + \sigma S_t dW_t$ we get

$$\int_0^T S_t dW_t = \frac{1}{\sigma}\Big[S_T - x - r\int_0^T S_u du\Big]$$

so that

$$\Lambda^\Delta = \frac{2}{\sigma x}\Big[\frac{S_T - x - r\int_0^T S_u du}{\sigma \int_0^T S_t dt} + \frac{\sigma}{2}\Big] = \frac{2}{\sigma x}\Big[\frac{S_T - x}{\sigma T F} - \frac{r}{\sigma} + \frac{\sigma}{2}\Big]$$

Sensitivity w.r.t. the correlation in a stochastic volatility model

Assume that S evolves as the following stochastic volatility version of the Black Scholes dynamics:

$$\begin{aligned} dS_t &= rS_t dt + \eta_t S_t dW_t^1, \quad S_0 = x \\ d\eta_t &= \kappa(\theta - \eta_t)dt + \beta dW_t^2, \end{aligned}$$

where W^1 and W^2 are two correlated Brownian motions with

$$d\langle W^1, W^2\rangle_t = \rho dt, \quad \rho \in [-1, 1].$$

We consider a digital option that has payoff $\mathbf{1}_{[K,\infty)}(S_T)$ and we want to compute the sensitivity of the option price w.r.t. ρ.
In this case is not possible to use Proposition 3.2.1 directly. So, one can proceed as follows.
First, one sets

$$W_t^1 = \sqrt{1-\rho^2}B_t^1 + \rho B_t^2, \quad W_t^2 = B_t^2$$

where B^1 and B^2 are two independent Brownian Motions (in fact, if $B = (B^1, B^2)$ is given by the inverse transformation, it is straightforward to show that it is a 2-dimensional Brownian motion). Therefore, the SDE for S and η becomes

$$\begin{aligned} dS_t &= rS_t dt + \eta_t S_t(\sqrt{1-\rho^2}dB_t^1 + \rho dB_t^2), \quad S_0 = x \\ d\eta_t &= \kappa(\theta - \eta_t)dt + \beta dB_t^2. \end{aligned}$$

Notice that S_T can be written as (apply Ito's formula)

$$S_T = x\exp\Big\{\int_0^t (r - \frac{1}{2}\eta_s^2)ds + \sqrt{1-\rho^2}\int_0^T \eta_t dB_t^1 + \rho\int_0^T \eta_t dB_t^2\Big\}. \tag{3.19}$$

Now, for any smooth function ϕ, one has

$$\partial_\rho \mathbb{E}(\phi(S_T)) = \mathbb{E}(\phi'(S_T)\partial_\rho S_T). \tag{3.20}$$

By (3.19) one gets easily

$$\partial_\rho S_T = S_T\Big(\int_0^T \eta_t dB_t^2 - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \eta_t dB_t^1\Big) = S_T\, G, \tag{3.21}$$

with

$$G = \int_0^T \eta_t dB_t^2 - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \eta_t dB_t^1. \tag{3.22}$$

We use the chain rule -$D\exp(F) = \exp(F)DF$,- the relation $D^i_s(\int_0^T \phi_r dW^i_r) = \phi(s)$ and the relation (3.21) in order to obtain

$$
\begin{aligned}
D^1_s S_T &= S_T D^1_s\Big(\sqrt{1-\rho^2}\int_0^T \eta_t dB^1_t + \rho\int_0^T \eta_t dB^2_t\Big) \\
&= S_T(\sqrt{1-\rho^2}\times \eta_s) = \partial_\rho S_T\,\frac{\eta_s\sqrt{1-\rho^2}}{G} \\
&= \partial_\rho S_T\,\frac{\eta_s\sqrt{1-\rho^2}}{G}.
\end{aligned} \tag{3.23}
$$

so that

$$D^1_s\phi(S_T) = \phi'(S_T)D^1_s S_T = \phi'(S_T)\partial_\rho S_T\,\frac{\eta_s\sqrt{1-\rho^2}}{G}.$$

Therefore

$$\phi'(S_T)\partial_\rho S_T = \frac{G}{\eta_s\sqrt{1-\rho^2}}D^1_s\phi(S_T) = \frac{1}{T\sqrt{1-\rho^2}}\int_0^T D^1_s\phi(S_T)\,\frac{G}{\eta_s}ds.$$

We use now the duality formula with respect to B^1 and we obtain

$$\mathbb{E}(\phi'(S_T)\partial_\rho S_T) = \frac{1}{T\sqrt{1-\rho^2}}\mathbb{E}\Big(\int_0^T D^1_s\phi(S_T)\,\frac{G}{\eta_s}ds\Big) = \mathbb{E}(\phi(S_T)\Theta_\rho)$$

where

$$\Theta_\rho = \frac{1}{T\sqrt{1-\rho^2}}\delta_1\Big(\frac{G}{\eta}\Big)$$

Now, by using the properties of the Skorohod integral of a special product and for adapted processes, one has

$$\delta_1\Big(\frac{G}{\eta}\Big) = G\delta_1\Big(\frac{1}{\eta}\Big) - \int_0^T D^1_s G\,\frac{1}{\eta_s}ds = \int_0^T \eta_s^{-1}dB^1_s - \int_0^T D^1_s G\,\frac{1}{\eta_s}ds.$$

