Introduction to Malliavin Calculus

Vlad Bally

26.03.2007

0.1. Introduction

We give a short introduction to Malliavin calculus which finishes with the proof of Hörmandr's theorem (which was the outstanding result obtained by P. Malliavin in his initial papers). We will not give all the complete proofs because this turns out to be rather heavy. In particular we shall skip the proofs of the convergence theorems which are used in order to check that a functional is smooth in Malliavin sense We will also skip proofs for the evaluation of the Sobolev norms on the Wiener space. But we will discuss rather thoroughly the non degeneracy conditions.

The paper is organized as follows. In the first section we introduce the simple functionals and the simple processes and we define the Malliaivin derivative and the Skorohod integral for these finite dimensional objects. Then we derive the duality formula. This permits to prove that the operators are closable and to extend them to the infinite dimensional space. Then we prove the integration by parts formula and the basic calculus rules. We give the application of this formula to the analysis of the density of the law of a random variable. And we also prove the Clark Ocone formula. So the line of this presentation si to consider first a finite dimensional setting and to define there the differential operators coming on in the Malliavin calculus and then to extend the calculus to functionals on the Wiener space. In the second section we give the expression of the differential operators in terms of the Wiener chaos decomposition and we precise the domains of the operators. In fact we keep to a rather basic level and we give just a minimum of results in order to understand the connection with the decomposition in Wiener chaos. But in order to get a real idea about what analysis on the Wiener space is one has to look in some other texts as Nualart or Shigekawa (see also M. Santz for a short and simple presentation of the Malliavin calculus in which the Wiener chaos decomposition represents the starting point). Finally in the third section we discuss the case of diffusion processes. We present first the elliptic case and then we give the result under Hörmander condition.

0.2. Basic notions

0.2.1. The Malliavin derivative and the Skorohod integral in the finite dimensional framework

On a probability space (Ω, F, P) we consider a d dimensional Brownian motion $W = (W^1, ..., W^d)$ and we denote by $F_t, t \geq 0$ the filtration associated to W that is $F_t = \sigma(W_s, s \leq t)$. We work with $t \in [0, 1]$. We will use the following notation. For each $n \in N$ we denote $t_n^k = k2^{-n}, k \in N$ and

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

We also denote $\Delta_n^k = (\Delta_n^{k,1}, ...\Delta_n^{k,d})$ and $\Delta_n = (\Delta_n^0, ..., \Delta_n^{2^n-1})$.

Simple Functionals. A simple functional of order n is a random variable of the form

$$F = f(\Delta_n), \quad f \in C_p^\infty(R^{d2^n}, R)$$

where $C_p^\infty(R^{d2^n}, R)$ designs the space of the infinitely differentiable functions which have polynomial growth and all their derivatives have also polynomial growth. We denote by S_n the space of the simple functionals of order n and by $S = \cup_{n\in N} S_n$ the space of the "simple functionals". Notice that $S_n \subset S_{n+1}$ and S is a linear subspace of $L^2(\Omega, F_1, P)$ which is **dense** in $L^2(\Omega, F_1, P)$ (here comes on the fact that $F_1 = \sigma(W_s, s \leq 1)$).

Simple Processes. A process $U : [0, 1] \times \Omega \to R$ is called a simple process of order n if

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

We denote by P_n the space of the simple processes of order n and $P = \cup_n P_n$. Notice that $P_n \subset P_{n+1}$..

Let $U \in P^d$. For each fixed $\omega \in \Omega$, $t \to U_t$ is an element of $L^2([0, 1], B([0, 1]), dt : R^d) =: H_d$. The scalar product on this space is given by

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

for $U = (U^1, ..., U^d), V = (V^1, ..., V^d)$. We will also denote

$$L^p(H_d) = \{U : \Omega \to H_d : E \|U\|_{H_d}^p = E(\int_0^1 \sum_{i=1}^d |U_s^i|^2 ds)^{p/2} < \infty\}.$$

Notice that for each $p \in N$, P^d is a linear subspace of $L^p(H_d)$ which is **dense** in $L^p(H_d)$ Notice also that a simple process is generally not adapted.

The Malliavin derivatives. $D : S_n \to P_n^d$. Let $F = f(\Delta_n) \in S_n$. We define

$$D_s^i F = \sum_{k=0}^{2^n-1} \frac{\partial f}{\partial x^{k,i}}(\Delta_n) 1_{[t_n^k, t_n^{k+1})}(s), \quad DF = (D^1 F, ..., D^d F).$$

A more intuitive notation: we denote

$$\Delta_n^{i,s} = \Delta_n^{k,i} \quad for \quad s \in [t_n^k, t_n^{k+1}).$$

Then

$$D_s^i F = \frac{\partial F}{\partial \Delta_n^{i,s}}(\Delta_n).$$

The operator D **does not depend on** n : for $F \in S_n \subset S_{n+1}$ we have

$$\frac{\partial F}{\partial \Delta_n^{i,s}}(\Delta_n) = \frac{\partial F}{\partial \Delta_{n+1}^{i,s}}(\Delta_{n+1}).$$

This permits to define (**well**) $D : S = \cup_n S_n \to P^d = \cup_n P_n^d$ by

$$D_s^i F = \frac{\partial F}{\partial \Delta_n^{i,s}}(\Delta_n) \quad for \quad F \in S_n.$$

The Skorohod integral (divergence operator). $\delta_i : P_n^d \to S_n, i = 1, ..., d :$ let $U = \sum_{k=0}^{2^n-1} u_k(\Delta_n) 1_{[t_n^k, t_n^{k+1})}(\circ) \in P_n$. We define

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

If $U = (U^1, ..., U^d) \in P_n^d$ then we define

$$\delta(U) = \sum_{i=1}^{d} \delta_i(U^\iota)$$

so $\delta : P^d \to S$. Finally we check that the definition of δ_i does not depend on n and we extend this operator to P. Now $\delta : P^d \to S$.

Remark 1. *Suppose that $U \in P_n$ is adapted. It follows that $u_k(\Delta_n)$ does not depend on Δ_n^k and consequently $\partial_{x^{k,i}} u_k = 0$. It follows that $\delta_i(U) = \sum_{k=0}^{2^n-1} u_k(\Delta_n)\Delta_n^{k,i} = \int_0^1 U_s dW_s^i$. So for adapted processes the Skorohod integral coincides with the Itô integral. For this reason we will use the notation*

$$\delta_i(U) = \int_0^1 U_s dW_s^i.$$

But we keep in mind that generally $\int_0^1 U_s dW_s^i$ is not an Itô integral but the Skorohod integral (an anticipative integral).

Proposition 1. *(The Duality relation) Let $F \in S$ and $U \in P^d$. We have*

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

Proof. We assume (without loss of generality) that $F \in S_n$ and $U \in P_n^d$ with the same n. We write

$$E(\langle DF, U\rangle) = \sum_{k=0}^{2^n-1}\sum_{i=1}^{d} E(\partial_{x^{k,i}} f(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i}) u_k^i(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i})\frac{1}{2^n}$$

where $\overline{\Delta}_n^{k,i}$ designs all the random variables involved in Δ_n, except $\Delta_n^{k,i}$. In particular $\overline{\Delta}_n^{k,i}, \Delta_n^{k,i}$ are independent and $\Delta_n^{k,i} \sim N(0,h)$ with $h = 2^{-n}$ so we obtain

$$\begin{aligned}
& E(\partial_{x^{k,i}} f(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i}) u_k^i(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i})) \\
= & \frac{1}{\sqrt{2\pi h}} E(\int_R \partial_{x^{k,i}} f(\overline{\Delta}_n^{k,i}, y) u_k^i(\overline{\Delta}_n^{k,i}, y) e^{-y^2/2h} dy) \\
= & -\frac{1}{\sqrt{2\pi h}} E(\int_R f(\overline{\Delta}_n^{k,i}, y)(\partial_{x^{k,i}} u_k^i(\overline{\Delta}_n^{k,i}, y) - \frac{y}{h} u_k^i(\overline{\Delta}_n^{k,i}, y)) e^{-y^2/2h} dy) \\
= & E(\partial_{x^{k,i}} f(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i})(u_k^i(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i})\frac{\Delta_n^{k,i}}{h} - \partial_{x^{k,i}} u_k^i(\overline{\Delta}_n^{k,i}, \Delta_n^{k,i}))).
\end{aligned}$$

We come back and we obtain the formula. □

0.2.2. Extension to the infinite dimensional framework.

We have

$$\begin{aligned}
D &: S \subset L^2(R) \to P^d \subset L^2(H_d), \\
\delta &: P^d \subset L^2(H_d) \to S \subset L^2(R)
\end{aligned}$$

and we want to "extend" these operators. These are **unbounded operators**. This means that one may not find a constant C such that

$$\|DF\|_{L^2(H_d)}^2 = E(\int_0^1 \sum_{i=1}^d \left|D_s^i F\right|^2 ds) \leq C \|F\|_2^2 = CE|F|^2, \quad \forall F \in S.$$

EX. Give an example of sequence $F_n \in S, n \in N$ such that $\sup_n \|F_n\|_2 < \infty$ and $\sup_n \|DF_n\|_{L^2(H_d)} = \infty$.

Then in order to be able to obtain a "well defined" extension of these operators we have to check that they are "closable", which means the following. Consider a sequence $F_n \in S, n \in N$ such that $\lim_n F_n = 0$ in $L^2(R)$ and $\lim_n DF_n = G$ in $L^2(H_d)$. Then $G = 0$.

Lemma 2. *D and δ are closable.*

Proof. We consider the sequence $F_n, n \in N$ as above and we want to prove that $G = 0$. We take $U \in P^d$ and we use the duality relation in order to obtain

$$E(\langle G, U\rangle) = \lim_n E(\langle DF_n, U\rangle) = \lim_n E(F_n \delta(U)) = 0.$$

Since P^d is dense in $L^2(H_d)$ we conclude that $G = 0$.□

EX. Prove the δ is closable.

Definition 3. *We say that $F \in DomD$ if there exists a sequence $F_n \in S, n \in N$ such that*

$$\lim_n F_n = F \text{ in } L^2(R), \quad \lim_n DF_n = G \text{ in } L^2(H_d)$$

for some $G \in L^2(H_d)$. In this case we define $DF = G = \lim_n DF_n$.

The definition does not depend on the sequence $F_n, n \in N$ because D is closable.

Remark 2. *Notice that we may just assume that $\lim_n DF_n = G$ in $L^2(H_d)$ weakly (that is $\lim_n E(\langle DF_n, U\rangle) = E(\langle G, U\rangle)$ for every $U \in L^2(H_d)$. In order to prove it we use Mazur's theorem: we pass to a subsequence, and then, for each $n \in N$ there exist $\lambda_k^n \geq 0, k = n, ..., k_n$ such that $\sum_{k=n}^{k_n} \lambda_k^n = 1$ and $G_n := \sum_{k=n}^{k_n} \lambda_k^n DF_k \to G$ strongly. Notice that $G_n = D\overline{F}_n$ with $\overline{F}_n = \sum_{k=n}^{k_n} \lambda_k^n F_k$, which is still a simple functional. And $\left\|F - \overline{F}_n\right\|_2 \leq \sum_{k=n}^{k_n} \lambda_k^n \|F - F_k\|_2 \leq \sup_{k \geq n} \|F - F_k\|_2 \to 0$.*

Notice that $F, F' \in L^2(R) \nRightarrow F \times F' \in L^2(R)$ and consequently $DomD$ is not closed to multiplication, and then it is not an algebra. This is unpleasant and so we define:

Definition 4. *Let $p \in N$. We say that $F \in Dom_pD$ if there exists a sequence $F_n \in S, n \in N$ such that*

$$\lim_n F_n = F \text{ in } L^p(R), \quad \lim_n DF_n = G \text{ in } L^p(H_d)$$

for some $G \in L^p(H_d)$. In this case we define $DF = G = \lim_n DF_n$.

So $DomD = Dom_2D$. And since $\|\circ\|_{p'} \leq \|\circ\|_p$ and $\|\circ\|_{L^{p'}(H_d)} \leq \|\circ\|_{L^p(H_d)}$ for $p \geq p'$, we have $Dom_pD \subset Dom_{p'}D$. It is easy to see that the definition of DF does not depend on p. We put $Dom_\infty D = \cap_{p\in N} Dom_pD$ and we then $Dom_\infty D$ is an algebra.

We would like to see Dom_pD as a normed space so we define

$$\|F\|_{1,p} = \|F\|_p + \|DF\|_{L^p(H_d)}.$$

Then it is clear that

$$Dom_pD = \overline{S}^{\|\circ\|_{1,p}}.$$

Notice that the significance of "$F \in \overline{S}^{\|\circ\|_{1,p}}$" is the following: there exist $F_n \in S, n \in N$ such that $F_n \to F$ in $L^p(R)$ and $(F_n)_{n\in N}$ is a Cauchy sequence in $\|\circ\|_{1,p}$. Then the above equality is obvious because $L^p(H_d)$ is a complete space. It is also obvious that Dom_pD is complete. Indeed consider a Cauchy sequence $(F_n)_{n\in N}$ with respect to $\|\circ\|_{1,p}$. In particular this sequence is Cauchy with respect to $\|\circ\|_p$ and so there exists $F \in L^p(R)$ such that $F_n \to F$ in $\|\circ\|_p$. Moreover, since $F_n \in Dom_pD$ we may find a sequence of simple functionals F'_n such that $\|F_n - F'_n\|_{1,p} \leq \frac{1}{n}$ so that $(F'_n)_{n\in N}$ is Cauchy with respect to $\|\circ\|_{1,p}$ and $F'_n \to F$ in $\|\circ\|_p$. So $F \in Dom_pD$.

In particular, for $p = 2$ the above norm is given by the scalar product

$$\langle F, G\rangle_{1,2} = E(FG) + E(\int_0^1 \sum_{i=1}^d D_s^i F \times D_s^i G ds)$$

so that Dom_2D is a Hilbert space. We denote

$$D^{1,p} = Dom_pD = \overline{S}^{\|\circ\|_{1,p}}, \quad D^{1,\infty} = \cap_{p\in N} D^{1,p}.$$

Sometimes it is unpleasant to compute $\lim_n DF_n = G$. So we give the following criterion.

Proposition 5. *Let $F \in L^2(R)$. Suppose that there exists a sequence $F_n \in DomD$ (in particular $F_n \in S$) such that $\lim_n F_n = F$ in $L^2(R)$ and $\sup_n \|F_n\|_{1,2} \leq C < \infty$. Then $F \in DomD$ and $\|F\|_{1,2} \leq C$. Moreover, if $\sup_n \|F_n\|_{1,p} \leq C_p$ then $\|F\|_{1,p} \leq C_p$.*

Proof. Any bounded set in a Hilbert space is relatively compact so we may find $F' \in D^{1,2}$ such that $F_n \to F'$ weakly (we pass to a subsequence). We use now Mazur's theorem: one may construct convex combinations of the functionals F_n in order to obtain "strong convergence" instead of week convergence. Let us be more precise: for each $n \in N$ there exist $\lambda_k^n \geq 0, k = n, ..., k_n$ such that $\sum_{k=n}^{k_n} \lambda_k^n = 1$ and $\overline{F}_n := \sum_{k=n}^{k_n} \lambda_k^n F_k \to F'$ strongly (with respect to $\|\circ\|_{1,2}$). In particular $\overline{F}_n \to F'$ in $L^2(R)$. Notice that

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

It follows that $F' = F$ and so $F \in D^{1,2}$.

We also have

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

Let us now prove the assertion concerning the p norm. Passing to a subsequence we may assume that $(\left|\overline{F}_n - F\right|, \left|D\overline{F}_n - DF\right|_{H_d}^2) \to (0,0)$ almost surely. Since $\sup_n \left\|\overline{F}_n\right\|_{1,p} \leq C_p$ we may use unifromly-integrability in order to derive $\overline{F}_n \to F$ in $\|\circ\|_{1,p'}$ for $p' < p$. Then $\|F\|_{1,p'} \leq \sup_n \left\|\overline{F}_n\right\|_{1,p'} \leq \sup_n \left\|\overline{F}_n\right\|_{1,p} \leq C_p$. And finally $\|F\|_{1,p} \leq \sup_{p'<p} \left\|\overline{F}\right\|_{1,p'} \leq C_p$.$\square$

Now we define $Dom_p\delta$ in a similar way:

Definition 6. *Let $p \in N$. We say that $U \in Dom_p\delta$ if there exists a sequence $U_n \in P^d, n \in N$ such that*

$$\lim_n U_n = U \text{ in } L^p(H_d), \quad \lim_n \delta(U_n) = F \text{ in } L^p(R)$$

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

On P^d we consider the norm

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

and we have

$$Dom_p\delta = \overline{P^d}^{\|\circ\|_{\delta,p}}.$$

Remark 3. *Let $p = 2$. We may replace the strong convergence $\lim_n \delta(U_n) = F$ in $L^2(R)$ by week convergence (the same reasoning based on Mazur's theorem as for the derivatives).*

And a similar argument as above gives

Proposition 7. *Let $U \in L^2(H_d)$. Suppose that there exists a sequence $U_n \in Dom_2\delta$ (in particular $U_n \in P^d$) such that $\lim_n U_n = U$ in $L^2(H_d)$ and $\sup_n \|U_n\|_{\delta,2} \leq C < \infty$. Then $U \in Dom_2\delta$ and $\|U\|_{\delta,2} \leq C$. Moreover, if $\sup_n \|U_n\|_{\delta,p} \leq C_p$ then $\|U\|_{\delta,p} \leq C_p$.*

Proof: EX.

By passage to the limit we readily obtain the extension of the duality relation:

Lemma 8. *Let $F \in DomD$ and $U \in Dom_2\delta$. Then*

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

We give now an alternative characterization of the domain of δ.

Proposition 9. *Let $U \in L^2(H_d)$. Then $U \in Dom_2(\delta)$ if and only if*

$$a) \quad |E(\langle DF, U\rangle)| \leq C\,\|F\|_2\,, \quad \forall F \in D^{1,2}.$$

In this case

$$b) \quad \|\delta(U)\|_2 = \sup_{\|F\|_2=1, F\in S} |E(\langle DF, U\rangle)|\,.$$

Proof. Suppose that $U \in Dom_2(\delta)$ and $F \in D^{1,2}$. Then using the duality relation we obtain $|E(\langle DF, U\rangle)| = |E(F\delta(U))| \leq \|\delta(U)\|_2 \|F\|_2$ so we have $a)$ with $C = \|\delta(U)\|_2$. Since S is dense in $L^2(R)$ we obtain

$$\sup_{\|F\|_2=1, F\in S} |E(\langle DF, U\rangle)| = \sup_{\|F\|_2=1, F\in S} |E(F\delta(U))| = \|\delta(U)\|_2\,.$$

Assume now that $a)$ holds true. Then using Riesz theorem we produce $G \in L^2(R)$ such that $E(\langle DF, U\rangle) = E(FG)$. We take now $U_n \in P^d$ such that $U_n \to U$ in $L^2(H_d)$ and consequently $E(\langle DF, U_n\rangle) \to E(\langle DF, U\rangle)$. It follows that $E(F\delta(U_n)) \to E(FG)$. And this is sufficient in order to obtain $U \in Dom_2\delta$ (see the remark after the definition of $Dom_2(\delta)$). $\square$

Finally we give the following computational rules:

Proposition 10. *(Chain rule) Let $F = (F_1, ..., F_m)$ with $F_i \in D^{1,2}$ and let $\phi \in C_b^1(R^m; R)$. Then $\phi(F) \in D^{1,2}$ and*

$$D\phi(F) = \sum_{k=1}^{m} \partial_k \phi(F) DF_k.$$

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

Proof. If $F_k \in S, k = 1, ..., m$ then $\phi(F) \in S$ and the standard chain rule gives the formula. And if $F_k \in D^{1,2}$ then we take $F_k^n \in S, n \in N$ such that $\|F_k^n - F_k\|_{1,2} \to 0$. Since ϕ has bounded derivatives we obtain $\|\phi(F^n) - \phi(F)\|_2 \to 0$. Next we write

$$\left\| D\phi(F^n) - \sum_{k=1}^{m} \partial_k \phi(F) DF_k \right\|_{L^2(H_d)} = \left\| \sum_{k=1}^{m} \partial_k \phi(F^n) DF_k^n - \sum_{k=1}^{m} \partial_k \phi(F) DF_k \right\|_{L^2(H_d)} \to 0.$$

This yields $\phi(F) \in Dom_2 D$ and $D\phi(F) = \sum_{k=1}^{m} \partial_k \phi(F) DF_k$. In order to prove the last convergence we have to treat terms of the form $a_n = E(|\partial_k \phi(F^n) - \partial\phi(F)|^2 \int_0^1 |D_s F_k|^2 ds)$ and $b_n = E(|\partial_k \phi(F^n)|^2 \int_0^1 |D_s(F_k - F_k^n)|^2 ds)$. Since $\partial_k \phi$ is bounded it is clear that $b_n \to 0$. In order to treat a_n we pass to a subsequence such that $F^n \to F$ almost surely and then we use Lebesgue's Theorem.

In the above argument the fact that $\partial_k \phi$ is bounded comes on crucially because we are not able to use Holder inequalities. But if $F_i \in D^{1,\infty}$ then we may use Holder's inequality in order to get the same conclusion in the case when $\partial_k \phi$ has polynomial growth. $\square$

Proposition 11. *Let $U \in Dom_\infty \delta$ and $F \in D^{1,\infty}$. Then $F \times U \in Dom_\infty \delta$ and*

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

Proof. EX. You test against $G \in S$ and you use the duality relation and the identity $D(FG) = FDG + GDF$.$\square$

Finally we introduce the Ornstein Uhlenbeck operator L. We define $L : S \to S$ as $L = -\delta D$. For every $F, G \in S$ we have the duality relation:

$$E(FLG) = -E(\langle DF, DG \rangle) = E(GLF).$$

Then the same arguments as above ensure that L is closable and we may extend L :

Definition 12. *$F \in DomL$ if there exists a sequence of simple functionals $F_n, n \in N$ such that $F_n \to F$ in $L^2(R)$ and $LF_n \to G$ in $L^2(R)$ for some $G \in L^2(R)$. 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 $Dom_\infty L = \cap_{p \geq 2} Dom_p L$.*

We may define on S the norm $\|F\|_{L,p}^p = \|F\|_p^p + \|LF\|_p^p$ and then $Dom_p L = \overline{S}^{\|\circ\|_{L,p}}$.

And we may relax the convergence assumption on $LF_n, n \in N$ and replace it by the assumption that $(F_n)_n$ and $(LF_n)_n$ bounded in $L^2(R)$ (respectively in $L^p(R)$).

Proposition 13. *Let $F = (F_1, ..., F_m)$ with $F_k \in Dom_\infty L, k = 1, ..., m$. Then, for every $\phi \in C_p^2(R^m, R)$ we have $\phi(F) \in Dom_\infty L$ and*

$$L\phi(F) = \sum_{k=1}^{m} \partial_k \phi(F) LF_k + \sum_{k,p=1}^{m} \partial_k \partial_p \phi(F) \langle DF_k, DF_p \rangle .$$

Proof: EX.

0.2.3. Examples

EX 1. $F = W_t^i$ with $t \in (0,1)$. Then $D_s^i W_t^i = 1_{[0,t)}(s)$

EX 2. $\phi \in L^2(0,1)$. Then $W^i(\phi) := \int_0^1 \phi_r dW_r^i \in D^{1,2}$ and $D_s^i W^i(\phi) = \phi(s)$ and $D_s^j W^i(\phi) = 0$ for $j \neq i$.

If ϕ is a step function on dyadic intervals then $W^i(\phi)$ is a simple functional and we compute directly the derivative. If $\phi \in L^2(0,1)$ is a general function we approximate it as follows: we denote by G_n the σ algebra generated by the dyadic intervals of order n. And we take ϕ_n to be the conditional expectation of ϕ with respect to G_n (with respect to the Lebesgue measure on $(0,1)$). Then $\phi_n \to \phi$ in $L^2(0,1)$ and we pass to the limit in order to obtain the formula.

EX 3. We denote $W(\phi) = (W^1(\phi), ..., W^d(\phi))$. We take $\phi_1, ..., \phi_m \in L^2(0,1)$ and $F = \Phi(W(\phi_1), ..., W(\phi_m))$ with $\Phi \in C_p^1(R^{dm}, R)$. Then $F \in D^{1,2}$ and

$$D_s^i F = \sum_{k=1}^m \partial_{x^{k,i}} \Phi(W(\phi_1), ..., W(\phi_m))\phi_k(s).$$

This is an immediate consequence of EX 2 and of the chain rule. In many textbooks one considers directly $F = \Phi(W(\phi_1), ..., W(\phi_m))$ to be the simple functionals and defines the Malliaivn derivative by the above formula. In our case we have restricted ourselves to ϕ which is an indicator function of a dyadic interval. But finally we obtain the same thing.

EX 4. Let U be an adapted process such that $E\int_0^1 |U_r|^2 dr < \infty$ and $I_i(U) = \int_0^1 U_r dW_r^i, i = 1, ..., d$ and $I_0(U) = \int_0^1 U_r dr$. We assume that for each $r \in [0,1]$ we have $U_r \in D^{1,2}$ and we assume that

$$i) \quad \sup_{r \leq 1} \|U_r\|_{1,2} < \infty.$$

Moreover we denote $\tau_n(r) = k2^{-n}$ for $k2^{-n} \leq r < (k+1)2^{-n}$ and $U_s^n := U_{\tau_n(s)}$. We assume that

$$ii) \quad \int_0^1 \|U_r - U_r^n\|_{1,2}^2 dr = E\int_0^1 (\left|U_r - U_{\tau_n(r)}\right|^2 + \int_0^1 \left|D_s U_r - D_s U_{\tau_n(r)}\right|^2 ds) dr \to 0.$$

Then $I_i(U) \in D^{1,2}, i = 0, ..., d$ and

$$(*) \quad D_s^j I_i(U) = \delta_{ij} U_s + \int_s^1 D_s^j U_r dW_r^i, \quad j = 1, ..., d$$

with the convention that $dW_r^0 = dr$.

In order to prove this we approximate $I_i(U)$ by $I_i(U^n)$ with $U_r^n = U_{\tau_n(r)}$. By $ii)$ we have $I_i(U^n) \to I_i(U)$ in $L^2(R)$. We check that $(*)$ holds for $I_i(U^n)$. And using $i)$ we obtain $\sup_n \|I_i(U^n)\|_{1,2} < \infty$. Then we use our criterion in order to get $I_i(U) \in D^{1,2}$. If we know $ii)$ we are able to prove that $(*)$ holds for $I_i(U)$ by passing to the limit.

EX 5. We consider the diffusion process

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

We assume that $\sigma_j^i, b^i \in C_b^1(R^m : R)$. Then we claim that $X_t^i \in D^{1,2}$ and the derivative verifies

$$\begin{aligned} D_s^l X_t^i &= \sigma_l^i(X_s) + \sum_{j=1}^d \int_s^t \sum_{k=1}^m \partial_k \sigma_j^i(X_r) D_s^l X_r^k dW_r^j \\ &+ \int_s^t \sum_{k=1}^m \partial_k b^i(X_r) D_s^l X_r^k dr. \end{aligned}$$

This may be proved using two types of approximations: 1. Euler scheme and 2. Fix point argument. In both cases it is easy to check that $\sup_n \left\| X_t^{n,i} \right\|_{1,2} < \infty$ (here $X_t^{n,i}$ designs the $n'th$ approximation - Euler scheme or fix point approximation). This is an immediate consequence of Burkholder's inequality and of Gronwall's lemma. But this arguments does not guarantee that the continuity hypothesis $ii)$ holds true. So we are not able to derive the equation verified by the derivative (we may use "formal derivation" but this is not rigorous). So finally we have to check the convergence of $(X^n, DX^n), n \in N$.

Another important point concerns the **"variance of constants method"** which permits to give a nice expression for $D_s X_t$. It is convenient to develop this argument for $SDE's$ written in Stratonovich form. So we assume that the coefficients $\sigma_j^i, j = 1, ..., d$ are twice differentiable. Then the equation of X reads

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

where $\circ dW_s^j$ designees the Stratonovich integral and $\overline{b} = b - \sum_{j=1}^d \sigma_j \times \nabla \sigma_j$. We define

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

Then Y satisfies the following SDE written in matrix notation

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

where $\partial \sigma_j$ is the $m \times m$ matrix with components $\partial_k \sigma_j^i, k, i = 1, ..., m$. We also consider Z solution of the SDE

$$Z_t = I - \sum_{j=1}^d \int_0^t Z_s \partial \sigma_j(X_s) \circ dW_s^j - \int_0^t Z_s \partial \overline{b}(X_s) ds.$$

Using Itô's formula one checks that $d(Z_tY_t)=0$ so that $Z_t=Y_t^{-1}$. It follows that

$$D_sX_t=Y_tZ_s\sigma(X_s).$$

In order to prove this one writes

$$Y_t=Y_s+\sum_{j=1}^{d}\int_s^t\partial\sigma_j(X_r)Y_r\circ dW_r^j+\int_s^t\partial\bar{b}(X_r)Y_rdr$$

and multiplies with $Z_s\sigma(X_s)$ the above equation in order to obtain

$$\begin{aligned}Y_tZ_s\sigma(X_s) &= Y_sZ_s\sigma(X_s)+\sum_{j=1}^{d}\int_s^t\partial\sigma_j(X_r)Y_rZ_s\sigma(X_s)\circ dW_r^j\\ &\quad+\int_s^t\partial\bar{b}(X_r)Y_rZ_s\sigma(X_s)dr.\end{aligned}$$

Since $Y_sZ_s\sigma(X_s)=\sigma(X_s)$ this is the equation of D_sX_t and so, using the uniqueness of the solutions of $SED's$ we get the equality.

0.2.4. The integration by parts formula

An important consequence of the duality formula is the following integration by parts formula. In order to give the statement we have to define the Malliavin covariance matrix. Let $F=(F_1,...,F_m)$ with $F_i\in D^{1,2}$. We define

$$\sigma_F^{ij}=\langle DF_i,DF_j\rangle=\int_0^1\sum_{k=1}^{d}D_s^kF_i\times D_s^kF_jds.$$

This is a symmetric positive definite matrix.

Notice that if $F_i=\sum_{k=1}^{d}\int_0^1\phi_i^k(s)dW_s^k$ with $\phi_i^k\in L^2([0,1])$ then the vector F is Gaussian and σ_F coincides with the covariance matrix of F. EX.

We introduce now the following non-degeneracy assumption:

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

In the Gaussian case σ_F is deterministic so this just means that $\det\sigma_F\neq 0$. This is equivalent with the fact that σ_F is invertible and this is a necessary and sufficient condition in order that the law of F is absolutely continuous with respect to the Lebesgue measure on R^m - so F is "non degenerated".

If $(N-D)$ holds true then σ_F is almost surely invertible. We denote $\gamma_F=\sigma_F^{-1}$.

Theorem 14. *Let $F = (F_1, ..., F_m)$ with $F_i \in D^{1,\infty}$ and $G \in D^{1,\infty}$. Suppose also that $\sigma_F^{i,j} \in D^{1,\infty}$ and $D^j F_i \in \cap_{p \in N} Dom_p \delta, j = 1, ..., d$ and that $(N - D)$ holds for F. Then for every $\phi \in C_b^1(R^m, R)$ we have*

$$E(\partial_i \phi(F) G) = E(\phi(F) H_i(F, G)), \quad i = 1, ..., m$$

with

$$H_i(F, G) = \sum_{j=1}^{m} \delta(G \gamma_F^{i,j} DF_j) = -\sum_{j=1}^{m} G \gamma_F^{i,j} LF_j - \left\langle D(G \gamma_F^{i,j}), DF_j \right\rangle.$$

Proof. Using the chain rule

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

We multiply with $D_s F$ and we integrate with respect to $s \in [0, 1]$ and we obtain

$$\langle D\phi(F), DF \rangle_{H^d} = \sigma_F \nabla \phi(F)$$

which yields

$$\nabla \phi(F) = \langle D\phi(F), \gamma_F DF \rangle.$$

Then we use the duality formula in order to obtain

$$E(\nabla \phi(F) G) = E(\langle D\phi(F), G \gamma_F DF \rangle) = E(\phi(F) \delta(G \gamma_F DF)).$$

□

We also have the following easy generalisation. Consider $\xi_i = (\xi_i^1, ..., \xi_i^d) \in (Dom_\infty(\delta))^m, i = 1, ..., m$ and denote

$$\sigma_{F,\xi}^{ij} = \left\langle DF_i, \xi_j \right\rangle = \int_0^1 \sum_{k=1}^{d} D_s^k F_i \times \xi_j^k(s) ds.$$

Proposition 15. *Let $F = (F_1, ..., F_m)$ with $F_i \in D^{1,\infty}$ and let $\xi_i = (\xi_i^1, ..., \xi_i^d) \in (Dom_\infty(\delta))^m, i = 1, ..., m$ such that $\sigma_{F,\xi}^{ij} \in D^{1,\infty}, ij, = 1, ..., m$ and $E((|\det \sigma_{F,\xi}|^{-p}) < \infty, \forall p \in N$. Then for every $G \in D^{1,\infty}$ and $\phi \in C_p^1(R^m, R)$ we have*

$$E(\partial_i \phi(F) G) = E(\phi(F) H_i^\xi(F, G), \quad i = 1, ..., m$$

with

$$H_i^\xi(F, G) = \sum_{j=1}^{m} \delta(G \gamma_{F,\xi}^{i,j} DF_j) = -\sum_{j=1}^{m} G \gamma_{F,\xi}^{i,j} LF_j - \left\langle D(G \gamma_{F,\xi}^{i,j}), DF_j \right\rangle.$$

Proof: Exactly the same but we multiply with ξ_j instead of DF_j.□

0.2.5. The Clark Ocone formula

We recall the martingale representation formula: if F is a square integrable random variable which is $F_1 = \sigma(W_s, s \leq 1)$ measurable then there exists a previsible process $\phi = (\phi_1, ..., \phi_d)$ such that $E\int_0^1 |\phi_s|^2 ds < \infty$ and

$$F = EF + \sum_{j=1}^{d} \int_0^1 \phi_j(s) dW_s^j.$$

In the case when F is differentiable in Malliavin sense we may give a precise expression for ϕ_j. This is done by the Clark Ocone formula:

Theorem 16. *Let $F \in D^{1,2}$. Then*

$$F = EF + \sum_{j=1}^{d} \int_0^1 E(D_s^j F \mid F_s) dW_s^j.$$

Proof. We assume that $EF = 0$ (if not we work with $F - EF$). We take $U = (U_1, ..., U_d)$ a simple process which is adapted and we use the duality relation:

$$E(F\delta(U)) = E(\int_0^1 \sum_{j=1}^{d} D_s^j F \times U_j(s) ds) = \int_0^1 \sum_{j=1}^{d} E(U_j(s) E(D_s^j F \mid F_s)) ds.$$

Moreover, as U_j are adapted, the Skorohod integral coincides with the Itô integral so that $\delta(U) = \sum_{j=1}^{d} \delta_j(U_j) = \sum_{j=1}^{d} \int_0^1 U_j(s) dW_s^j$. Using the representation theorem we have $F = \sum_{j=1}^{d} \int_0^1 \phi_j(s) dW_s^j$ so we obtain

$$\begin{aligned} E(F\delta(U)) &= E((\sum_{j=1}^{d} \int_0^1 \phi_j(s) dW_s^j) \times (\sum_{j=1}^{d} \int_0^1 U_j(s) dW_s^j)) \\ &= E(\int_0^1 \sum_{j=1}^{d} U_j(s) \phi_j(s) ds). \end{aligned}$$

It follows that

$$E(\int_0^1 \sum_{j=1}^{d} U_j(s)(E(D_s^j F \mid F_s) - \phi_j(s)) ds) = 0.$$

Since the adapted simple processes are dense (in L^2 sense) in the adapted square integrable processes we conclude that $E(D_s^j F \mid F_s) = \phi_j(s)$ almost surely (with respect to $P \times ds$).$\square$

Consequences:

1. Let $F \in D^{1,2}$. Then $DF = 0 \Longleftrightarrow F = C$ (constant).

2. Let $A \in F_1$. Then $1_A \in D^{1,2} \Longleftrightarrow P(A) = 1$ or $P(A) = 0$. In particular $D^{1,2}$ is strictly included in L^2.