Moreover, it easy to see that

$$D^1_s G = -\frac{\rho}{\sqrt{1-\rho^2}}\eta_s,$$

so that, in conclusion, we can say that

$$\partial_\rho\mathbb{E}(\phi(S_T)) = \mathbb{E}(\phi(S_T)\Theta_\rho)$$

with

$$\Theta_\rho = G\int_0^T \eta_s^{-1}dB^1_s + \frac{\rho T}{\sqrt{1-\rho^2}}.$$

Sensitivity w.r.t. the volatility for change options

Let S^1 and S^2 be two financial assets following

$$
\begin{aligned}
dS^1_t &= rS^1_t dt + \sigma_1 S^1_t dW^1_t,\ \ S^1_0 = x^1 \\
dS^2_t &= rS^2_t dt + \sigma_2 S^2_t dW^2_t,\ \ S^2_0 = x^2
\end{aligned}
$$

where W^1 and W^2 are two correlated Brownian Motions with $d\langle W^1, W^2\rangle_t = \rho dt$, $\rho \in (-1, 1)$.
Consider an option which pays one dollar if $S_T^1 > S_T^2$, i.e. the payoff is $\phi = \mathbf{1}_{\{S_T^1 > S_T^2\}}$. We want here to compute the sensitivity w.r.t. the volatility, for example $\partial_{\sigma^1}\Pi$.
First we write

$$W_t^1 = \sqrt{1-\rho^2}B_t^1 + \rho B_t^2, \quad W_t^2 = B_t^2$$

where B^1 and B^2 are two independent Brownian Motions. Then W^1 and W^2 are Brownian Motions with correlation ρ and the SDE associated to (S^1, S^2) becomes

$$\begin{aligned} dS_t^1 &= rS_t^1 dt + \sigma_1 S_t^1\left(\sqrt{1-\rho^2}dB_t^1 + \rho dB_t^2\right), \quad S_0^1 = x^1 \\ dS_t^2 &= rS_t^2 dt + \sigma_2 S_t^2 dB_t^2, \quad S_0^2 = x^2 \end{aligned}$$

At time T one has

$$\begin{aligned} S_T^1 &= x^1 \exp\left\{\left(r - \tfrac{1}{2}\sigma_1^2\right)T + \sigma_1\sqrt{1-\rho^2}B_T^1 + \sigma_1\rho B_T^2\right\} \\ S_T^2 &= x^2 \exp\left\{\left(r - \tfrac{1}{2}\sigma_2^2\right)T + \sigma_2 B_T^2)\right\} \end{aligned}$$

We now put

$$S_T = \frac{S_T^1}{S_T^2} = \frac{x^1}{x^2}\exp\left\{\frac{1}{2}(\sigma_1^2 - \sigma_2^2)T + \sigma_1\sqrt{1-\rho^2}B_T^1 + (\sigma_1\rho - \sigma_2)B_T^2\right\}, \quad (3.24)$$

so we want compute

$$\partial_{\sigma_1}\mathbb{E}(\mathbf{1}_{[1,\infty)}(S_T)).$$

Setting $\Pi = \mathbb{E}(\phi(S_T))$, we have

$$\partial_{\sigma_1}\Pi = \mathbb{E}[\phi(S_T)\partial_{\sigma_1}S_T]. \quad (3.25)$$

Notice that

$$\partial_{\sigma_1}S_T = S_T(\sqrt{1-\rho^2}B_T^1 + \rho B_T^2) \quad (3.26)$$

then

$$S_T = \frac{\partial_{\sigma_1}S_T}{(\sqrt{1-\rho^2}B_T^1 + \rho B_T^2)}. \quad (3.27)$$

Now,

$$\begin{aligned} D_s^1 S_T \quad &= S_T D_s^1(\sigma_1\sqrt{1-\rho^2}B_T^1 + (\sigma_1\rho - \sigma_2)B_T^2) \\ &= S_T\,(\sigma_1\sqrt{1-\rho^2}) = \partial_{\sigma_1}S_T\,\frac{(\sigma_1\sqrt{1-\rho^2})}{(\sqrt{1-\rho^2}B_T^1 + \rho B_T^2)}, \end{aligned}$$

so that

$$D_s^1(\phi(S_T)) = \phi'(S_T)D_s^1 S_T = \phi'(S_T)\partial_{\sigma^1}S_T\,\frac{\sigma_1\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}B_T^1 + \rho B_T^2}.$$

Therefore

$$\phi'(S_T)\partial_{\sigma_1} S_T = D^1_s(\phi(S_T))\frac{\sqrt{1-\rho^2}B^1_T + \rho B^2_T}{\sigma_1\sqrt{1-\rho^2}}. \tag{3.28}$$

By using the duality formula with respect to B^1, we have

$$\begin{aligned}\mathbb{E}(\phi'(S_T)\partial_{\sigma_1} S_T) &= \mathbb{E}\Big(\frac{1}{T}\int_0^T \frac{D^1_s(\phi(S_T))(\sqrt{1-\rho^2}B^1_T + \rho B^2_T)}{\sigma_1\sqrt{1-\rho^2}}ds\Big) \\ &= \frac{1}{T\sigma_1\sqrt{1-\rho^2}}\mathbb{E}(\phi(S_T)\delta_1(G))\end{aligned}$$

where

$$G = \sqrt{1-\rho^2}B^1_T + \rho B^2_T.$$

Now, $\delta_1(G) = G\delta_1(1) - \int_0^T D^1_s G ds$ and since $D^1_s G = \sqrt{1-\rho^2}\mathbf{1}_{s<T}$ we obtain

$$\delta_1(G) = (\sqrt{1-\rho^2}B^1_T + \rho B^2_T)B^1_T - T\sqrt{1-\rho^2}.$$

In conclusion, $\partial_{\sigma_1}\Pi = \mathbb{E}(\phi(S_T)\Theta_\sigma)$ with

$$\Theta_\sigma = \frac{(\sqrt{1-\rho^2}B^1_T + \rho B^2_T)B^1_T - T\sqrt{1-\rho^2}}{T\sigma_1\sqrt{1-\rho^2}}$$

3.3 Conditional expectation

To avoid a too technical machinery, we consider here only the case of the Black and Scholes model. The general case can be considered as well, see e.g. the paper by Bouchard, Ekeland and Touzi [6].

Consider the risk neutral world, so that X is driven by

$$dX^i_t = (r-\eta^i)X^i_t dt + \sum_{j=1}^d \sigma_{ij}X^i_t dW^j_t, \quad \text{with} \quad X^i_0 = x^i, \quad i = 1,\dots,d$$

where: $x = (x^1,\dots,x^d) \in \mathbb{R}^d_+$ denotes the vector of the initial asset values; r is the (constant) spot rate and $\eta \in \mathbb{R}^d$ being the vector of the dividends of the option; σ denotes the $d \times d$ volatility matrix which we suppose to be non-degenerate; W is a d-dimensional correlated Brownian motion.
Without loss of generality, one can suppose that σ is a sub-triangular matrix, that is $\sigma_{ij} = 0$ whenever $i < j$, and that W is a standard d-dimensional Brownian motion. Thus, any component of X_t can be written as

$$X^i_t = x^i \exp\Big(h^i t + \sum_{j=1}^i \sigma_{ij}W^j_t\Big), \qquad i = 1,\dots,d \tag{3.29}$$

where from now on we set $h^i = r - \eta^i - \frac{1}{2}\sum_{j=1}^i \sigma^2_{ij}$, $i = 1,\dots,d$. The aim is to study the conditional expectation, that is

$$\mathbb{E}(\Phi(X_t)\,|\,X_s = \alpha),$$

where $0 < s < t$, $\alpha \in \mathbb{R}^d_+$ and $\mathscr{E}_b(\mathbb{R}^d)$ denotes the class of the measurable functions with polynomial growth, that is $|\Phi(y)| \le C(1+|y|^m)$ for some m.
In few words, to this goal it suffices to consider an auxiliary process $\tilde X$ with independent components for which a formula for the conditional expectation immediately follows as a product. In a second step, such a formula can be adapted to the original process X by means of an (invertible) function giving X from the auxiliary process $\tilde X$.

3.3.1 Diagonalization procedure and first formulas

To our purposes, let us set

$$\tilde X^i_t = x^i \exp\Big(h^i\, t + \sigma_{ii} W^i_t\Big), \qquad i = 1, \dots, d. \tag{3.30}$$

As a first result, we study a transformation allowing to handle the new process $\tilde X$ in place of the original process X:

Lemma 3.3.1. *For any $t \ge 0$ there exists an invertible function $F_t(\cdot) : \mathbb{R}^d_+ \to \mathbb{R}^d_+$ such that $X_t = F_t(\tilde X_t)$ and $\tilde X_t = G_t(X_t)$, where (set $\prod_{j=1}^0 \stackrel{def}{=} 1$)*

$$F^i_t(y) = y^i \prod_{j=1}^{i-1} \Big(\frac{y^j}{x^j}\, e^{-h^j t}\Big)^{\tilde\sigma_{ij}} \quad \textit{and} \quad G^i_t(z) = z^i \prod_{j=1}^{i-1} \Big(\frac{z^j}{x^j}\, e^{-h^j t}\Big)^{\widehat\sigma_{ij}} \tag{3.31}$$

as $i = 1, \dots, d$ and $y, z \in \mathbb{R}^d_+$, where

$$\tilde\sigma_{ij} = \frac{\sigma_{ij}}{\sigma_{jj}}, \ i, j = 1, \dots, d, \quad \textit{and} \quad \widehat\sigma = \tilde\sigma^{-1} \tag{3.32}$$

The proof is straightforward and we omit it. Let us add a further result.