In order to prove 2. we use the chain rule and we get $D1_A = D1_A^2 = 2D1_A \times 1_A$. If $D1_A \neq 0$ then we may simplify and we get $1 = 2 \times 1_A$ which is not possible. So $D1_A = 0$ and we conclude that 1_A is a constant - so zero or one.

0.2.6. Bismut-Elworthy formula and sensitivity computations

Let us first present the problem of sensitivity computations - which is central in mathematical finance and is known as "greeks computation". Assume that you have a financial asset S_t with the dynamics given by the Black Scholes model:

$$dS_t = \sigma S_t dW_t, \quad S_0 = x$$

where W is an one dimensional Brownian motion. For simplicity we take the interest rate $r = 0$. The price of an option of maturity T and payoff ϕ on the asset S is given by

$$\Pi = E(\phi(S_T)).$$

For some special payoffs (as for example for call's and put's) one has an explicit expression for π (a closed form) but generally such a formula is not available and one has to compute Π using a Monte Carlo method. This is a first stage. But in the second stage one needs to compute the sensitivity of the price with respect to certain parameters as the initial value of S, that is $S_0 = x$ (this derivative is called "delta" of the option) or the second derivative with respect to x, or the derivative with respect to σ... Each of these derivatives is denoted by a greek letter and this is why we speak about computation of Greeks. Anyway what we want to compute is $\partial_x \Pi$. And there are several ways to do it - the simpler and the most common one is using finite differences. But it turns out that if the payoff function is singular such methods do not work very well (because they have a big variance) and then Malliavin calculus is an efficient instrument in order to avoid the difficulty. Let us be more precise. We write

$$\partial_x \Pi = \partial_x E(\phi(S_T)) = E(\phi'(S_T)\partial_x S_T).$$

If ϕ is smooth then ϕ' makes sense and we may compute the above quantity using the Monte Carlo method. This is for example the case when $\phi(x) = (x - K)_+$ which is the payoff of a call option. But if $\phi(x) = 1_{[K,\infty)}(x)$ (a digital option) then $\phi' = \delta_K$ is a Dirac function and we are no more able to compute the above quantity. Then we have to use integration by parts in order to "regularize" ϕ'. But if we go straight away we obtain $E(\phi'(S_T)\partial_x S_T) = E(\phi(S_T)\delta(U))$ where U is some stochastic process which is not adapted. So we have to deal with an anticipative integral and this is hard to simulate. We would like to get the same type of formula but with a process U which is adapted so that $\delta(U)$ becomes an Ito integral - which is simple to simulate. And this is what Bismut Elworthy formula produces.

Now the motivation is given and we go on to the result itself. We consider an one dimensional diffusion process

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt.$$

We assume that σ and b are smooth so that $X_t \in D^{1,2}$. We denote $Y_t = \partial_x X_t$, $Z_t = 1/Y_t$ and by the variance of constants method we obtain $D_s X_t = Y_t Z_s \sigma(X_s)$.

Proposition 17. *Suppose that $\sigma, b \in C_p^2$ and $E(|\sigma(X_t)|^{-p}) < \infty$. Then for every $G \in D^{1,\infty}$ and $\phi \in C^1$*

$$E(\phi'(X_T)Y_T G) = \frac{1}{T} E(\phi(X_T)(G \int_0^T \frac{Y_s}{\sigma(X_s)} dW_s - \int_0^T \frac{D_s G \times Y_s}{\sigma(X_s)} ds))$$

where $Y_t = \partial_x X_t$.

Proof. We write

$$D_s \phi(X_T) = \phi'(X_T) D_s X_T = \phi'(X_T) Y_T Z_s \sigma(X_s)$$

which gives

$$\phi'(X_T)Y_T = D_s\phi(X_T)\frac{Y_s}{\sigma(X_s)}.$$

Since the equality holds true for each $s \in [0,T]$ we may integrate and obtain

$$\phi'(X_T)Y_T = \frac{1}{T}\int_0^T D_s\phi(X_T)\frac{Y_s}{\sigma(X_s)} ds.$$

We use now the duality formula in order to obtain

$$\begin{aligned} E(\phi'(X_T)Y_TG) &= \frac{1}{T}E(\int_0^T D_s\phi(X_T)\frac{GY_s}{\sigma(X_s)}ds) = \frac{1}{T}E(\phi(X_T)\delta(\frac{GY_\circ}{\sigma(X_\circ)})) \\ &= \frac{1}{T}E(\phi(X_T)(G\int_0^T \frac{Y_s}{\sigma(X_s)}dW_s - \int_0^T \frac{D_sG \times Y_s}{\sigma(X_s)}ds)). \end{aligned}$$

□

Finally we give some exercises concerning sensitivity computations.

EX 1. Assume that S follows the Black Scholes dynamics

$$dS_t = \sigma S_t dW_t, \quad S_0 = x$$

and let $\Pi = E(\phi(S_T))$. Compute $\partial_x \Pi$ and $\partial_\sigma \Pi$ for $\phi = 1_{[K,\infty)}$.

EX 2. (Stochastic volatility model) Assume that S follows the Black Scholes dynamics

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

where W^1 and W^2 are two correlated Brownian motions with $d\langle W^1, W^2\rangle_t = \rho dt, \rho \in [-1,1]$. Compute $\partial_x \Pi$ and $\partial_\rho \Pi$ for $\phi = 1_{[K,\infty)}$.

EX 3. (Change option). Let S^1 and S^2 be two financial assets with dynamics

$$\begin{aligned} dS_t^1 &= \sigma_1 S_t^1 dW_t^1, \quad S_0^1 = x^1, \\ dS_t^2 &= \sigma_2 S_t^2 dW_t^2, \quad S_0^2 = x^2. \end{aligned}$$

where W^1 and W^2 are two correlated Brownian motions with $d\langle W^1, W^2\rangle_t = \rho dt, \rho \in [-1,1]$. We consider an option which pays one dollar if $S_T^1 \geq S_T^2$. Compute the sensitivity of the price with respect to σ_1 and to x^1.

Solution of EX 2. The first step is to decorate the Brownain motions: 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, B^2 are two independent Brownian motions. Then clearly W^1 and W^2 are Brownian motions with correlation ρ so the $SDE's$ become

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

We want to compute

$$\partial_\rho \Pi = \partial_\rho E(1_{[K,\infty)}(S_T)) = E(1'_{[K,\infty)}(S_T))\partial_\rho S_T).$$

This problem do not enter directly in the framework of the Bismut Elworthy formula but we will use the same strategy as there in order to solve it. First of all we solve the first equation:

$$S_T = x\exp(\sqrt{1-\rho^2}\int_0^T \eta_t B_t^1 + \rho\int_0^T \eta_t B_t^2 - \frac{1}{2}\int_0^T \eta_t^2 dt).$$

Then

$$\begin{aligned} \partial_\rho S_T &= S_T \times (\int_0^T \eta_t B_t^2 - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \eta_t B_t^1) = S_T \times G \quad with \\ G &= \int_0^T \eta_t B_t^2 - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \eta_t B_t^1. \end{aligned}$$

We compute now

$$D_s^1 S_T = S_T \times \sqrt{1-\rho^2} \times \eta_s = \partial_\rho S_T \times \frac{\eta_s\sqrt{1-\rho^2}}{G}.$$

It follows that (with $\phi = 1_{[K,\infty)}$)

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

Then

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

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

$$\begin{aligned} E(\phi'(S_T)\partial_\rho S_T) &= \frac{1}{T\sqrt{1-\rho^2}}E(\int_0^T D_s^1\phi(S_T) \times \frac{G}{\eta_s}ds) \\ &= \frac{1}{T\sqrt{1-\rho^2}}E(\phi(S_T) \times \delta_1(\frac{G}{\eta})) \\ &= \frac{1}{T\sqrt{1-\rho^2}}E(\phi(S_T) \times (G\delta_1(\frac{1}{\eta}) - \left\langle D^1 G, \frac{1}{\eta}\right\rangle)). \end{aligned}$$

Since η is adapted we have $\delta_1(\frac{1}{\eta}) = \int_0^T \eta_s^{-1} dB_s^1$ and we also have

$$D_s^1 G = -\frac{\rho}{\sqrt{1-\rho^2}}\eta_s.$$

We conclude that, up to an error, we have

$$E(\phi'(S_T)\partial_\rho S_T) = \frac{1}{T\sqrt{1-\rho^2}} E(\phi(S_T) \times (G\int_0^T \eta_s^{-1} dB_s^1 + \frac{\rho T}{\sqrt{1-\rho^2}})).$$

□

0.2.7. Higher order derivatives

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

$$D_{s_1,s_2}^{(i,j)} F = D_{s_1}^i D_{s_2}^j F = \frac{\partial^2 F}{\partial \Delta_n^{i,s_1} \partial \Delta_n^{j,s_2}}.$$

The definition does not depend on n. Moreover we have the following duality relation. For $U_1, U_2 \in P_n$ we have

$$E(\int_0^1 \int_0^1 D_{s_1,s_2}^{(i,j)} F \times U_1(s_1) \times U_2(s_2) ds_1 ds_2) = E(F \int_0^1 (\int_0^1 U_1(s_1) \times U_2(s_2) \widehat{d}W_{s_1}^i) \widehat{d}W_{s_2}^j)).$$

In the above formula dW^i and dW^j designees Skorohod integrals. 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. And using the above formula permits to 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 from of the space of second order differentiable functionals in terms of Sobolev norms. We define on S the norm

$$\|F\|_{2,p}^p = E|F|^p + E((\int_0^1 \sum_{i=1}^d |D_s^i F|^2 ds)^{p/2}) + E((\int_{(0,1)^2} \sum_{i,j=1}^d |D_{s_1,s_2}^{i,j} F|^2 ds_1 ds_2)^{p/2}).$$

And we put

$$D^{2,p} = \overline{S}^{\|\circ\|_{2,p}}, \quad D^{2,\infty} = \cap_{p\in N} D^{2,p}.$$

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

$$D^{\alpha}_{s_1,...,s_k}F=D^{\alpha_1}_{s_1}...D^{\alpha_k}_{s_k}F=\frac{\partial^k F}{\partial\Delta_n^{\alpha_1,s_1}...\partial\Delta_n^{\alpha_k,s_k}}.$$

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

$$\|F\|^p_{k,p}=E\,|F|^p+\sum_{r=1}^{k}E((\int_{(0,1)^r}\sum_{|\alpha|=r}\left|D^{\alpha}_{s_1,...s_r}F\right|^2ds_1...ds_r)^{p/2})$$

and

$$D^{k,p}=\overline{S}^{\|\circ\|_{k,p}},\quad D^{k,\infty}=\cap_{p\in N}D^{k,p},\quad D^{\infty}=\cap_{k\in N}\cap_{p\in N}D^{k,p}.$$

The space D^{∞} is the "good" space in order to work because we will be able to iterate the integration by parts formula. It represents the analogues of C^{∞} in the standard analysis. But notice that in Malliavin calculus (and in particular in the expression of the weight coming on in the integration by parts formula) we have two basic operators: the Malliavin derivative D and the Ornstein Uhlenbeck operator L. And for $F\in D^{\infty}$ we are sure that we may use D as many times as we want - but not L! It will be proved in the following section that for $F\in D^{\infty}$ we may also compute L^kF for any $k\in N$. And moreover, D^{∞} is an algebra. So a recurrence procedure based on the previous integration by parts formula gives:

Theorem 18. *Let $F=(F_1,...,F_m)$ with $F_i\in D^{\infty}$, $F\in DomL$ and $LF\in D^{\infty}$ and let $G\in D^{\infty}$. Suppose that $(N-D)$ holds true for F. Then for every $\phi\in C_p^{\infty}(R^m,R)$ and every multi-index $\alpha=(\alpha_1,...,\alpha_p)\in\{1,...,m\}^p$ we have*

$$E(\partial_{\alpha}\phi(F)G)=E(\phi(F)H_{\alpha}(F;G))$$

with $H_i(F,G),i=1,...,m$ given in the basic integration by parts formula and with $H_{\alpha}(F;G)$ constructed by

$$H_{(\alpha_1,...,\alpha_p)}(F;G)=H_{\alpha_p}(F,H_{(\alpha_1,...,\alpha_{p-1})}(F;G)).$$

Proof. We have to check that for each $\alpha,H_{\alpha}(F;G)\in D^{\infty}$. Once we know this we apply recursively the basic integration by parts formula. So we have to check that $G\in D^{\infty}\Rightarrow H_i(F;G)\in D^{\infty}$. And this is true because D^{∞} is an algebra and $F_j\in D^{\infty}\Rightarrow F_j\in DomL$ and $LF_j\in D^{\infty}$ (the last two assertions have to be proved - they follow from Mayer's inequalities). $\square$

Remark 4. *We have the following useful estimates. For every multi-index* α *and every* $p, k \in N$ *there exists some constants* C, q, q', p' *(depending on* α *and on* p, k*) such that*

$$\|H_\alpha(F;G)\|_{k,p} \leq C\,\|G\|_{k+1,p'}^q\,(1 + \|F\|_{k+2,p'} + \|LF\|_{k,p'})^q (E(\det \sigma_F)^{-q'})^q.$$

0.2.8. Link between the integration by parts formula and the density of the law of a random variable

We discuss this problem in an abstract framework. On a probability space (Ω, F, P) we consider some random variables $F = (F_1, ..., F_m)$ and G and a multi-index $\alpha = (\alpha_1, ..., \alpha_k)$. We say that we have an integration by parts formula if there exists some random variable $H_\alpha(F, G) \in L^1$ such that

$$IP_\alpha(F,G) \quad E(\partial_\alpha \phi(F) G) = E(\phi(F) H_\alpha(F,G)) \quad \phi \in C_b^\infty(R^m, R).$$

Notice that $H_\alpha(F, G)$ is not unique. The Malliavin calculus produces such a weight, but maybe other methods produce other weights. Notice also that F and G play not a symmetric part in this formula.

We will discuss the link between such integration by parts formulas and the density of the random variable F. For simplicity we consider first the one dimensional case $m = 1$. Now $IP_k(F, G)$ means

$$IP_k(F,G) \quad E(\phi^{(k)}(F) G) = E(\phi(F) H_k(F,G)) \quad \phi \in C_b^\infty(R, R).$$

Suppose first that we have $IP_1(F, 1)$. Then we write down the following formal computation. If p_F is the density of the law of F then

$$p_F(x) = E(\delta_0(F - x)) = E(1'_{[0,\infty)}(F - x)) = E(1_{[0,\infty)}(F - x) H_1(F, 1)).$$

So we obtain an integral representation for p_F. We prove now that this computation may be done rigorously.

Proposition 19. *Suppose that we have* $IP_1(F, 1)$ *Then the law of* F *is absolutely continuous with respect to the Lebesgue measure and the density verifies*

$$p_F(x) = E(1_{[0,\infty)}(F - x) H_1(F, 1)).$$

In particular $x \to p_F(x)$ *is contiguous.*

Moreover suppose that $IP_1(F,G)$ with $G = H_1(F,1)$ holds true also and $H_1(F,H_1(F,1)) \in L^{1+\delta}$ for some $\delta > 0$. Then $x \to p_F(x)$ is continuously differentiable and we have

$$p'_F(x) = E(1_{[0,\infty)}(F-x)H_1(F,H_1(F,1))).$$

Proof. We take $\phi \in C^\infty(R,R_+)$ which is symmetric, non negative and null outside $(-1,1)$. We also assume that $\int \phi = 1$. Then we define $\phi_\varepsilon(x) = \varepsilon^{-1}\phi(x\varepsilon^{-1})$ which converge weakly to the Dirac function. Moreover we denote $\Phi_\varepsilon(x) = \int_{-\infty}^x \phi_\varepsilon(y)dy$ which converge to $\overline{1}_{(0,\infty)}(x) = 1_{(0,\infty)}(x) + \frac{1}{2}\delta_0(x)$. Take now $f \in C_b$. Since $f_\varepsilon = f * \phi_n \to f$ we have $E(f(F)) = \lim_\varepsilon E(f_\varepsilon(F))$. We use the integration by parts formula in order to obtain

$$\begin{aligned} E(f_\varepsilon(F)) &= E(\int_R f(y)\phi_\varepsilon(F-y)dy) = \int_R f(y)E(\Phi'_\varepsilon(F-y))dy \\ &= \int_R f(y)E(\Phi_\varepsilon(F-y)H_1(F,1))dy \to \int_R f(y)E(\overline{1}_{(0,\infty)}(F-y)H_1(F,1))dy \end{aligned}$$

This proves that the law of F is absolutely continuos. But then we may change $\overline{1}_{(0,\infty)}$ with $1_{(0,\infty)}$ because the Lebesgue measure does not charge points. So we have the integral representation. And the continuity follows from the Lebesgue convergence theorem.