Lemma 3.3.2. *Suppose $d = 1$: $X_t = x\, e^{\mu t + \sigma W_t}$, being $\mu \in \mathbb{R}$, $x, \sigma \in \mathbb{R}_+$ and W a one dimensional Brownian motion. Suppose $f, g : \mathbb{R} \to \mathbb{R}$, where f has a polynomial growth and g has a continuous derivative. Then for any $0 < s < t$ one has:*

$$\mathbb{E}(f(X_t)\, g'(X_s)) = \mathbb{E}\Big(f(X_t) g(X_s) \frac{\Delta W_{s,t}}{\sigma s(t-s) X_s}\Big)$$

where $\Delta W_{s,t} = (t-s)(W_s + \sigma\, s) - s(W_t - W_s)$. As a consequence, for any fixed $\alpha \in \mathbb{R}$, the following formulas hold:

$$\mathbb{E}(f(X_t)\, g'(X_s - \alpha)) = \mathbb{E}\Big(f(X_t)\, \frac{g(X_s - \alpha)}{\sigma s(t-s) X_s}\, \Delta W_{s,t}\Big).$$

Proof. The proof consists in applying twice the MIbP formula, first on the time interval $[0, s]$ and secondly over $[s, t]$.

1) MIbP formula over $[0, s]$**.**

One has $D_r g(X_s) = g'(X_s)\sigma X_s$ for any $r < s$. Therefore,

$$g'(X_s) = \int_0^s \frac{D_r g(X_s)}{\sigma s X_s}\, dr$$

and by duality,

$$\begin{aligned}\mathbb{E}(f(X_t)g'(X_s)) = &\ \mathbb{E}\Big(\int_0^s D_r g(X_s)\cdot\frac{f(X_t)}{\sigma s X_s}\,dr\Big) = \mathbb{E}\Big(g(X_s)\delta\Big(\frac{f(X_t)}{\sigma s X_s}\Big)\Big)\\ = &\ \mathbb{E}\Big(g(X_s)\Big[\frac{f(X_t)}{\sigma s X_s}W_s - \int_0^s D_r\Big(\frac{f(X_t)}{\sigma s X_s}\Big)\,dr\Big]\Big)\end{aligned}$$

in which we have used the fact that $\delta(F) = F\delta(1) - \int_0^s D_rF\,dr = FW_s - \int_0^s D_rF\,dr$ applied to $F = f(X_t)/(\sigma s X_s)$. Now, recalling that $D_rX_u = \sigma X_u$ for $r < u$, one obtains

$$D_r\Big(\frac{f(X_t)}{\sigma s X_s}\Big) = -f(X_t)\,\frac{1}{sX_s} + f'(X_t)\frac{X_t}{sX_s}$$

Then,

$$\mathbb{E}(f(X_t)g'(X_s)) = \mathbb{E}\Big(f(X_t)g(X_s)\frac{W_s+\sigma s}{\sigma s X_s}\Big) - \underbrace{\mathbb{E}\Big(f'(X_t)g(X_s)\frac{X_t}{X_s}\Big)}_{(*)}.$$

We have now to handle the term (*), which is “bad” because of the presence of the derivative of f: we are now going to drop it.

2) MIbP formula over $[s,t]$.

By using arguments similar to the ones developed above but over $[s,t]$, we can write

$$\begin{aligned}\mathbb{E}\Big(f'(X_t)g(X_s)\frac{X_t}{X_s}\Big) = &\ \mathbb{E}\Big(\int_s^t \frac{g(X_s)}{\sigma(t-s)X_s}D_rf(X_t)\,dr\Big)\\ = &\ \mathbb{E}\Big(f(X_t)\delta\Big(\frac{g(X_s)}{\sigma(t-s)X_s}\Big)\Big)\\ = &\ \mathbb{E}\Big(f(X_t)\frac{g(X_s)}{\sigma(t-s)X_s}\,(W_t-W_s)\Big)\end{aligned}$$

in which we have used the fact that $D_r(g(X_s)/(\sigma(t-s)X_s)) = 0$ for any $r\in(s,t)$. By inserting this term in place of the term (*), in conclusion we obtain

$$\mathbb{E}(g'(X_s)\,f(X_t)) = \mathbb{E}\Big(f(X_t)g(X_s)\Big[\frac{\Delta W_{s,t}}{\sigma s(t-s)X_s}\Big]\Big).$$

Let us observe that to achieve this representation one has implicitly assumed that f is regular (C^1), which is not true in general. But this is not really a problem: one can regularize f with some suitable mollifier and by using density arguments, the statement follows. $\square$

We are now ready to state the main result of this section

Theorem 3.3.3. ***[Representation formulas I: without localization]*** *Let $0<s<t$, $\Phi\in\mathscr{E}_b(\mathbb{R}^d)$ and $\alpha\in\mathbb{R}^d_+$ be fixed. Set: $\tilde X_s = G_s(X_s)$ and $\tilde\alpha_s = G_s(\alpha)$, G_s being defined in (3.31), $H(\xi) = \mathbf{1}_{\xi\ge 0}$, $\xi\in\mathbb{R}$, $\widehat\sigma$ as in (3.32) and*

$$\Delta W^i_{s,t} = (t-s)(W^i_s + \sigma_{ii}s) - s(W^i_t - W^i_s),\qquad i=1,\dots,d. \tag{3.33}$$

Then, the following representation formula for the conditional expectation holds:

$$\mathbb{E}(\Phi(X_t)\,|\,X_s=\alpha)=\frac{\mathbb{T}_{s,t}[\Phi](\alpha)}{\mathbb{T}_{s,t}[1](\alpha)}$$

where

$$\mathbb{T}_{s,t}[f](\alpha)=\mathbb{E}\Big(f(X_t)\prod_{i=1}^d\frac{H(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\,\Delta W^i_{s,t}\Big).\tag{3.34}$$

Proof. Let us set $\tilde\Phi_t(y)\equiv\tilde\Phi(y)=\Phi\circ F_t(y)$, $y\in\mathbb{R}^d_+$, being F_t defined in (3.31). Since $X_t=F_t(\tilde X_t)$ for any t, one obviously has

$$\mathbb{E}(\Phi(X_t)\,|\,X_s=\alpha)=\mathbb{E}(\tilde\Phi(\tilde X_t)\,|\,\tilde X_s=G_s(\alpha)),$$

(recall that $G_s=F_s^{-1}$). Thus, setting $\tilde\alpha_s=G_s(\alpha)$, it is sufficient to prove that

$$\mathbb{E}(\tilde\Phi(\tilde X_t)\,|\,\tilde X_s=\tilde\alpha_s)=\frac{\tilde{\mathbb{T}}_{s,t}[\tilde\Phi](\tilde\alpha)}{\tilde{\mathbb{T}}_{s,t}[1](\tilde\alpha)}\tag{3.35}$$

where

$$\tilde{\mathbb{T}}_{s,t}[f](\tilde\alpha)=\mathbb{E}\Big(f(\tilde X_t)\prod_{i=1}^d\frac{H(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\,\Delta W^i_{s,t}\Big)$$

Let us firstly suppose that $\tilde\Phi(y)=\tilde\Phi_1(y_1)\cdots\tilde\Phi_d(y_d)$, that is $\tilde\Phi$ can be separated in the product of d functions each one depending only on a single variable and belonging to $\mathscr{E}_b(\mathbb{R})$. In such a case, one obviously has

$$\mathbb{E}(\tilde\Phi(\tilde X_t)\,|\,\tilde X_s=\tilde\alpha_s)=\prod_{i=1}^d\mathbb{E}\Big(\tilde\Phi_i(\tilde X^i_t)\,\Big|\,\tilde X^i_s=\tilde\alpha^i_s\Big).$$

Now, let us consider $\mathbb{E}(\tilde\Phi_i(\tilde X^i_t)\,|\,\tilde X^i_s=\tilde\alpha^i_s)$, for any fixed $i=1,\dots,d$. Let $\{h_n\}_n$ be a sequence of C^∞ probability density functions on $\mathbb{R}$ weakly converging to the Dirac mass in 0 as $n\to\infty$. Then one can write

$$\mathbb{E}(\tilde\Phi_i(\tilde X^i_t)\,|\,\tilde X^i_s=\tilde\alpha^i_s)=\lim_{n\to\infty}\frac{\mathbb{E}(\tilde\Phi_i(\tilde X^i_t)\,h_n(\tilde X^i_s-\tilde\alpha^i_s))}{\mathbb{E}(h_n(\tilde X^i_s-\tilde\alpha^i_s)).}$$