Now we use one more integration by parts and we obtain

$$p_F(x) = E(1_{[0,\infty)}(F-x)H_1(F,1)) = E((F-x)_+H_1(F,H_1(F,1))).$$

And if we now that $H_1(F,H_1(F,1)) \in L^{1+\delta}$ then we may differentiate under the integral. So p_F is differentiable and $p_F(x) = E(1_{[0,\infty)}(F-x)H_1(F,H_1(F,1)))$.□

Let us now give a proposition which is somehow the converse of the above one.

Proposition 20. *Suppose that the law of F is absolutely continuous with respect tot the Lebegue measure and that $x \to p_F(x)$ is continuous and $E\,|p'_F(F)| < \infty$.. Then for every $g \in C_b^1$ and $G = g(F)$ we have the following integration by parts formula*

$$E(\phi'(F)G) = E(\phi(F)H_1(F,G))$$

with $H_1(F,G) = -(g' + g\partial \ln p_F)$.

Proof. We use standard integration by parts on R in order to get

$$\begin{aligned} E(\phi'(F)G) &= \int_R \phi'(x)g(x)p_F(x)dx = -\int_R \phi(x)(g'(x)p_F(x) + g(x)p'_F(x))dx \\ &= -\int_R \phi(x)(g'(x)p_F(x) + g(x)\partial(\ln p_F(x))p_F(x)dx \\ &= -E(\phi(F)(g' + g\partial \ln p_F)(F)). \end{aligned}$$

□

The multi-dimensional statement is the following.

Proposition 21. *Let $F = (F_1, ..., F_m)$ be a random variable such that for each multi-index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., m\}^k$ the integration by parts formula $I_\alpha(F, 1)$ given in the beginning of this section, holds true. Then $P \circ F^{-1}(x) = p_F(x)dx$ and one has*

$$p_F(x) = E(\prod_{i=1}^m 1_{[0,\infty)}(F_i - x_i)H_{(1,...,m)}(F, 1)).$$

Moreover $x \to p_F(x)$ is of class C^∞ and

$$\partial_\alpha p_F(x) = E(\prod_{i=1}^m 1_{[0,\infty)}(F_i - x_i)H_{(1,...,m,\alpha_1,...,\alpha_k)}(F, 1)).$$

The proof follows the same idea as above so we skip it.

0.3. Wiener chaos decomposition

For notational convenience we will work with an one dimensional Brownian motion in this section. We denote by $L^2_{s,k}$ the space of the functions $f_k : [0, 1]^k \to R$ which are symmetric (we mean that $f(t_1, ..., t_k) = f(t_{\pi_1}, ..., t_{\pi_k})$ for any permutation π) and such that $\int_{[0,1]^k} |f_k|^2 ds_1...ds_k =: \|f_k\|_k^2 < \infty$. For such a function we define the iterated integral

$$\int_0^1 \int_0^{t_1} ... \int_0^{t_{k-1}} f_k(t_1, ..., t_k)dW_{t_k}...dW_{t_1}$$

and the "multiple stochastic integral"

$$I_k(f_k) := k! \int_0^1 \int_0^{t_1} ... \int_0^{t_{k-1}} f_k(t_1, ..., t_k)dW_{t_k}...dW_{t_1}.$$

If we take f_k to be constant on rectangles we may see that the iterated integral is the integral on a simplex - and since the function is symmetric, we have to multiply with $k!$ (the number of simplexes) in order to obtain the integral on the whole cube $[0,1]^k$. EX: $k = 2$....

We have the following basic relations:

$$\begin{aligned} i) \quad E(I_k(f_k) \times I_p(f_p)) &= 0 \quad if \quad k \neq p, \\ ii) \quad E(I_k(f_k)I_k(g_k)) &= k! \left\langle f_k, g_k \right\rangle_k \quad so \quad E(I_k^2(f_k)) = k! \left\| f_k \right\|_k^2. \end{aligned}$$

The second relation is obtained using the isometry property k times (after replacing the multiple integral by iterated integrals). In order to obtain the first relation suppose that $k < p$. We use the isometry property k times and the stochastic integrals in $I_k(f_k)$ disappear - we get a deterministic quantity. But we have still $p - k \geq 1$ stochastic integrals from $I_p(f_p)$. And they are of expectation zero.

We define the "chaos of order k" as

$$H_k = \{I_k(f_k) : f_k \in L^2_{k,s}\} \subset L^2(\Omega, F_1, P).$$

This is a linear closed subspace of $L^2(\Omega, F_1, P)$. In order to see that H_k is closed one employes the isometry property $ii)$: so if $F_n = I_k(f_k^n), n \in N$ is a Cauchy sequence in $L^2(\Omega, F_1, P)$ then $f_k^n, n \in N$ is a Cauchy sequence in $L^2_{k,s}$ and we get a limit f_k.

Moreover, using $i)$ we see that H_k and H_p are orthogonal for $k \neq p$. And we have (with the convention $H_0 = R =$ constants)

$$L^2(\Omega, F_1, P) = \oplus_{k=0}^{\infty} H_k.$$

This is the chaos decomposition theorem (which I do not prove). So for every $F \in L^2(\Omega, F_1, P)$ there exist some kernels $f_k \in L^2_{k,s}, k \in N$ such that $\sum_{k=0}^{\infty} k! \left\| f_k \right\|_k^2 < \infty$ and

$$F = \sum_{k=0}^{\infty} I_k(f_k), \quad \left\| F \right\|_2^2 = \sum_{k=0}^{\infty} k! \left\| f_k \right\|_k^2.$$

We denote $\mathrm{Pr}_k(F) = I_k(f_k)$.

Proposition 22. *A. For every $k \in N$ one has $H_k \subset D^{1,2}$ and*

$$D_s I_k(f_k) = k \times I_{k-1}(f_k(s, \circ)).$$

Moreover $H_k \subset \cap_{p\in N} D^{p,2}$ *and for* $p \le k$

$$D^p_{s_1,...,s_p} I_k(f_k) = k(k-1)...(k-p+1) \times I_{k-p}(f_k(s_1,...,s_p,\circ)).$$

B. Let $f_k \in L^2_{s,k}$. *Then* $(I_{k-1}(f_k(s,\circ)))_{s\in[0,1]} \in Dom_2\delta$ *and*

$$\int_0^1 I_{k-1}(f_k(s,\circ))dW_s = \delta(I_{k-1}(f_k(s,\circ)) = I_k(f_k).$$

C. $H_k \subset DomL$ *and*

$$LI_k(f_k) = kI_k(f_k).$$

Proof. A. We denote

$$J_k(f_k)(T) = \int_0^T \int_0^{t_1} ... \int_0^{t_{k-1}} f_k(t_1,...,t_k)dW_{t_k}...dW_{t_1}$$

and we prove that $J_k(f_k)(T) \in D^{1,2}$ and that $D_s J_k(f_k)(T) = J_{k-1}(f_k(s,\circ))(T)$. Since $I_k(f_k) = J_k(f_k)(1)$ A follows. We proceed by recurrence on k. Suppose that the assertion is true for $k-1$ and let us prove it for k. We will use the rule of derivation of stochastic integrals given in Ex 4. In order to do it we need to check some continuity properties but we postpone this for the end of the proof. For the moment we assume that we are allowed to use that rule. We have

$$J_k(f_k)(T) = \int_0^T J_{k-1}(f_k(t_1,\circ))(t_1)dW_{t_1}$$

so we take derivatives and we obtain

$$D_s J_k(f_k)(T) = J_{k-1}(f_k(s,\circ))(s) + \int_s^T D_s J_{k-1}(f_k(t_1,\circ))(t_1)dW_{t_1}.$$

We use the recurrence hypothesis and we develop $J_{k-1}(f_k(s,\circ))(s)$ in order to get

$$D_s J_k(f_k)(T) = \int_0^s J_{k-2}(f_k(s,t_1,\circ))(t_1)dW_{t_1} + \int_s^T J_{k-2}(f_k(t_1,s,\circ))(t_1)dW_{t_1}.$$

Now we use symmetry and get $f_k(t_1,s,\circ) = f_k(s,t_1,\circ)$ so that

$$D_s J_k(f_k)(T) = \int_0^T J_{k-2}(f_k(s,t_1,\circ))(t_1)dW_{t_1} = J_{k-1}(f_k(s,\circ)).$$

So the formula is proved. We will now check that we have the needed continuity properties for $U_t = J_{k-1}(f_k(t, \circ))(t)$:

$$E \int_0^T \left| U_t - U_{\tau_n(t)} \right|^2 dt + E \int_0^T \int_0^T \left| D_s U_t - D_s U_{\tau_n(t)} \right|^2 dt ds \to 0.$$

It is easy to check that this is true if f_k is continuos. So we have the result for continuous kernels.

Take now some general $f \in L^2_{s,k}$ and approximate it by a sequence of continuos kernels f_k^n (because the continuous functions are dense in $L^2[0,1]^k$. But these functions may be not symmetric. Then we take the symmetrization:

$$\overline{f}_k^n(t_1, ..., t_k) = \frac{1}{k!} \sum_\pi f_k^n(t_{\pi_1}, ..., t_{\pi_k})$$

where the sum is over all the permutation π. Denote $f_k^n \circ \pi(t_1, ..., t_k) =: f_k^n(t_{\pi_1}, ..., t_{\pi_k})$. Since f_k is symmetric we have $f_k = f_k \circ \pi$ so that

$$\left\| \overline{f}_k^n - f_k \right\|_k \leq \frac{1}{k!} \sum_\pi \left\| f_k^n \circ \pi - f_k \circ \pi \right\|_k = \left\| f_k^n - f_k \right\|_k \to 0.$$

Now we have $I_k(\overline{f}_k^n) \to I_k(f_k)$ and $D_s I_k(\overline{f}_k^n) = k I_k(\overline{f}_k^n(s, \circ)) \to I_k(f_k(s, \circ))$ so we conclude that $I_k(f_k) \in D^{1,2}$ and $D_s I_k(f_k) = I_k(f_k(s, \circ))$.

In order to obtain the formula for higher order derivatives we iterate the procedure.

B. Suppose for a moment that we know that $I_{k-1}(f_k(s, \circ)) \in Dom_2 \delta$ and let us prove the formula . We will taste against multiple integrals. Using the duality relation we obtain

$$\begin{aligned} E(I_p(g_p)\delta(I_{k-1}(f_k(s, \circ)))) &= E(\int_0^1 D_s I_p(f_p) \times I_{k-1}(f_k(s, \circ)) ds) \\ &= p \int_0^1 E(I_{p-1}(f_p(s, \circ)) \times I_{k-1}(f_k(s, \circ)) ds \end{aligned}$$

and this is zero if $k \neq p$ and is $k \langle g_k, f_k \rangle$ if $k = p$. Moreover we have $E(I_p(g_p)\delta(I_k(f_k)) = 0$ if $k \neq p$ and $E(I_p(g_p)\delta(I_k(f_k)) = k \langle g_k, f_k \rangle$ if $k = p$. So the two coincides and this proves that $\delta(I_{k-1}(f_k(s, \circ)) = I_k(f_k)$.

Let us now prove that $I_{k-1}(f_k(s,\circ)) \in Dom_2\delta$. Suppose first that f_k is a step functuion. Then $I_{k-1}(f_k(s,\circ))$ is a simple process so it belongs to the domain of δ. Then by approximation we get the result for f_k continuos and then for a general f_k (the fact that we know that $\delta(I_{k-1}(f_k(s,\circ)) = I_k(f_k)$ is useful in order to prove convergence).

C. Just composition $L = -\delta D.\square$

We are now able to characterize the domains of the operators.

Proposition 23. *We have* $D^{1,2} = \{F : \sum_{k=1}^{\infty} k \times k! \times \|f_k\|_k^2 < \infty\}$ *and*

$$D_s F = \sum_{k=1}^{\infty} k \times I_{k-1}(f_k(s,\circ)), \quad \|F\|_{1,2}^2 = E(F^2) + \sum_{k=1}^{\infty} k \times k! \times \|f_k\|_k^2 .$$

Moreover $D^{p,2} = \{F : \sum_{k=p}^{\infty} \frac{k!}{(k-p)!} \times k! \times \|f_k\|_k^2 < \infty\}$ *and*

$$\begin{aligned} D^p_{s_1,...,s_p} F &= \sum_{k=p}^{\infty} \frac{k!}{(k-p)!} \times I_{k-p}(f_k(s_1,...,s_p,\circ)), \\ \|F\|_{p,2}^2 &= E(F^2) + \sum_{p'=1}^{p} \sum_{k=p'}^{\infty} \frac{k!}{(k-p')!} \times k! \times \|f_k\|_k^2 . \end{aligned}$$

Proof. Let $A := \{F : \sum_{k=1}^{\infty} k \times k! \times \|f_k\|_k^2 < \infty\}$. If $F \in A$ then we approximate it by $F_n = \sum_{k=0}^{n} I_k(f_k)$ and we obtain that $F \in D^{1,2}$. So $A \subset D^{1,2}$ and it is easy to check that it is a closed linear subspace. In order to see that $A = D^{1,2}$ we will check that every $F \in D^{1,2}$ which is orthogonal (with respect to $\langle \circ, \circ \rangle_{1,2}$) on all the multiple integrals, is null. Indeed

$$\begin{aligned} 0 &= \langle I_k(f_k), F \rangle_{1,2} = E(I_k(f_k)F) + E(\langle DI_k(f_k), DF \rangle) = E((I_k(f_k) + \delta(DI_k(f_k))F) \\ &= (1+k)E(I_k(f_k)F). \end{aligned}$$

And this implies $F = 0$. $\square$

Corollary 24. *(Strook's formula) If* $F \in D^{p,2}$ *then*

$$f_p(s_1,...,s_p) = \frac{1}{p!} E(D^p_{s_1,...,s_p} F).$$

Proof. We take expectation in the formula of the derivative operator and we get

$$E(D^p_{s_1,...,s_p}F) = \sum_{k=p}^{\infty} \frac{k!}{(k-p)!} \times E(I_{k-p}(f_k(s_1,...,s_p,\circ))).$$

For $k > p$ we the expectation is null. And for $k = p$ we get $k!f_k(s_1,...,s_k)$.□

EX. Write down the chaos decomposition of $F = W_1^3$ using two methods: 1. Strook formula. 2. Clark Ocone formula.

Proposition 25. *$Dom_2(L) = \{F : \sum_{k=1}^{\infty} k^2 \times k! \times \|f_k\|_k^2 < \infty\}$. And we have $LF = \sum_{k=1}^{\infty} k \times I_k(f_k)$.and $E|LF|^2 = \sum_{k=0}^{\infty} k^2 \times k! \|f_k\|_k^2$.*

The proof is the same as for the derivatives, but we work with the norm $\|F\|_L^2 = \|F\|_2^2 + \|LF\|_2^2$.□

An easy computation shows that $\|F\|_L = \|F\|_{2,2}$ so $Dom_2(L) = D^{2,2}$.

One may define other norms in the following way. We have $(I-L)(F) = \sum_{k=1}^{\infty}(1-k) \times I_k(f_k)$ so we may define $(I-L)^s(F) = \sum_{k=1}^{\infty}(1-k)^s \times I_k(f_k)$, $s > 0$. Then one defines

$$|||F|||_{q,p} = \left\|(I-L)^{q/2}(F)\right\|_{L^p} = (E(\left|(I-L)^{q/2}(F)\right|^p))^{1/p}.$$

Notice that for $q = 2$ we have

$$\begin{aligned} |||F|||_{2,2}^2 &= E((I-L)(F))^2 = E(F^2) - 2E(FLF) + E(LF)^2 \\ &= \|F\|_L^2 + 2E\langle DF, DF\rangle = \|F\|_{2,2}^2 + 2E\langle DF, DF\rangle \end{aligned}$$

so we obtain

$$\|F\|_{2,2}^2 \leq |||F|||_{2,2}^2 \leq 3\|F\|_{2,2}^2.$$

This inequalities may be generalized (completely non trivial) for every $p \geq 1$ and $q \in N$.

Theorem 26. *(Meyer's inequalities [7]) For every $p \geq 1$ and $q \in N$ there exists some $c_{q,p}, C_{q,p}$ such that*

$$c_{q,p}\|F\|_{q,p} \leq |||F|||_{q,p} \leq C_{q,p}\|F\|_{q,p}.$$

We will now discuss the domain of δ. Let $u \in L^2(\Omega, L^2([0,1])$. Then for almost every $t \in [0,1]$ we have $u(t,\circ) \in L^2(\Omega)$ and consequently $u(t,\circ) = \sum_{k=0}^{\infty} I_k(f_k(t,\circ))$ and the sum converges in $L^2(\Omega, L^2([0,1])$. Having in mind a result given previously we believe that $I_k(f_k(t,\circ)) \in Dom_2\delta$ and that $\delta(I_k(f_k(t,\circ))) = I_{k+1}(f_k)$. But the trouble is that now f_k is a function of $k+1$ variables which is not symmetric with respect to the first variable. So we have to come back and extend the previous result in this case. We define the symmetrization of f_k:

$$\overline{f}_k(t_1, ..., t_{k+1}) := \frac{1}{(k+1)!} \sum_{\pi} f_k(t_{\pi_1}, ..., t_{\pi_{k+1}}).$$

Notice that if g_{k+1} is any symmetric function then we have

$$\int_{[0,1]^{k+1}} g_{k+1} f_k = \int_{[0,1]^{k+1}} g_{k+1} \overline{f}_k.$$

Then we have the following result: the process $(I_k(f_k(t,\circ))_{t \in [0,1]}$ is in the domain of δ (this is obtained by approximation and we do not insist on it) and we have $\delta(I_k(f_k(t,\circ)) = (\int_0^1 I_k(f_k(t,\circ) \widehat{d}W_t = I_{k+1}(\overline{f}_k)$. In order to prove the last equality we test against $I_p(g_p)$:

$$E(\delta(I_k(f_k(t,\circ))I_p(g_p)) = pE \int_0^1 I_k(f_k(t,\circ) I_{p-1}(g_p(t,\circ)))dt.$$

This quantity is zero if $p \neq k+1$ and if $p = k+1$ we get

$$\begin{aligned} & (k+1) \int_0^1 E(I_k(f_k(t,\circ) I_k(g_{k+1}(t,\circ))dt \\ = \; & (k+1)k! \int_0^1 \int_{[0,1]^k} f_k(t,\circ) g_{k+1}(t,\circ))dt = (k+1)! \int_{[0,1]^{k+1}} \overline{f}_k g_{k+1} \\ = \; & E(I_k(\overline{f}_k) I_{k+1}(g_{k+1})) \end{aligned}$$

and we are done.