Setting H_n the probability distribution function associated with h_n, we have to handle something like $\mathbb{E}(f(\tilde X^i_t)\,H_n'(\tilde X^i_s-\tilde\alpha^i_s))$. Since the process $\tilde X^i$ is of the same type studied in Lemma 3.3.2, we can apply it:

$$\mathbb{E}(f(\tilde X^i_t)\,H_n'(\tilde X^i_s-\tilde\alpha^i_s))=\mathbb{E}\Big(f(\tilde X^i_t)\,\frac{H_n(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\Delta W^i_{s,t}\Big),$$

where $\Delta W^i_{s,t}=(t-s)(W^i_s+\sigma_{ii}s)-(t-s)(W^i_t-W^i_s)$. By using the Lebesgue dominated convergence theorem, one has

$$\begin{aligned}\mathbb{E}(\tilde\Phi_i(\tilde X^i_t)\,|\,\tilde X^i_s=\tilde\alpha^i_s)=&\lim_{n\to\infty}\frac{\mathbb{E}\Big(\tilde\Phi_i(\tilde X^i_t)\,\frac{H_n(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\Delta W^i_{s,t}\Big)}{\mathbb{E}\Big(\frac{H_n(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\Delta W^i_{s,t}\Big)}\\ =&\frac{\mathbb{E}\Big(\tilde\Phi_i(\tilde X^i_t)\,\frac{H(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\Delta W^i_{s,t}\Big)}{\mathbb{E}\Big(\frac{H(\tilde X^i_s-\tilde\alpha^i_s)}{\sigma_{ii}s(t-s)\tilde X^i_s}\Delta W^i_{s,t}\Big),}\end{aligned}$$

where $H(\xi) = \lim_{\delta\to 0} H_\delta(\xi) = \mathbf{1}_{\xi\geq 0}$. Therefore,

$$\mathbb{E}(\tilde\Phi(\tilde X_t^i)\,|\,\tilde X_s = \tilde\alpha_s) = \prod_{i=1}^d \mathbb{E}(\tilde\Phi_i(\tilde X_t^i)\,|\,\tilde X_s^i = \tilde\alpha_s^i) = \frac{\tilde{\mathbb{T}}_{s,t}[\tilde\Phi](\tilde\alpha)}{\tilde{\mathbb{T}}_{s,t}[1](\tilde\alpha)}$$

so that (3.35) holds when $\tilde\Phi(y) = \tilde\Phi_1(y_1)\cdots\tilde\Phi_d(y_d)$. In the general case, the statement holds by using density arguments: for any $\tilde\Phi \in \mathscr{E}_b(\mathbb{R}^d)$ there exists a sequence of functions $\{\tilde\Phi^n\}_n \subset \mathscr{E}_b(\mathbb{R}^d)$ such that $\tilde\Phi^n(\tilde X_t) \to \tilde\Phi(\tilde X_t)$ in L^2 and such that each $\tilde\Phi^n$ is a linear combination of functions which separate the variables as above. Since representation (3.35) holds for any $\tilde\Phi^n$, it finally holds for $\tilde\Phi$ as well, as it immediately follows by passing to the limit. $\square$

3.3.2 Localized formulas

Let us now discuss formulas involving localization functions. If we restrict our attention to product-type localizing function, then we can first state a localized formula for the operators $\mathbb{T}_{s,t}[f](\alpha)$.
Set $\mathscr{L}_1 = \{\psi : \mathbb{R} \to [0,+\infty)\,;\, \psi \in C^1(\mathbb{R}),\ \psi(+\infty) = 0 \text{ and } \int_{\mathbb{R}} \psi(t)\,dt = 1\}$, and $\mathscr{L}_d = \{\psi : \mathbb{R}^d \to [0,+\infty)\,;\, \psi(x) = \prod_{i=1}^d \psi_i(x^i),\ \psi_i \in \mathscr{L}_1, \text{ for any } i\}$.
One has

Theorem 3.3.4. ***[Representation formulas II: with localization]*** *For any $0 \leq s < t$, $\Phi \in \mathscr{E}_b$, $\alpha \in \mathbb{R}^d_+$ and for any $\psi \in \mathscr{L}_d$, one has*

$$\mathbb{E}\Big(\Phi(X_t)\,\Big|\,X_s = \alpha\Big) = \frac{\mathbb{T}^\psi_{s,t}[\Phi](\alpha)}{\mathbb{T}^\psi_{s,t}[1](\alpha)}$$

where

$$\mathbb{T}^\psi_{s,t}[f](\alpha) = \mathbb{E}\Big(f(X_t)\prod_{i=1}^d \Big[\psi_i(X_s-\alpha)) + \frac{(H-\Psi_i)(\tilde X_s^i - \tilde\alpha_s^i)}{\sigma_{ii}s(t-s)\tilde X_s^i}\,\Delta W^i_{s,t}\Big]\Big) \quad (3.36)$$

where Ψ_i denotes the probability distribution function associated with ψ_i: $\Psi_i(y) = \int_{-\infty}^y \psi_i(\xi)d\xi$.