The general result is now the following.

Proposition 27. *Let $u \in L^2(\Omega, L^2([0,1])$ such that $u(t,\circ) = \sum_{k=0}^{\infty} I_k(f_k(t,\circ)$. Then $u \in Dom_2(\delta)$ if and only if $\sum_k (k+1)! \left\| \overline{f}_k \right\|_{k+1}^2 < \infty$ (which is equivalent with the fact that $\sum_k I_{k+1}(\overline{f}_k)$ converges in L^2). In this case $\delta(u) = \sum_k I_{k+1}(\overline{f}_k)$.*

Proof. Suppose that $\sum_k I_{k+1}(\overline{f}_k)$ converges. Then $u_n(t,\circ) = \sum_{k=0}^n I_k(f_k(t,\circ) \to u(t,\circ)$ and $\delta(u_n) = \sum_{k\leq} I_{k+1}(\overline{f}_k) \to \sum_k I_{k+1}(\overline{f}_k)$ so we have the result. Assume now that $u \in Dom_2(\delta)$ and let us prove that the above sum is convergent. Since $\delta(u) \in L^2$ we have $\delta(u) = \sum_k I_k(\widehat{f_k})$ for some kernels $\widehat{f_k}$ and the above sum is convergent in L^2. We write now

$$\begin{aligned} E(I_p(\widehat{f_p})I_p(g_p)) &= E(\delta(u)I_p(g_p)) = pE\int_0^1 u(t)I_{p-1}(g_p(t,\circ))dt \\ &= pE\int_0^1 I_{p-1}(f_p(t,\circ)I_{p-1}(g_p(t,\circ))dt \end{aligned}$$

and the same calculus as above gives $E(I_p(\overline{f}_p)I_p(g_p))$. So we conclude that $\overline{f}_p = \widehat{f_p}$ and we are done. □

0.4. Diffusion processes

We consider the diffusion process

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

We assume that $\sigma_j^i, b^i \in C_b^\infty(R^m : R)$. A first result concerns the regularity of X_t in Malliavin sense.

Proposition 28. *Suppose that $\sigma_j^i, b^i \in C_b^\infty(R^m : R)$. Then $X_t^i \in Dom_p L, \forall p \in N$ and $X_t^i, LX_t^i \in D^\infty$ and one has*

$$\left\|X_t^i\right\|_{k,p} + \left\|LX_t^i\right\|_{k,p} \leq C_{k,p}(1+|x|) \quad \forall k,p \in N.$$

where $C_{k,p}$ depends on the bounds of σ, b and of their derivatives up to order $k+2$.

The proof is strightforward but rather long so we skip it.

Consider now the Malliavin covariance matrix σ_{X_t}. We have seen in a previous section that we have

$$D_s X_t = Y_t Z_s \sigma(X_s)$$

where

$$\begin{aligned} Y_t &= I + \sum_{j=1}^{d} \int_0^t \partial\sigma_j(X_s) Y_s \circ dW_s^j + \int_0^t \partial\overline{b}(X_s) Y_s ds \quad and \\ Z_t &= I - \sum_{j=1}^{d} \int_0^t Z_s \partial\sigma_j(X_s) \circ dW_s^j - \int_0^t Z_s \partial\overline{b}(X_s) ds. \end{aligned}$$

Here $\overline{b} = b - \sum_{j=1}^{d} \sigma_j \times \nabla\sigma_j$ and $\circ dW_s^j$ designs the Stratonovich integral.

Now the integration by parts formula from Malliavin calculus reads:

Theorem 29. *Suppose that* $\sigma_j^i, b^i \in C_b^\infty(R^m : R)$ *and moreover*

$$E\left(\frac{1}{(\det \sigma_{X_t})^p}\right) < \infty, \quad p \in N.$$

Then for every $G \in D^\infty$ *and every* $\phi \in C_p^\infty(R^m, R)$ *we have*

$$E(\partial_\alpha \phi(X_t) G) = E(\phi(X_t) H_\alpha(X_t, G)) \tag{1}$$

where $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., m\}^k$ *is a multi-index,* ∂_α *is the derivative corresponding to* α *and the weights* $H_\alpha(X_t, G)$ *are defined by recurrecne:*

$$\begin{aligned} H_i(X_t, G) &= \sum_{j=1}^{m} (-G\gamma_{X_t}^{ij} L X_t^j - \langle D(G\gamma_{X_t}^{ij}, DX_t^j \rangle), \\ H_\alpha(X_t, G) &= H_{\alpha_k}(X_t, H_{(\alpha_1, \ldots, \alpha_{k-1}}(X_t, G)) \end{aligned}$$

with $\gamma_{X_t} = \sigma_{X_t}^{-1}$.

We also have the following useful estimation:

Lemma 30. *Let* $k, p \in N$. *There exists some constants* $C = C_{k,p}, q = q_{k,p}, q' = q'_{k,p}$ *(depending on the bounds of* σ, b *and their derivatives up to order* $k+2$*) such that for every multi-index* α *with* $|\alpha| \leq k$ *one has*

$$E\left|H_\alpha(X_t, G)\right|^p \leq C(1 + |x|) \left\|G\right\|_{k,q}^{q'} \times (E((\det \sigma_{X_t})^{-q}))^{q'}. \tag{2}$$

This is another strightforward but long computation which we will skip. In fact, in a first step one proves the general inequality

$$E\left|H_{\alpha}(X_t,G)\right|^p \leq C(1+\|F\|_{k,q}+\|LF\|_{k,q})^{q'}\|G\|_{k,q}\times E((\det\sigma_{X_t})^{-q}))^{q'}$$

and then employs the estimation of $\|F\|_{k,q}, \|LF\|_{k,q}$ given above.

So the problem we have now is to estimate $E((\det\sigma_{X_t})^{-q})$. A first step is to use the variance of constants method in order to obtain

$$E((\det\sigma_{X_t})^{-q}) \leq C_q E((\det U_t)^{-2q}) \tag{3}$$

where the matrix U_t is given by

$$U_t^{i,j} = \int_0^t (Z_s\sigma\sigma^*(X_s)Z_s^*)^{i,j} ds.$$

This follows in the following way: we have $\sigma_{X_t} = Y_t \times U_t \times Y_t^*$ and then $\det\sigma_{X_t} = (\det Y_t)^2 \times \det U_t$. One has $Y_tZ_t = I$ so that $(\det Y_t)^{-1} = \det Z_t$. And since Z_t has finite moments of any order we obtain $E(\det Y_t)^{-4q} = E(\det Z_t)^{4q} \leq C$ Then we use Schwartz inequality and we obtain $E((\det\sigma_{X_t})^{-q}) \leq C_q E((\det U_t)^{-2q})$.

So now on we will focus on U. We will first discuss the elliptic case - which is much simpler then the general case given by Hörmander's theorem. We fix the starting point $x \in R^m$ and we assume that for some $c_* > 0$ we have

$$\sigma\sigma^*(x) \geq c_* I. \tag{4}$$

Then we have the following estimate:

Proposition 31. *Suppose that $\sigma\sigma^*(x) \geq c_* I$. Then*

$$E((\det U_t)^{-q}) \leq 1 + (\frac{4}{tc_*}+1)^{mq+1} + \frac{C_q}{c_*^{2(mq+2)}} \tag{5}$$

where C_q is a constant which depends on q, m and on the bounds of σ and b and of their derivatives of order one and two.

Proof. **Step 1**. We compute first

$$\begin{aligned}\langle U_t\xi,\xi\rangle &= \int_0^t \langle Z_s\sigma\sigma^*(X_s)Z_s^*\xi,\xi\rangle\, ds = \int_0^t \langle \sigma^*(X_s)Z_s^*\xi, \sigma^*(X_s)Z_s^*\xi\rangle\, ds \\ &= \sum_{j=1}^d \int_0^t \left\langle \sigma_j^*(X_s)Z_s^*,\xi\right\rangle^2 ds.\end{aligned}$$

Then the smallest eigenvalue λ_t of U_t is

$$\lambda_t := \inf_{|\xi|=1} \langle U_t\xi, \xi\rangle = \inf_{|\xi|=1} \sum_{j=1}^{d} \int_0^t \left\langle \sigma_j^*(X_s)Z_s^*, \xi\right\rangle^2 ds.$$

Since $\det U_t \geq \lambda_t^m$ we have $E((\det U)^{-q}) \leq E(\lambda_t^{-mq})$ so we estimate this last quantity.

Step 2. Using the elementary inequality $(a+b)^2 \geq \frac{1}{2}a^2 - b^2$ we obtain

$$\begin{aligned}
\sum_{j=1}^{d} \left\langle \sigma_j^*(X_s)Z_s^*, \xi\right\rangle^2 &\geq \frac{1}{2}\sum_{j=1}^{d} \left\langle \sigma_j^*(x), \xi\right\rangle^2 - \sum_{j=1}^{d} \left\langle \sigma_j^*(X_s)Z_s^* - \sigma_j^*(x), \xi\right\rangle^2 \\
&\geq \frac{1}{2}c_* - \sum_{j=1}^{d} \left|\sigma_j^*(X_s)Z_s^* - \sigma_j^*(x)\right|^2.
\end{aligned}$$

The second inequality hods true if $|\xi| = 1$. In order to dominate the first term we have used the ellipticity assumption: $\sum_{j=1}^{d} \left\langle \sigma_j^*(x), \xi\right\rangle^2 = \langle \sigma\sigma^*(x)\xi, \xi\rangle \geq c_* |\xi|^2$.

Step 3 We take now $\varepsilon > 0$ and $a > 0$ and we write

$$\begin{aligned}
P(\lambda_t &\leq \varepsilon) \leq P(\inf_{|\xi|=1} \sum_{j=1}^{d} \int_0^{a\varepsilon} \left\langle \sigma_j^*(X_s)Z_s^*, \xi\right\rangle^2 ds \leq \varepsilon) \\
&\leq P(\int_0^{a\varepsilon} (\frac{1}{2}c_* - \sum_{j=1}^{d} \left|\sigma_j^*(X_s)Z_s^* - \sigma_j^*(x)\right|^2) ds \leq \varepsilon) \\
&\leq P(\frac{ac_*}{2}\varepsilon - \sup_{s\leq a\varepsilon} \sum_{j=1}^{d} \left|\sigma_j^*(X_s)Z_s^* - \sigma_j^*(x)\right|^2 a\varepsilon \leq \varepsilon).
\end{aligned}$$

We take $a = 4/c_*$ such that $\frac{ac_*}{2} = 2$ and we conclude that

$$\begin{aligned}
P(\lambda_t &\leq \varepsilon) \leq P(\sup_{s\leq a\varepsilon} \sum_{j=1}^{d} \left|\sigma_j^*(X_s)Z_s^* - \sigma_j^*(x)\right|^2 a\varepsilon \geq \varepsilon) \qquad (6) \\
&= P(\sup_{s\leq a\varepsilon} \sum_{j=1}^{d} \left|\sigma_j^*(X_s)Z_s^* - \sigma_j^*(x)\right|^2 \geq \frac{c_*}{4}) \leq \frac{4^p}{c_*^p} \sum_{j=1}^{d} E(\sup_{s\leq a\varepsilon} \left|\sigma_j^*(X_s)Z_s^* - \sigma_j^*(x)\right|^{2p}) \\
&\leq \frac{C_p}{c_*^p}(a\varepsilon)^p = \frac{C_p}{c_*^{2p}} \times \varepsilon^p
\end{aligned}$$

where C_p is a constant which depends on p and on the bounds of σ and b and of their derivatives of order one and two.

Step 4. We estimate now $E(\lambda_t^{-mq})$. We write

$$\begin{aligned} E(\lambda_t^{-mq}) &= E(\lambda_t^{-mq}1_{\{\lambda_t>1\}}) + \sum_{k=1}^{\infty} E(\lambda_t^{-mq}1_{\{1/k\geq\lambda_t>1/(k+1)\}}) \\ &\leq 1 + \sum_{k=1}^{\infty}(k+1)^{mq}P(\lambda_t \leq \frac{1}{k}). \end{aligned}$$

We will now use (6). Notice that this inequality holds true only if $a\varepsilon < t$ (because we have used the inequality $\int_0^t \geq \int_0^{a\varepsilon}$). So, if $\varepsilon = \frac{1}{k}$, we need $\frac{4}{c_* k} < t$ which gives $k > \frac{4}{tc_*}$. So we come back and write

$$\begin{aligned} E(\lambda_t^{-mq}) &\leq 1 + \sum_{k\leq\frac{4}{tc_*}}(k+1)^{mq} + \sum_{k>\frac{4}{tc_*}}^{\infty}(k+1)^{mq}P(\lambda_t \leq \frac{1}{k}) \\ &\leq 1 + (\frac{4}{tc_*}+1)^{mq+1} + \frac{C_p}{c_*^{2p}}\sum_{k>\frac{4}{tc_*}}^{\infty}(k+1)^{mq} \times \frac{1}{k^p}. \end{aligned}$$

In order to get a convergent sum we take $p = mq + 2$ and the above inequality gives

$$E(\lambda_t^{-mq}) \leq 1 + (\frac{4}{tc_*}+1)^{mq+1} + \frac{C_{mq+2}}{c_*^{2(mq+2)}}.$$

□

We turn no to the application of the above result to the study of the density of the law of X_t.

Theorem 32. *Suppose that $\sigma^{i,j}, b^i \in C_b^\infty(R^m, R)$ and $\sigma\sigma^*(x) \geq c_* I$. Then the law of $X_t(x)$ is absolutely continuous with respect to the Lebesgue measure and the density $p_t(x,y)$ is of class C^∞ with respect to $y \in R^m$. Moreover*

$$p_t(x,y) \leq C(\frac{1}{(c_* t)^p} + 1)\exp(-\frac{C' \left|x-y\right|^2}{t})$$

where C, C', p are constants which depend on m and on the bounds of σ, b and of their derivatives . The same inequality holds true for the derivatives of the density with respect to y.

Proof. We have the integration by parts formula

$$\begin{aligned} E(\partial_\alpha \phi(X_t)) &= E(\phi(X_t) H_\alpha(X_t, 1)) \quad with \\ \|H_\alpha(X_t, 1)\|_p &\leq C_p(1 + |x|)(\frac{1}{c_* t} + 1)^{q_p}, \quad \forall p \in N. \end{aligned}$$

Let us first study the differentiability of $p_t(x, y)$. We will use a formal calculus but it is clear that one may get a rigorous presentation using a regularization procedure. We take

$$\Phi(y) = \prod_{i=1}^{m-1} 1_{[0,\infty)}(y^i) \times (y^m)_+$$

and $\alpha = (1, 2, ..., m-1, m, m)$ so that $\partial_\alpha \Phi = \delta_0$. Then using integration by parts one obtains

$$\begin{aligned} p_t(x, y) &= E(\delta_0(X_t - y)) = E(\partial_\alpha \Phi(X_t - y)) \\ &= E(\Phi(X_t - y) H_\alpha(X_t, 1)). \end{aligned}$$

Then it is easy to check that we may differentiate the term in the right hand side and we obtain

$$\partial_{y^m} p_t(x, y) = \partial_{y^m} E(\Phi(X_t - y) H_\alpha(X_t, 1)) = E(\prod_{i=1}^{m} 1_{[0,\infty)}(X_t^i - y^i) H_\alpha(X_t, 1)).$$

For higher order derivatives one proceeds in a similar way.

Let us now prove the estimates for the tails of the density. If we work with X_t directly there will be some $(1 + |x|)$ coming on in our estimate (because this quantity appears in the estimate of $\|H_\alpha(X_t, 1)\|_p$). This is why it is more clever to look to $F := X_t - x$. Notice that we have $\|H_\alpha(X_t - x, 1)\|_p \leq C_p(1 + c_* t)^{-q_p}$ without $1 + |x|$. This is because $\|X_t - x\|_p \leq C_p$. Let us denote $q_t(x, y)$ the density of $F = X_t - x$. We have $p_t(x, y) = q_t(x, y - x)$. So if we prove that $q_t(x, y) \leq C \exp(-\frac{C'|y|^2}{t})$ then we get $p_t(x, y) \leq C \exp(-\frac{C'|y-x|^2}{t})$.

We take y such that $y^i > x^i, i = 1, ..., m$ and we write

$$\begin{aligned} q_t(x, y) &= E(\prod_{i=1}^{m} 1_{[0,\infty)}(F^i - y^i) H_\alpha(F, 1)) \leq (P(\cap_{i=1}^m \{F^i > y^i\}))^{1/2} \times \|H_\alpha(X_t - x, 1)\|_2 \\ &\leq (\min_{i=1,m} P(F^i > y^i))^{1/2} \times C_2(1 + (c_* t)^{-q_2}). \end{aligned}$$

We use now Bernstein inequality in order to prove that

$$P(F^i > y^i) = P(\sum_{j=1}^d \int_0^t \sigma^{i,j}(X_s)dW_s^j + \int_0^t b^i(X_s)ds > y^i) \leq Ce^{-C'(y^i)^2/t}.$$

Then

$$\min_{i=1,m} P(F^i > y^i) \leq C\exp(-C' \max(y^i)^2/t) \leq C\exp(-C' \left|y\right|^2/t).$$

□

We present now the Hörmander condition. For two differentiable functions $f, g : R^m \to R^m$ we define the Lie bracket $[f,g] := f\partial g - g\partial f$ which reads on components $[f,g]^i := \sum_{j=1}^m f^j \partial_j g^i - g^j \partial_j f^i$. Then we denote $\sigma_j = (\sigma_j^1, ..., \sigma_j^d), j = 1, ..., d$ and $\sigma_0 = (\overline{b}^1, ..., \overline{b}^m)$ so that the equation of our diffusion process reads (with Stratonovich integrals and with the convention $W_t^0 = t$) :

$$dX_t = \sum_{j=0}^d \sigma_j(X_t) \circ dW_t^j.$$

We define now recursively

$$S_0 = \{\sigma_1, ..., \sigma_d\}, \quad S_{k+1} = \{[\phi, \sigma_j], j = 0, ..., d\}, \quad S_\infty = \cup_{k\in N} S_k.$$

Then, given $x \in R^m$, the Hörmander condition in x is

$$(H_x) \quad Span\{\phi(x) : \phi \in S_\infty\} = R^m.$$

Notice that the ellipticity condition $\sigma\sigma^*(x) \geq c_* I$ is equivalent with $SpanS_0(x) = R^m$.

We explain now how this condition appears. Recall that we deal with the matrix $\int_0^t Z_s^* \sigma^* \sigma(X_s) Z_s ds$ with Z solution of the equation

$$dZ_t = -\sum_{j=0}^d Z_t \partial\sigma_j(X_t) \circ dW_t^j.$$

Using Itô's formula we obtain for every $f \in C^2(R^m, R^m)$:

$$d(Z_t f(X_t)) = \sum_{j=0}^d Z_t[\sigma_j, f](X_t) \circ dW_t^j.$$

Then we use this formula in order to develop $Z_t\sigma_j(X_t)$ in "stochastic series". We take $k = 1, ..., d$ and we write

$$\begin{aligned} Z_t\sigma_k(X_t) &= \sigma_k(x) + \sum_{j=0}^{d} \int_0^t Z_s[\sigma_j, \sigma_k](X_s) \circ dW_s^j \\ &= \sigma_k(x) + \sum_{j=0}^{d} [\sigma_j, \sigma_k](x) W_t^j + \sum_{j=0}^{d} \int_0^t (Z_s[\sigma_j, \sigma_k](X_s) - [\sigma_j, \sigma_k](x)) \circ dW_s^j. \end{aligned}$$

If we continue to develop $Z_s[\sigma_j, \sigma_k](X_s) - [\sigma_j, \sigma_k](x)$ then we obtain, for every $N \in N$:

$$\begin{aligned} Z_t\sigma_k(X_t) &= \sigma_k(x) + \sum_{1 \leq |\alpha| \leq N} T^\alpha(\sigma_k)(x) W_t^{\circ,\alpha} \\ &\quad + \sum_{|\alpha|=N} \sum_{j=0}^{d} \int_0^t \int_0^{t_1} ... \int_0^{t_{N-1}} (T^\alpha(\sigma_k)(X_{t_N}) - T^\alpha(\sigma_k)(x)) \circ dW_s^{\alpha_N} ... \circ dW^{\alpha_N} \end{aligned}$$

where $T^\alpha(\sigma_k)$ and $W_t^{\circ,\alpha}$ are defined recursively in the following way. If $\alpha = (\alpha_1, ..., \alpha_p) \in \{0, ..., d\}^p$ is a multi-index, we denote $\overleftarrow{\alpha} = (\alpha_1, ..., \alpha_{p-1})$ and $\overrightarrow{\alpha} = \alpha_p$. Then we define

$$\begin{aligned} T^j(f) &= [f, \sigma_j], \quad j = 0, ..., d, \\ T^\alpha(f) &= [T^{\overleftarrow{\alpha}}(f), \sigma_{\alpha_p}] = [[...[f, \sigma_{\alpha_1}]....], \sigma_{\alpha_p}] \end{aligned}$$

and

$$W_t^{\circ,\alpha} = \int_0^t W_s^{\circ,\overleftarrow{\alpha}} \circ dW_s^{\alpha_p}.$$

Notice that $S_N = \{T^\alpha(\sigma_1), ..., T^\alpha(\sigma_d), |\alpha| = N\}$. So the vectors which appear in Hörmander's conditions are exactly the coefficients of the development of $Z_t\sigma_k(X_t), k = 1, ..., d$ in stochastic series. Notice that S_0 contains only $\sigma_1, ..., \sigma_d$ and not σ_0 because in the covariance matrix σ_0 does not appear. In the following section we will discuss the non degeneracy of a matrix of the form $\int_0^t V_s ds$ where V is a matrix valued process which may be developed in Taylor sums. As an immediate consequence of Theorem 32 given there we obtain the following.

Theorem 33. *Suppose that $\sigma, b \in C_b^\infty$ and that (H_x) holds true. Then the law of X_t^x is absolutely continuous with respect to the Lebesgue measure and $y \to p_t(x, y)$ is of class C^∞.*

Remark 5. *One may also prove that $p_t(x,y)$ is of class C^∞ with respect to $x \in R^m$ and $t>0$. And one may obtain very precise exponential bounds for $p_t(x,y)$ and its derivatives with respect to t,x,y. This is a rather heavy computation - see Kusuoka-Strook.*

0.5. Stochastic series

We are on a probability space and $W=(W^1,...,W^d)$ is a standard Brownian motion. We introduce some notation. For notational convenience we denote $W_t^0=t$. We work with multi-indexes $\alpha=(\alpha_1,...,\alpha_k)\in\{0,1,...,d\}^k$ and we denote

$$|\alpha|=k,\quad p(\alpha)=card\{i:\alpha_i\in\{1,...,d\}\}+2card\{i:\alpha_i=0\}.$$

Given an adapted and square integrable process X we define the iterated stochastic integral

$$\begin{aligned} I_t^\alpha(X) &= \int_0^t ... \int_0^{t_{k-1}} X_{t_k} dW_{t_k}^{\alpha_k} ... dW_{t_1}^{\alpha_1}, \quad and \\ W_t^\alpha &= I_t^\alpha(1) = \int_0^t ... \int_0^{t_{k-1}} 1 dW_{t_k}^{\alpha_k} ... dW_{t_1}^{\alpha_1}. \end{aligned}$$

An important point concerns the scaling property

$$(Scaling-k)\quad W_t^\alpha \sim \varepsilon^{k/2} W_{t\varepsilon}^\alpha \quad with \quad k=p(\alpha).$$

The specific property of W_t^α is that it behaves as $t^{p(\alpha)/2}$ as $t\to 0$. In order to give a precise definition of this fact we introduce some notation.

A function $f:R\to R$ will be called **flat** if $\lim_{\varepsilon\to 0}\varepsilon^{-p}|f(\varepsilon)|=0, \forall p\in N$. The nice thing about this definition is the following simple remark. Let F be a non negative random variable. Then $EF^{-p}<\infty, \forall p\in N$ if and only if the function $\varepsilon\to P(F<\varepsilon)$ is flat.

Moreover we say that a process X has order larger then N and we write $O(X)\geq N$ if $\varepsilon\to P(X_\varepsilon^*>\varepsilon^{N/2-\delta})$ is flat. Here we use the standard notation $X_\varepsilon^*:=\sup_{t\leq\varepsilon}|X_t|$. And we say that $O(X)\leq N$ if $\varepsilon\to P(X_\varepsilon^*<\varepsilon^{N/2+\delta}), \forall\delta>0$ is flat. We say that $O(X)=N$ if $O(X)\leq N$ and $O(X)\geq N$. Notice that we do not assume that we may define $O(X)$ for any process X. We need just inequalities. But when we say that "$O(X)=k$" this means that we are able to prove both inequalityes.

Let us list some basic properties.

Lemma 34. *i) If a process X has the Scaling $-$ k property then $O(X) \leq N$ (respectively $O(X) \geq N$) is equivalent to the fact that $\varepsilon \rightarrow P(\sup_{t\leq 1}|X_t| < \varepsilon)$ is flat (respectively $\varepsilon \rightarrow P(\sup_{t\leq 1}|W_t^{\alpha}| > \varepsilon^{-1})$ is flat).*

ii) Let b be a standard one dimensional Brownian motion. Then $O(b) = 1$

iii) For each X, Y we have

$$O(X) + O(Y) \leq O(X \times Y)$$

in the sense that $O(X) \geq k, O(Y) \geq p \Rightarrow O(X \times Y) \geq k + p$.

iv) If $\|X_{\varepsilon}^*\|_p \leq C_p \varepsilon^k$ for every $p \in N$ and every $\varepsilon > 0$ then $O(X) \geq k$.

v) If $O(X) < O(Y)$ then $O(X+Y) = O(X)$ (in the sense: if $O(X) \leq k < O(Y)$ then $O(X+Y) \leq k$).

Proof. The only non -trivial property is ii). And this property is crucial because it permits to handle all the other processes coming on in the calculus. The fact that $O(b) \geq 1$ is trivial - one employs the scaling property and iv). Let us prove that $O(b) \leq 1$. So we have to check that $\varepsilon \rightarrow P(\sup_{t\leq 1}|b_t| < \varepsilon)$ is flat. We write

$$\begin{aligned} P(b_1^* &< \varepsilon) \leq P(b_{1/2}^* < \varepsilon, \left|b_1 - b_{1/2}\right| < 2\varepsilon) = P(b_{1/2}^* < \varepsilon)P(\left|b_1 - b_{1/2}\right| < 2\varepsilon) \\ &\leq C\varepsilon P(b_{1/2}^* < \varepsilon). \end{aligned}$$

And if we continue like this we obtain $P(b_1^* < \varepsilon) \leq C_p \varepsilon^p$.□

We introduce now a class of stochastic processes which may be developed in stochastic series; We denote

$$C_0 = \{X : \text{ continuous and adapted process such that } E(X_T^*)^p < \infty, \forall p \in N, T > 0\}.$$

Then we define by recurrence C_k to be the class of the stochastic processes of the form

$$(*) \quad X_t = X_0 + \sum_{i=0}^{d} \int_0^t X_s^i dW_s^i$$

with $X^i \in C_{k-1}$. It is easy to check by recurrence that $C_k \subset C_{k-1}$. We denote $C_\infty = \cap_k C_k$.

Moreover if X has the representation $(*)$ we define $\Pr_i(X) = X^i$ and we iterate this definition: for a multi-index $\alpha = (\alpha_1, ..., \alpha_k)$ and $X \in C_\infty$ we define $\Pr_\alpha(X) = \Pr_{\alpha_k}(\Pr_{(\alpha_1,...,\alpha_{k-1})}(X))$. Finally we put $pr_\alpha(X) = \Pr_\alpha(X)(0)$. In order to get unitary notation we put $pr_\emptyset(X) = X_0$.

Let us now take $X \in C_\infty$ and develop it in stochastic series. We write

$$\begin{aligned} X_t &= X_0 + \sum_{i=0}^{d} \int_0^t X_s^i dW_s^i = X_0 + \sum_{i=1}^{d} pr_i(X) W_t^i + \sum_{i=1}^{d} \int_0^t (X_s^i - pr_i(X)) dW_s^i + \int_0^t X_s^0 dW_s^0 \\ &= \sum_{p(\alpha) \le 1} pr_\alpha(X) W_t^\alpha + \sum_{p(\alpha)=1} I_\alpha(\Pr_\alpha(X) - pr_\alpha(X)) + \sum_{p(\alpha)=1-1=0} I_{(\alpha,0)}(\Pr_\alpha(X)). \end{aligned}$$

We wrote the last expression in order to understand what we will obtain in the general case. And we obtain

$$\begin{aligned} X_t &= \sum_{p(\alpha) \le N} pr_\alpha(X) W_t^\alpha + R_N(X)(t) \quad with \quad R_N(X)(t) = R'_N(X)(t) + R''_N(X)(t), \\ R'_N(X)(t) &:= \sum_{p(\alpha)=N} I_\alpha(\Pr_\alpha(X) - pr_\alpha(X)), \quad R''_N(X)(t) := \sum_{p(\alpha)=N-1} I_{(\alpha,0)}(\Pr_\alpha(X)). \end{aligned}$$

The interest of this decomposition is that it permits us to compute $O(X)$. We have:

Proposition 35.

$$O(X) = \min\{p(\alpha) : pr_\alpha(X) \neq 0\}$$

Proof. Suppose that $pr_\alpha(X) = 0$ if $p(\alpha) < N$ and there exists α with $pr_\alpha(X) \neq 0$ and $p(\alpha) = N$. Then we have $X = \sum_{p(\alpha)=N} pr_\alpha(X) W^\alpha + R_N(X)$. It is easy to check that $\|R_N^*(X)(\varepsilon)\|_p \le C_p \varepsilon^{(N+1)/2}$ for every $p \in N$ so that $O(R_N(X)) \ge N+1$. So it suffice to prove that $O(\sum_{p(\alpha)=N} pr_\alpha(X) W^\alpha) = N$. And by the scaling property this amounts to check that $\varepsilon \to P((\sum_{p(\alpha)=N} pr_\alpha(X) W^\alpha)(1) \le \varepsilon)$ is flat. And this is a consequence of (7) in the appendix. □

The followning lemma which we prove in the appendix, is crucial.

Lemma 36. *Let*

$$\begin{aligned} D_q(c)(s,t) &= \sum_{p(\alpha)+p(\beta)=q} c_{\alpha,\beta} W_s^\alpha W_t^\beta \quad and \\ Y_q(t) &= \int_0^t |D_q(c)(s,t)|^2 \, ds. \end{aligned}$$

with $c_{\alpha,\beta} \in R$. If there exists at least one $c_{\alpha,\beta} \neq 0$ then $O(Y) = 2q+2$.

As a consequence we obtain

Proposition 37. *Let* $U, V \in C_\infty^n$ *and* $Y_t := \int_0^t \langle U_s, V_t \rangle^2 ds$. *Then* $O(Y) = 2q + 2$ *where* $q = \min\{p(\alpha) + p(\beta) : \langle pr_\alpha(U), pr_\beta(V) \rangle \neq 0\}$.

Proof. Take $N = q + 1$ and write

$$\begin{aligned} U &= S_N(U) + R_N(U), \quad S_N(U) = \sum_{p(\alpha) \leq N} pr_\alpha(U) W^\alpha \\ V &= S_N(V) + R_N(V), \quad S_N(V) = \sum_{p(\alpha) \leq N} pr_\alpha(V) W^\alpha. \end{aligned}$$

Then

$$\begin{aligned} \langle U_s, V_t \rangle &= \langle S_N(U)(s), S_N(V)(t) \rangle + R_N(s,t) \quad \textit{with} \\ R_N(s,t) &= \langle S_N(U)(s), R_N(V)(t) \rangle + \langle R_N(U)(s), S_N(V)(t) \rangle + \langle R_N(U)(s), R_N(V)(t) \rangle. \end{aligned}$$

It is easy to see that $O(\int_0^t R_N^2(s,t)ds) \geq 2(N+1) + 2 > 2q + 2$ so we may ignore this term. Moreover we write

$$\begin{aligned} & \langle S_N(U)(s), S_N(V)(t) \rangle \\ = & \sum_{p(\alpha) \leq N} \sum_{p(\beta) \leq N} \langle pr_\alpha(U), pr_\beta(V) \rangle W_s^\alpha W_t^\beta \\ = & \sum_{p(\alpha)+p(\beta)=q} \langle pr_\alpha(U), pr_\beta(V) \rangle W_s^\alpha W_t^\beta + \sum_{p(\alpha)+p(\beta)>q} \langle pr_\alpha(U), pr_\beta(V) \rangle W_s^\alpha W_t^\beta \\ = & : D_q(s,t) + D'_q(s,t). \end{aligned}$$

We have $O(\int_0^t \left| D'_q(s,t) \right|^2 ds) \geq 2(q+1) + 2 > 2q + 2$ so we ignore this term also. And by the previous lemma $O(\int_0^t \left| D_q(s,t) \right|^2 ds) = 2q + 2$ so we are done. □

We consider now a matrix valued process $U \in M_{n \times m}(C_\infty)$ and we denote by $U_j, j = 1, ..., m$ the columns of U, that is $U_j = (U_j^1, ..., U_j^n) \in C_\infty^n$. We define

$$E_k = Span\{pr_\alpha(U_j) : j = 1, ..., m, p(\alpha) \leq k\} \subset R^n.$$

And we consider the matrix

$$A_U(t) = \int_0^t U_s U_s^* ds.$$

Lemma 38. *The following are equivalent:*

$$\begin{array}{rcl} i) \quad E_q & = & R^n \\ ii) \quad O(\langle A_U \times X, X\rangle) & \leq & 2q+2+2\min_{i=1,\ldots,n} O(X_i), \quad \forall X=(X_1,...,X_n)\in C_\infty^n. \end{array}$$

Proof. Suppose that $E_q = R^n$. Take $X \in C_\infty^n$ and let $p = \min_{i=1,\ldots,n} O(X_i)$. Then there exists some β with $p(\beta) = p$ such that $pr_\beta(X) \neq 0$. And then there exist some $j \in \{1,...,m\}$ and some α with $p(\alpha) \leq q$ such that $\langle pr_\alpha(U_j), pr_\beta(X)\rangle \neq 0$. We have

$$\langle A_U(t) \times X_t, X_t\rangle = \sum_{k=1}^{m} \int_0^t \langle U_k(s), X_t\rangle^2 \, ds$$

so that, by the previous lemma

$$O(\langle A_U(t) \times X_t, X_t\rangle) \leq O(\int_0^t \langle U_j(s), X_t\rangle^2 \, ds) \leq 2(p(\alpha)+p(\beta))+2 \leq 2(q+p)+2$$

so $ii)$ is proved.