Proof. The proof follows from the elementary fact: in dimension $d = 1$, Lemma 3.3.2 gives

$$\begin{aligned}\mathbb{E}(f(X_t)\,g'(X_s-\alpha)) = {} & \mathbb{E}(f(X_t)\,(g-\Psi)'(X_s-\alpha)) + \mathbb{E}(f(X_t)\,\psi(X_s-\alpha)) \\ = {} & \mathbb{E}\Big(f(X_t)(g-\Psi)(X_s-\alpha)\frac{\Delta W_{s,t}}{\sigma s(t-s)X_s}\Big) \\ & +\mathbb{E}(f(X_t)\,\psi(X_s-\alpha))\end{aligned}$$

so that

$$\mathbb{E}(f(X_t)\,g'(X_s-\alpha)) = \mathbb{E}\Big(f(X_t)\Big[\psi(X_s-\alpha) + (g-\Psi)(X_s-\alpha)\frac{\Delta W_{s,t}}{\sigma s(t-s)X_s}\Big]\Big)$$

Now, by using this equality, the proof of Theorem 3.3.3 can be repeated and the statement holds. $\square$

Remark 3.3.5. *Notice that, in principle, one could take different localizing functions for each operator, that is:*

$$\mathbb{E}\Big(\Phi(X_t)\,\Big|\,X_s=\alpha\Big)=\frac{\mathbb{T}^{\psi_1}_{s,t}[\Phi](\alpha)}{\mathbb{T}^{\psi_2}_{s,t}[1](\alpha)}$$

Let us add some more details about the localizing function.
First, one has to consider it because in practice (for example to price American contingent claims) the non localized formula does not work (in fact, the pricing algorithm blows up).
Then, the question is: how to choose it? Let us give a discussion about this. The proofs of the following statements are omitted and can be found in Bally, Caramellino and Zanette [3].
Let us start from the result in Theorem 3.3.4: to compute $\mathbb{E}(\Phi(X_t)\,|\,X_s=\alpha)$ one has to evaluate

$$\mathbb{T}^{\psi}_{s,t}[f](\alpha)=\mathbb{E}\Big(f(X_t)\prod_{i=1}^{d}\Big[\psi_i(\tilde{X}^i_s-\tilde{\alpha}^i_s)+\frac{(H-\Psi_i)(\tilde{X}^i_s-\tilde{\alpha}^i_s)}{\sigma_{ii}s(t-s)\tilde{X}^i_s}\,\Delta W^i_{s,t}\Big]\Big),$$

with $f=\Phi$ and $f=1$. Such an expectation is evaluated in practice through the empirical mean from many replications. The aim is now to choose the localizing function ψ allowing to reduce the variance. To this purpose, one can follow the optimization criterium introduced in the one-dimensional case by Kohatsu-Higa and Petterson [11]. It deals in looking for the localizing function ψ which minimizes the integrated variance, given by

$$I^f_d(\psi)=\int_{\mathbb{R}^d}\mathbb{E}\Big(f^2(X_t)\prod_{i=1}^{d}\Big[\psi_i(\tilde{X}^i_s-\tilde{\alpha}^i)+\frac{(H-\Psi_i)(\tilde{X}^i_s-\tilde{\alpha}^i)}{\sigma_{ii}s(t-s)\tilde{X}^i_s}\,\Delta W^i_{s,t}\Big]^2\Big)\,d\tilde{\alpha}, \tag{3.37}$$

up to the constant (with respect to the localizing function ψ) term coming out from $\mathbb{T}^{\psi}_{s,t}[f](\alpha)=\mathbb{T}_{s,t}[f](\alpha)$. Then the following result holds:

Proposition 3.3.6. *Set* $\mathscr{L}_1=\{\psi:\mathbb{R}\to[0,+\infty)\,;\,\psi\in C^1(\mathbb{R}),\ \psi(+\infty)=0$ *and* $\int_{\mathbb{R}}\psi(t)\,dt=1\}$, *and* $\mathscr{L}_d=\{\psi:\mathbb{R}^d\to[0,+\infty)\,;\,\psi(x)=\prod_{i=1}^d\psi_i(x^i)$, $\psi_i\in\mathscr{L}_1$, *for any* $i\}$. *Then*

$$\inf_{\psi\in\mathscr{L}_d}I^f_d(\psi)=I^f_d(\psi^*)$$

where $\psi^*(x)=\prod_{j=1}^d\psi^*_j(x^j)$, *with* $\psi^*_j(\xi)=\lambda^*_j\,e^{-\lambda^*_j\,|\xi|}/2$ *is a Laplace probability density function on* $\mathbb{R}$ *and* $\lambda^*_j=\lambda^*_j[f]$ *enjoys the following system of nonlinear equations:*

$$\lambda^{*2}_j=\frac{\mathbb{E}\Big(f^2(X_t)\,\Theta^2_{s,t;j}\,\prod_{i\,:\,i\neq j}\Big[\lambda^{*2}_i+\Theta^2_{s,t;i}\Big]\Big)}{\mathbb{E}\Big(f^2(X_t)\,\prod_{i\,:\,i\neq j}\Big[\lambda^{*2}_i+\Theta^2_{s,t;i}\Big]\Big)},\quad j=1,\dots,d, \tag{3.38}$$

where $\Theta_{s,t;i}=\Delta W^i_{s,t}/(\sigma_{ii}s(t-s)\tilde{X}^i_s)$, $i=1,\dots,d$.