Suppose now that $E_q \subset R^n$. Then we take $x \in R^n$ with $x \neq 0$. We have $\min_{i=1,\ldots,n} O(x_i) = 0$ so that, if $ii)$ is true, then $O(\langle A_U \times x, x\rangle) \leq 2q+2$. But if we look to the coefficients of the development in stochastic series we see that all the coefficients with $p(\alpha) \leq q$ are null. Then $O(\langle A_U \times x, x\rangle) \geq 2(q+1)+2 > 2q+2.\square$

We need also the following algebra lemma.

Lemma 39. *Let I be a comutative ring and let $O : I \to R_+ \cup \{\infty\}$ be a function such that $O(0) = \infty, O(x) = O(-x)$ and*

$$(a) \quad O(x) + O(y) \leq O(xy), \forall x, y \in I.$$

Consider a matrix $a \in M_{n\times n}(I)$ which verifies

$$(b) \quad \langle ax, x\rangle \leq k + 2\min_{i=1,\ldots,n} O(x_i), \quad \forall x \in I^n.$$

Then $O(\det a) \leq nk$.

Proof. We procced by recurrence on n. Suppose first that $n = 1$. Then $a \in I$ and by (a) first and then by (b)

$$O(a) + 2O(x) \leq O(axx) = O(\langle ax, x\rangle) \leq k + 2O(x).$$

So $O(a) \leq k$.

We suppose now that the assertion is true for $n-1$ and we prove it for n. We denote $A^i_j = (a^p_k)_{k\neq j, p\neq i}$ and $\Gamma^i_j = (-1)^{i+j} \det A^i_j$. We also denote $\widehat{a} = (\Gamma^i_j)_{i,j=1,n}$. And we have $a \times \widehat{a} = \det a \times Id$.

We prove now that $A^n_n \in M_{n-1,n-1}$ satisfies (b). We take $x = (x_1, ..., x_{n-1}) \in I^{n-1}$ and we denote $\widehat{x} = (x_1, ..., x_{n-1}, 0) \in I^n$. Then $\langle A^n_n x, x\rangle = \langle a\widehat{x}, \widehat{x}\rangle$ so that

$$O(\langle A^n_n x, x\rangle) = O(\langle a\widehat{x}, \widehat{x}\rangle) \leq k + \min_{j=1,\dots,n-1} O(x_i)$$

the last inequality being a consequence of (b) for a and of the fact that $O(0) = \infty$. So be the recurrence hypothesis we get $O(\det A^n_n) \leq k(n-1)$.

We take now $x = (0, ..., 0, 1) \in I^n$ and we denote $y = \widehat{a}x = (\Gamma^1_n, ..., \Gamma^n_n)$. Then

$$\langle ay, y\rangle = \langle a(\widehat{a}x), \widehat{a}x\rangle = \langle \det(a)x, \widehat{a}x\rangle = \det(a)\Gamma^n_n.$$

Using (a)

$$\begin{aligned} O(\det(a)) + O(\Gamma^n_n) &\leq O(\det(a)\Gamma^n_n) = O(\langle ay, y\rangle) \leq k + 2\min_{i=1,\dots,n} O(y_i) \\ &\leq k + 2O(y_n) = k + O(\Gamma^n_n). \end{aligned}$$

And then $O(\det(a)) \leq k + O(\Gamma^n_n) = k + O(A^n_n) \leq k + k(n-1) = kn.\Box$

Proposition 40. *Let $h(U) = \min\{k : E_k = R^n\}$ and $h(U) = \infty$ is the set is void. Then*

$$O(\det A_U) \leq (2h(U) + 1)n.$$

Proof. Suppose that $h(U) = q < \infty$. Then $E_q = R^n$ and then by the previous lemma $O(\langle A_U X, X\rangle) \leq 2q+2+2\min_{i=1,\dots,n} O(X_i)$. Now, using the algebra lemma we get $O(\det A_U) \leq (2q+1)n.\Box$

Corollary 41. *If $h(U) < \infty$ then*

$$E(|\det A_U(T)|^{-p}) < \infty, \quad \forall p \in N.$$

Proof. Let $\lambda_U(t)$ be the smallest eigenvalue of $A_U(t)$. Since $\det(A_U(t)) = \lambda_U(t) \times V_t$ for some process $V \in C_0$ we have $O(\lambda_U) \leq O(\det(A_U)) < \infty$. It follows that there exists q such that $\varepsilon \to P(\lambda^*_U(\varepsilon) \leq \varepsilon^q)$ is flat. It follows that $\varepsilon \to P(\lambda^*_U(T) \leq \varepsilon^q)$ is also flat and since $t \to \Lambda_U(t)$ is non decreasing we get that $\varepsilon \to P(\lambda_U(T) \leq \varepsilon^q)$ is flat. And this implies $E(|\lambda_U(T)|^{-p}) < \infty, \forall p \in N$. Since $\det A_U(T) \geq \lambda^n_U(T)$ the proof is compleated. $\Box$

0.5.1. Appendix.

We give now a several of lemmas which precise the asymptotic behavior of linear combinations of iterated stochastic integrals. In order to do it we need some more notation concerning multi-indexes. Let $\alpha = (\alpha_1, ..., \alpha_k) \in \{0, ..., d\}^k$. We denote

$$\overleftarrow{\alpha} = (\alpha_1, ..., \alpha_{k-1}) \quad and \quad \overrightarrow{\alpha} = \alpha_k.$$

Moreover we denote

$$\widehat{\alpha} = (\alpha_k, \alpha_{k-1}, ..., \alpha_1, 0, 0, ...).$$

On the set of the multi-indexes we consider the following order relation. Let $\alpha = (\alpha_1, ..., \alpha_k)$ and $\beta = (\beta_1, ..., \beta_p)$. We say that $\alpha \succeq \beta$ if the following is true. Case 1. We have $p(\alpha) < p(\beta)$. Case 2. We have $p(\alpha) = p(\beta)$ but $\widehat{\alpha}$ is larger than $\widehat{\beta}$ in the lexicographic order. This means that there exists some $i \in N$ such that $\widehat{\alpha}_j = \widehat{\beta}_j$ for $1 \leq j \leq i-1$ and $\widehat{\alpha}_i > \widehat{\beta}_i$. It is easy to see that this relation is complete, that is: for each α and β we have $\alpha \succeq \beta$ or $\beta \succeq \alpha$. And if the two inequalities hold true then $\alpha = \beta$. The idea of this definition is the following: we prefer multi-indexes with small order. And if we have the same order we prefer the one which has larger components. But we begin to compare them starting from the last component.

Moreover we denote by Λ the set of all the multi-indexes and we consider a set of numbers $c = (c_\alpha)_{\alpha \in \Lambda}$. We denote $|c| = (\sum_{\alpha \in \Lambda} |c_\alpha|^2)^{1/2}$. Clearly $|c| < \infty$ if and only if c has a finite number of non null components. For c with $|c| < \infty$ we define $c_{\max}$ in the following way. We define $\alpha_{\max}(c) = \max\{\alpha : \alpha \in \Lambda, c_\alpha \neq 0\}$. The maximum is taken with respect to the order defined above and since this order relation is complete and the set is finite, there exists a unique element which is strictly larger then all the other ones. So $\alpha_{\max}(c)$ is well defined. Finally we put

$$c_{\max} = \left| c_{\alpha_{\max}(c)} \right|.$$

Lemma 42. *A. Let $S_N(c,t) = \sum_{p(\alpha) \leq N} c_\alpha W_t^\alpha$. For every $p \in N$, there exists some universal constants $C = C_{p,N}, q = q_{p,N}$ such that*

$$\begin{aligned} i) \quad (E(|S_N^*(c,1)|^p))^{1/p} &\leq C \times |c| && (7) \\ ii) \quad (E(\sup_{\tau \leq t \leq \tau+\varepsilon} |S_N(c,t) - S_N(c,\tau)|^p))^{1/p} &\leq C \times |c| \times \varepsilon^{1/2} && \forall \tau \text{ stopping time.} \end{aligned}$$

B. Let $Q_N(c,t) = \sum_{p(\alpha)=N} c_\alpha W_t^\alpha$. For every $p \in N$, there exist some universal *constants* $C = C_{p,N}, q = q_{p,N}$ *such that*

$$iii) \quad P(Q_N^*(c,1) \ < \ \varepsilon) \leq C(|c|^q + c_{\max}^{-8^{N-1}p}) \times \varepsilon^p \quad and \tag{8}$$
$$iv) \quad P(\int_0^1 |Q_N(c,s)|^2 \, ds \ < \ \varepsilon) \leq C(|c|^q + c_{\max}^{-8^N p}) \times \varepsilon^p.$$

Proof. The properties $i), ii)$ are standard. Let us prove $iii)$ and $iv)$ by recurrence. If $N = 1$ then $Q_1(c,t) = \sum_{i=1}^d c_i W_t^i$ has the same law as $|c| \, b_t$ where b is a standard Brownian motion. So $P(Q_1^*(c,1) < \varepsilon) = P(b_1^* \leq \varepsilon / |c|) \leq C_p(\varepsilon / |c|)^p \leq C_p(\varepsilon / c_{\max})^p$. We treat now the second term. We take $u = \frac{1}{8}$ and we define

$$T_\varepsilon = \inf\{t : |Q_1(c,s)| \geq 2\varepsilon^u\} \quad and \quad T'_\varepsilon = \inf\{t \geq T_\varepsilon : |Q_1(c,s)| < \varepsilon^u\}.$$

Then we take $v = \frac{1}{2}$ and we write

$$\begin{aligned} P(\int_0^1 |Q_1(c,s)|^2 \, ds \ &< \ \varepsilon) \leq P(T_\varepsilon > \frac{1}{2}) + P(T_\varepsilon \leq \frac{1}{2}, T'_\varepsilon - T_\varepsilon \leq \varepsilon^v) \\ +P(T_\varepsilon \ &\leq \ \frac{1}{2}, T'_\varepsilon - T_\varepsilon > \varepsilon^v, \int_0^1 |Q_N(c,s)|^2 \, ds < \varepsilon) =: \sum_{i=1}^3 I_\varepsilon^i. \end{aligned}$$

We take $p' = 8p$ and we write

$$\begin{aligned} P(T_\varepsilon \ &> \ \frac{1}{2}) = P(Q_1^*(c, \frac{1}{2}) < \varepsilon^u) = P(Q_1^*(c,1) < \varepsilon\sqrt{2}) \leq C_{p'}(\varepsilon^u / c_{\max})^{p'} \\ &= \ \frac{C_{p'}}{c_{\max}^{8p}} \times \varepsilon^p. \end{aligned}$$

Since $|Q_1(c,T_\varepsilon)| - |Q_1(c,T'_\varepsilon)| = 2\varepsilon^u - \varepsilon^u = \varepsilon^u$ we have

$$\begin{aligned} P(T_\varepsilon \ &\leq \ \frac{1}{2}, T'_\varepsilon - T_\varepsilon \leq \varepsilon^v) \leq P(\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |Q_1(c,t) - Q_1(c,T_\varepsilon)| \geq \varepsilon^u) \\ &\leq \ \frac{1}{\varepsilon^{p'u}} E(\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |Q_1(c,t) - Q_1(c,T_\varepsilon)|^{p'}) \leq \frac{1}{\varepsilon^{p'u}} \times |c|^{p'} \varepsilon^{p'v/2} = |c|^{8p} \varepsilon^p. \end{aligned}$$

For $T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v \leq T'_\varepsilon$, we have $|Q_N(c,s)| \geq \varepsilon^u$ so that

$$\int_0^1 |Q_N(c,s)|^2 \, ds \geq \int_{T_\varepsilon}^{T_\varepsilon + \varepsilon^v} |Q_N(c,s)|^2 \, ds \geq \varepsilon^{v+2u} > \varepsilon$$

so that $I_\varepsilon^3 = 0$.

Suppose now that the property holds true for $N-1$. Let $c^i, i = 0, ..., d$ be the vector of the coefficients corresponding to multi-indexes $\alpha = (\alpha_1, ..., \alpha_k)$ with $\alpha_k = i$, that is: $c_\alpha^i = c_\alpha$ is $\overrightarrow{\alpha} = i$ and $c_\alpha^i = 0$ is $\overrightarrow{\alpha} \neq i$. Then we have

$$\begin{aligned} Q_{N-1}(c^i, s) &= \sum_{p(\alpha)=N, \overrightarrow{\alpha}=i} c_\alpha W_s^{\overleftarrow{\alpha}}, \quad for \quad i = 1, ..., d, \\ Q_{N-1}(c^0, s) &= \sum_{p(\alpha)=N, \overrightarrow{\alpha}=0} c_\alpha W_s^{\overleftarrow{\alpha}}. \end{aligned}$$

Then

$$Q_N(c,t) = \sum_{i=1}^d \int_0^t Q_{N-1}(c^i, s) dW_s^i + \int_0^t Q_{N-2}(c^0, s) ds =: m_t + d_t.$$

Case 1. Suppose that there exists α with $\overrightarrow{\alpha} = i \in \{1, ..., d\}$ such that $c_\alpha \neq 0$. We take $i^* = \max\{\overrightarrow{\alpha} : c_\alpha \neq 0, \overrightarrow{\alpha} \in \{1, ..., d\}\}$. Notice that we have $\alpha_{\max}(c) = \alpha_{\max}(c^{i^*})$ and consequently $c_{\max} = c_{\max}^{i^*}$. We take $u = \frac{1}{8}$ and we define

$$T_\varepsilon = \inf\{t : \left|Q_{N-1}(c^{i^*}, s)\right| \geq 2\varepsilon^u\} \quad and \quad T'_\varepsilon = \inf\{t \geq T_\varepsilon : \left|Q_{N-1}(c^{i^*}, s)\right| < \varepsilon^u\}.$$

We take $p' = 8p$ and we use the recurrence hypothesis in order to obtain

$$\begin{aligned} P(T_\varepsilon &> \frac{1}{2}) = P(Q_{N-1}^*(c^{i^*}, \frac{1}{2}) < \varepsilon^u) \leq C(\left|c^{i^*}\right|^q + \frac{1}{(c_{\max}^{i^*})^{8^{N-2}p'}})\varepsilon^{up'} \\ &\leq C(|c|^q + \frac{1}{(c_{\max})^{8^{N-1}p}})\varepsilon^p. \end{aligned}$$

Moreover, we take $v = \frac{3}{2}$ and $p'' = \frac{8p}{5}$ and we use the same reasoning as above and obtain

$$\begin{aligned} P(T_\varepsilon &\leq \frac{1}{2}, T'_\varepsilon - T_\varepsilon \leq \varepsilon^v) \leq P(\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} \left|Q_{N-1}(c^{i^*}, t) - Q_{N-1}(c^{i^*}, T_\varepsilon)\right| \geq \varepsilon^u) \\ &\leq \frac{1}{\varepsilon^{p''u}} E(\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} \left|Q_{N-1}(c^{i^*}, t) - Q_{N-1}(c^{i^*}, T_\varepsilon)\right|^{p''}) \leq \frac{C}{\varepsilon^{p''u}} \times \left|c^{i^*}\right|^{p''} \varepsilon^{p''v/2} = C\left|c\right|^{p''} \varepsilon^p. \end{aligned}$$

Finally we take $r = \frac{1}{2}$ and $p''' = 2p$ we write

$$P(Q_{N-2}^*(c^0, 1) \geq \varepsilon^{-r}) \leq C\left|c^0\right|^{p'''} \varepsilon^{p'''r} \leq C\left|c\right|^{p'''} \varepsilon^p.$$

We conclude that

$$
\begin{aligned}
P(Q_N^*(c,1) &< \varepsilon) \leq C(|c|^q + \frac{1}{(c_{\max})^{8^{N-1}p}})\varepsilon^p + I_\varepsilon \quad with \\
I_\varepsilon &= P(T_\varepsilon \leq \frac{1}{2}, T'_\varepsilon - T_\varepsilon > \varepsilon^v, Q_{N-2}^*(c^0,1) < \varepsilon^{-r}, Q_N^*(c,1) < \varepsilon).
\end{aligned}
$$

We take $T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v$ and we write

$$
\begin{aligned}
|Q_N(c,t)| &\geq |Q_N(c,t) - Q_N(c,T_\varepsilon)| - |Q_N(c,T_\varepsilon)| \\
&\geq |m_t - m_{T_\varepsilon}| - |d_t - d_{T_\varepsilon}| - |Q_N(c,T_\varepsilon)|
\end{aligned}
$$

which gives

$$
\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |m_t - m_{T_\varepsilon}| \leq 2Q_N^*(c,1) + \sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |d_t - d_{T_\varepsilon}|.
$$

If we are on the set $Q_{N-2}^*(c^0,1) < \varepsilon^{-r}$ then $\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |d_t - d_{T_\varepsilon}| \leq \varepsilon^{v-r} = \varepsilon$. And if we also have $Q_N^*(c,1) < \varepsilon$ the above inequality becomes $\sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |m_t - m_{T_\varepsilon}| \leq 3\varepsilon$. So

$$
I_\varepsilon \leq P(T_\varepsilon \leq \frac{1}{2}, T'_\varepsilon - T_\varepsilon > \varepsilon^v, \sup_{T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v} |m_t - m_{T_\varepsilon}| \leq 3\varepsilon).
$$

We write now $m(t) = b(\langle m \rangle_t)$ where b is a standard Brownian motion. Notice that

$$
\langle m \rangle_{T_\varepsilon + \varepsilon^v} - \langle m \rangle_{T_\varepsilon} = \int_{T_\varepsilon}^{T_\varepsilon + \varepsilon^v} \sum_{i=1}^d Q_{N-1}^2(c^i,s)ds \geq \int_{T_\varepsilon}^{T_\varepsilon + \varepsilon^v} Q_{N-1}^2(c^{i^*},s)ds \geq \varepsilon^{v+2u}
$$

the last inequality being true because $Q_{N-1}^2(c^{i^*},s) > \varepsilon^{2u}$ for $T_\varepsilon < s < T_\varepsilon + \varepsilon^v \leq T'_\varepsilon$. This gives

$$
I_\varepsilon \leq P(\sup_{0<t<\varepsilon^{v+2u}} |b(t)| < \varepsilon) = P(b^*(1) < \varepsilon^{1-(u+v/2)}) = P(b^*(1) < \varepsilon^{1/8}) \leq C_p \varepsilon^p
$$

the last inequality being a consequence of the basic property of the Brownian motion. So the proof of $iii)$ is completed in this case.