In the case $f=1$, the optimal values can be explicitly written:

Corollary 3.3.7. *One has*

$$\lambda_j^*[1] = \frac{e^{-h^j s + \sigma_{jj}^2 s}}{x^j} \sqrt{\frac{t + \sigma_{jj}^2 s(t-s)}{\sigma_{jj}^2 s(t-s)}}, \quad j = 1, \dots, d.$$

For practical purposes, numerical evidence shows that the choice $\lambda^* = 1/\sqrt{t-s}$ works good enough, thus avoiding to weight the algorithm with the computation of further expectations. When $f = 1$, this kind of behavior is clear from Corollary 3.3.7. In the general case, the theoretical justification is given by the following

Proposition 3.3.8. *For any $j = 1, \dots, d$, one has $\lambda_j^*[f] = O(1/\sqrt{t-s})$ as $t \to s$. Moreover, if f is continuous, then*

$$\lim_{\sigma \to 0} \lim_{t \to s} \frac{\lambda_j^*[f]}{\lambda_j^*[1]} = 1.$$

Bibliography

[1] V. Bally: *An elementary introduction to Malliavin calculus*. Rapport de recherche 4718. INRIA, 2003.

[2] V. Bally, M.P. Bavouzet, M. Messaoud: Integration by parts formula for locally smooth laws and applications to sensitivity computations. *Annals of Applied Probability*, **17**, 33-66, 2007.

[3] V. Bally, L. Caramellino, A. Zanette: Pricing and Hedging American Options by Monte Carlo methods using a Malliavin calculus approach. *Monte Carlo Methods and Applications*, **11**, 121-137, 2005.

[4] M.P. Bavouzet-Morel, M. Messaoud: Computation of Greeks uning Malliavin's calculus in jump type market models. *Electronic Journal of Probability*, **11**, 276-300, 2006.

[5] K. Bichteler, J.-B. Gravereaux, J. Jacod. *Malliavin calculus for processes with jumps*. Gordon and Breach Science Publishers, New York, 1987.

[6] B. Bouchard, I. Ekeland, N. Touzi: On the Malliavin Approach to Monte Carlo Approximation of Conditional Expectations. *Finance and Stochastics*, **8**, 45-71, 2004.

[7] N. Chen, P. Glasserman. Malliavin Greeks without Malliavin calculus. *Stochastic Processes and their Applications*, **117**, 1689-1723, 2007.

[8] E. Fournié, J.M. Lasry, J. Lebouchoux, P.-L. Lions, N. Touzi: Applications of Malliavin Calculus to Monte Carlo methods in finance. *Finance and Stochastics*, **3**, 391-412, 1999.

[9] E. Fournié, J.M. Lasry, J. Lebouchoux, P.-L. Lions: Applications of Malliavin Calculus to Monte Carlo methods in finance II. *Finance and Stochastics*, **5**, 201-236, 2001.

[10] P.E. Kloeden, E. Platen: *Numerical Solutions of Stochastic Differential Equations*. Applications of Mathematics, Stochastic Modeling and Applied Probability 23, Springer, 1991.

[11] A. Kohatsu-Higa, R. Petterson: Variance Reduction Methods for Simulation of Densities on Wiener Space. *SIAM Journal of Numerical Analysis*, **4**, 431-450, 2002.

[12] S. Kusuoka, D. Strook: Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **32**, 1–76, 1985.

[13] N. Ikeda, S. Watanabe: *Stochastic differential equations and diffusion processes.* North Holland, second ed. 1989.

[14] D. Lamberton, B. Lapeyre. *Introduction to stochastic calculus applied to finance.* Chapman & Hall, London, 1996.

[15] P-L. Lions, H. Regnier: Calcul du Prix et des Sensibilités d'une Option Américaine par une Méthode de Monte Carlo. Preprint, 2000.

[16] P. Malliavin: *Stochastic analysis.* Springer, 1997.

[17] P. Malliavin, A. Thalmaier: *Stochastic calculus of variations in mathematical finance.* Springer-Verlag, Berlin, 2006.

[18] D. Nualart: *The Malliavin calculus and related topics.* Springer-Verlag, 1995.

[19] M. Sanz-Solé: *Malliavin calculus, with applications to stochastic partial differential equations.* EPFL Press, 2005.