Case 2. We have $c_\alpha = 0$ for any α with $\overrightarrow{\alpha} \neq 0$. Then

$$
Q_N(c,t) = \int_0^t Q_{N-2}(c^0,s)ds.
$$

Notice that $c_{\max} = c^0_{\max}$. We take $u = \frac{1}{8}$ and we define

$$T_\varepsilon = \inf\{t : \left|Q_{N-2}(c^0, s)\right| \geq 2\varepsilon^u\} \quad \textit{and} \quad T'_\varepsilon = \inf\{t \geq T_\varepsilon : \left|Q_{N-2}(c^0, s)\right| < \varepsilon^u\}.$$

We follow now the same strategy as above. Using the recurrence hypothesis $iii)$ with $p' = 8p$ we obtain

$$P(T_\varepsilon > \frac{1}{2}) \leq C(\left|c^0\right|^q + \frac{1}{c_{\max}^{p'8^{N-3}}})\varepsilon^{up'} \leq C(\left|c^0\right|^q + \frac{1}{c_{\max}^{p8^{N-2}}})\varepsilon^p.$$

We take $v = \frac{1}{2}$ and $p'' = 8p$. Then, the same argument as above gives

$$P(T_\varepsilon \leq \frac{1}{2}, T'_\varepsilon - T_\varepsilon \leq \varepsilon^v) \leq C\left|c\right|^q \varepsilon^{p''(\frac{v}{2}-u)} = C\left|c\right|^q \varepsilon^p.$$

It follows that

$$\begin{aligned} P(Q^*_N(c,1) &< \varepsilon) \leq C(\left|c^0\right|^q + \frac{1}{c_{\max}^{p8^{N-2}}})\varepsilon^p \\ +P(T_\varepsilon &\leq \frac{1}{2}, T'_\varepsilon - T_\varepsilon > \varepsilon^v, Q^*_N(c,1) < \varepsilon). \end{aligned}$$

Notice that for $T_\varepsilon \leq t \leq T_\varepsilon + \varepsilon^v \leq T'_\varepsilon$, the function $t \to Q_{N-2}(c^0, t)$ keeps the same sign and $|Q_{N-2}(c^0, s)| > \varepsilon^u$. It follows that

$$\left|\int_{T_\varepsilon}^{T_\varepsilon+\varepsilon^v} Q_{N-2}(c^0, t)dt\right| = \int_{T_\varepsilon}^{T_\varepsilon+\varepsilon^v} \left|Q_{N-2}(c^0, t)\right| dt \geq \varepsilon^{v+u} = \varepsilon^{3/4}.$$

But we also have

$$\left|\int_{T_\varepsilon}^{T_\varepsilon+\varepsilon^v} Q_{N-2}(c^0, t)dt\right| \leq \left|\int_0^{T_\varepsilon+\varepsilon^v} Q_{N-2}(c^0, t)dt\right| + \left|\int_0^{T_\varepsilon} Q_{N-2}(c^0, t)dt\right| \leq 2\varepsilon.$$

Since $2\varepsilon < \varepsilon^{3/4}$ this is impossible so the proof finishes.

Let us now prove $iv)$. The proof is identical with the one in the case $N = 1$ so we skip it. □

We consider now the double sum

$$D_k(c, s, t) = \sum_{p(\alpha)+p(\beta)=k} c_{\alpha,\beta} W^\alpha_s W^\beta_t, \quad 0 \leq s \leq t \leq 1.$$

We define

$$\begin{aligned} \overline{\alpha}_{\max}(c) &= \max\{\alpha : \exists\beta \text{ such that } c_{\alpha,\beta} \neq 0\}, \quad \overline{\beta}_{\max}(c) = \max\{\beta : c_{\overline{\alpha}_{\max}(c),\beta} \neq 0\}, \\ c_{\max} &= c_{\overline{\alpha}_{\max}(c),\overline{\beta}_{\max}(c)}. \end{aligned}$$

Lemma 43. *For every $p, k \in N$, there exists some universal constants $C = C_{p,k}$ and $q = q_{p,k}$ such that*

$$P(\sup_{t \leq 1} \int_0^t D_k^2(c, s, t) ds \leq \varepsilon) \leq C(|c|^q + c_{\max}^{-p8^k}) \times \varepsilon^p. \tag{9}$$

Proof. For every $l, r \in \{1, ..., k\}$ we define

$$Q_{l,r}(c, s, t) = \sum_{p(\alpha)=l} \sum_{p(\beta)=r} c_{\alpha,\beta} W_s^\alpha W_t^\beta$$

with the convention $c_{\alpha,\beta} = 0$ if $p(\alpha) + p(\beta) \neq k$. We take $N = p(\alpha_{\max}(c))$ and we notice that for $p(\alpha) < N$ we have $c_{\alpha,\beta} = 0, \forall \beta$. So $Q_{l,k-l} = 0$ for $l < N$. So $D_k = \sum_{l=N}^k Q_{l,k-l}$. We also denote $M = k - N$.

We take $\delta = \varepsilon^v$ with $v > 0$ (to be chosen latter on) and for $\beta = (\beta_1, ..., \beta_k)$ we define

$$W_t^{\beta,\delta} = \int_0^t .... \int_0^{t_{k-1}} 1_{\{t_k > \delta\}} dW_{t_k}^{\beta_k} ... dW_{t_1}^{\beta_1}, \quad \overline{W}_t^{\beta,\delta} = \int_0^t .... \int_0^{t_{k-1}} 1_{\{t_k \leq \delta\}} dW_{t_k}^{\beta_k} ... dW_{t_1}^{\beta_1}$$

so that $W_t^\beta = W_t^{\beta,\delta} + \overline{W}_t^{\beta,\delta}$. Moreover, for α with $p(\alpha) = N$ we denote (recall that $M = k - N$)

$$d_\alpha^\delta(t) = \sum_{p(\beta)=M} c_{\alpha,\beta} W_t^{\beta,\delta}, \quad \overline{d}_\alpha^\delta(t) = \sum_{p(\beta)=M} c_{\alpha,\beta} \overline{W}_t^{\beta,\delta}$$

so that $Q_{N,M}(c, s, t) = Q_{N,M}^\delta(c, s, t) + \overline{Q}_{N,M}^\delta(c, s, t)$ with

$$Q_{N,M}^\delta(c, s, t) = \sum_{p(\alpha)=N} d_\alpha^\delta(t) W_s^\alpha \quad \textit{and} \quad \overline{Q}_{N,M}^\delta(c, s, t) = \sum_{p(\alpha)=N} \overline{d}_\alpha^\delta(t) W_s^\alpha.$$

Then, for $t > \delta$ we write (we use the elementary inequalities $(a + b)^2 \geq \frac{1}{2} a^2$ and $(a + b)^2 \leq 2a^2 + 2b^2$)

$$\begin{aligned} \int_0^t D_k^2(c, s, t) ds &\geq \int_0^\delta D_k^2(c, s, t) ds \\ &\geq \frac{1}{2} \int_0^\delta \left| Q_{N,M}^\delta(c, s, t) \right|^2 ds - C \int_0^\delta \left| \overline{Q}_{N,M}^\delta(c, s, t) \right|^2 ds \\ &\quad - C \sum_{l=N+1}^k \int_0^\delta \left| Q_{l,k-l}(c, s, t) \right|^2 ds \end{aligned}$$

where C is an universal constant. We obtain

$$P(\sup_{t\leq 1}\int_0^t D_k^2(c,s,t)ds\leq\varepsilon)\leq I_{\varepsilon,\delta}+\overline{I}_{\varepsilon,\delta}+\sum_{l=N+1}^{k}J_\varepsilon^l$$

with

$$\begin{aligned}
I_{\varepsilon,\delta} &= P(\sup_{\delta<t\leq 1}\int_0^\delta \left|Q_{N,M}^\delta(c,s,t)\right|^2 ds\leq C\varepsilon),\quad \overline{I}_{\varepsilon,\delta}=P(\sup_{\delta<t\leq 1}\int_0^\delta \left|\overline{Q}_{N,M}^\delta(c,s,t)\right|^2 ds\geq\varepsilon),\\
J_\varepsilon^l &= P(\sup_{0\leq t\leq 1}\int_0^\delta \left|Q_{l,k-l}(c,s,t)\right|^2 ds\geq\varepsilon).
\end{aligned}$$

Let us estimate $I_{\varepsilon,\delta}$. Notice that $d_\alpha^\delta(t)$ is measurable with respect to $F_{\delta,1}:=\sigma\{W_s-W_\delta,\delta<s\leq 1\}$. So it is independent of $W_s,s\leq\delta$. So

$$Q_{N,M}^\delta(c,s,t)=\sum_{p(\alpha)=N}d_\alpha^\delta(t)W_s^\alpha=Q_N(d^\delta(t),s)$$

is a quantity which is similar to the one in the previous lemma but which has random coefficients which are independent of W_s^α. Let $\alpha^*:=\overline{\alpha}_{\max}(c)$. We take $u>0$ and we define

$$T_\varepsilon^\delta=\inf\{t>\delta:\left|d_{\alpha^*}^\delta(t)\right|\geq\varepsilon^u\}.$$

This random time is $F_{\delta,1}$ measurable so it is independent of $W_s,s\leq\delta$. We write

$$I_{d,\varepsilon}\leq P(T_\varepsilon^\delta>\frac{1}{2})+P(T_\varepsilon^\delta\leq\frac{1}{2},\int_0^\delta\left|Q_N(d^\delta(T_\varepsilon^\delta),s)\right|^2ds\leq C\varepsilon)=:I'_{d,\varepsilon}+I''_{d,\varepsilon}.$$

We treat first

$$I'_{d,\varepsilon}=P(\sup_{\delta<t\leq 1/2}\left|d_{\alpha^*}^\delta(t)\right|<\varepsilon^u)=P(\sup_{0<t\leq 1/2-\delta}\left|Q_M(c^*,t)\right|<\varepsilon^u)$$

with $c^*=(c_{\alpha^*,\beta})_{p(\beta)=M}$. The equality is true because $W_t^{\beta,\delta}$ has the same law as $W_{t-\delta}^\beta$. By definition we have $\alpha_{\max}(c^*)=\overline{\beta}_{\max}(c)$ so that $c^*_{\max}=c_{\max}$. We take now $p'=p/u$ and we use the point $iii)$ of the previous lemma and we obtain

$$I'_{d,\varepsilon}\leq C(|c^*|^q+(c^*_{\max})^{-8^{M-1}p'})\times\varepsilon^{up'}\leq C(|c|^q+c_{\max}^{-8^{M-1}p/u})\times\varepsilon^p.$$

We estimate now $I''_{d,\varepsilon}$. We write

$$I''_{d,\varepsilon}=E(1_{\{T_\varepsilon^\delta\leq\frac{1}{2}\}}P(\int_0^\delta |Q_N(d,s)|^2\,ds\leq C\varepsilon \mid d^\delta(T_\varepsilon^\delta)=d)).$$

Notice that α^* is the largest multi-index for which there exists β with $c_{\alpha^*,\beta}\neq 0$. So for $\alpha\succ\alpha^*$ we have $c_{\alpha,\beta}=0,\forall\beta$ and consequently $d_\alpha^\delta(T_\varepsilon^\delta)=0$. So $\alpha^*\succeq\alpha_{\max}(d^\delta(T_\varepsilon^\delta))$. By the definition of T_ε^δ we have $\left|d_{\alpha^*}^\delta(t)\right|=\varepsilon^u>0$ so we conclude that $\alpha^*=\alpha_{\max}(d^\delta(T_\varepsilon^\delta))$. And consequently if $d=d^\delta(T_\varepsilon^\delta)$ then $d_{\max}=d_{\alpha^*}^\delta(T_\varepsilon^\delta)=\varepsilon^u$. We take some $p''\in N$ and we use the scaling property and the point $iv)$ from the previous lemma and we obtain

$$\begin{aligned}P(\int_0^\delta |Q_N(d,s)|^2\,ds &\leq C\varepsilon\mid d^\delta(T_\varepsilon^\delta)=d)=P(\delta^{N+1}\int_0^1|Q_N(d,s)|^2\,ds\leq C\varepsilon\mid d^\delta(T_\varepsilon^\delta)=d)\\ &\leq C(|d|^q+\varepsilon^{-8^Nu})(\varepsilon\delta^{-(N+1)})^{p''}.\end{aligned}$$

Since $E(\left|d^\delta(T_\varepsilon^\delta)\right|^q)\leq C'\,|c|^{q'}$ we obtain

$$I''_{d,\varepsilon}\leq C(|c|^q+1)\times\varepsilon^{-8^Nu}\times(\varepsilon\delta^{-(N+1)})^{p''}.$$

We estimate now $\overline{I}_{\varepsilon,\delta}$. We write

$$\int_0^\delta\left|\overline{Q}_{N,M}^\delta(c,s,t)\right|^2ds\leq C\,|c|^2\times\delta\times\sum_{p(\alpha)=N}\sum_{p(\beta)=M}\sup_{s<\delta}|W_s^\alpha|^2\times\sup_{t\leq 1}\left|\overline{W}_t^{\beta,\delta}\right|^2$$

so that

$$\begin{aligned}\overline{I}_{\varepsilon,\delta} &= P(\sup_{\delta<t\leq 1}\int_0^\delta\left|\overline{Q}_{N,M}^\delta(c,s,t)\right|^2ds\geq\varepsilon)\\ &\leq \max_{p(\alpha)=N}\max_{p(\beta)=M}P(\sup_{s<\delta}|W_s^\alpha|^2\times\sup_{t\leq 1}\left|\overline{W}_t^{\beta,\delta}\right|^2\geq\frac{\varepsilon}{C\,|c|^2\,\delta}).\end{aligned}$$

We have

$$(E(\sup_{s<\delta}\left|\overline{W}_t^\alpha\right|^{4p'''}))^{1/2}\leq C\delta^{Np'''},\quad (E(\sup_{t\leq 1}\left|\overline{W}_t^{\beta,\delta}\right|^{4p'''}))^{1/2}\leq C\delta^{p'''}$$

so, using Chebyshev inequality first and Schwarz inequality then we obtain

$$\overline{I}_{\varepsilon,\delta}\leq(\frac{C\,|c|^2\,\delta}{\varepsilon}\times\delta^{N+1})^{p'''}.$$

Let us now estimate J_ε^k. We have

$$\int_0^\delta |Q_{l,k-l}(c,s,t)|^2\,ds \le C\,|c|^2 \times \delta \times \sum_{p(\alpha)=l}\ \sum_{p(\beta)=k-l} \sup_{s<\delta} |W_s^\alpha|^2 \times \sup_{t\le 1} \left|W_t^\beta\right|^2.$$

Notice that for $p(\alpha) = l \ge N+1$ we have

$$(E(\sup_{s<\delta} |W_t^\alpha|^{4p'''}))^{1/2} \le C\delta^{(N+1)p'''}, \quad (E(\sup_{t\le 1} \left|W_t^\beta\right|^{4p'''}))^{1/2} \le C$$

so the same estimation as above gives

$$J_\varepsilon^k \le (\frac{C\,|c|^2\,\delta}{\varepsilon} \times \delta^{N+1})^{p'''}.$$

We choose now our parameters. We take now $\delta = \varepsilon^v$ with

$$v = \frac{1}{2}(\frac{1}{N+1} + \frac{1}{N+2}) = \frac{2N+3}{2(N+1)(N+2)}.$$

Then $v(N+2) - 1 = 1/2(N+1)$ so we take $p''' = 2(N+1)p$ and we obtain

$$\overline{I}_{\varepsilon,\delta} + \sum_{l=N+1}^{k} J_\varepsilon^k \le C\,|c|^{2p'''}\,\varepsilon^{p'''/2(N+1)} = C\,|c|^{4(N+1)p} \times \varepsilon^p.$$

We take now $u = 1$ and $p'' = 2(p + 8^N)(N+2)$. Since $1 - v(N+1) = 1/2(N+2)$ we obtain

$$-8^N u + p''(1 - v(N+1)) = p.$$

This gives

$$I''_{d,\varepsilon} \le C(|c|^q + 1) \times \varepsilon^{-8^N u} \times (\varepsilon\delta^{-(N+1)})^{p''} = C(|c|^q + 1) \times \varepsilon^p$$

and finally

$$I'_{d,\varepsilon} \le C(|c|^q + c_{\max}^{-8^M p/u}) \times \varepsilon^p = C(|c|^q + c_{\max}^{-8^M p}) \times \varepsilon^p.$$

We put all this together and we obtain (9). □

0.6. References

[1] Bally V. On the connection between the Malliavin covarience matrix and Hörmander condition. J. Func. Ann. 96, pp 219-255, (1991).

[2] Bell D. R. The Malliavin calculus. Pitman Monographs and Surveys in Pure and Applied Math. 34, Longman and Wiley (1987)

[3] Hörmander L : Hypoeliptic second order differential equations. Acta math. 119, pp 147-171 (1967)

[4] Ikeda N. and Watanabe S. Stochastic differential equations and diffusion processes. North Holand, second ed. (1989).

[5] Kusuoka R. and Stroock D.W.: Application of the Malliavin calculus II: J. Fac. Sci. Univ. Tokyo Sect IA Math. 32pp. 391-442 (1987).

[6] Malliavin P. Stochastic calculus of variations and hypoelliptic operators. In: Proc. Inter. Symp. on Stoch. Diffe. Equations, Kyoto (1976), Wiley, 313, pp. 195-263.

[7] Meyer P-A. Transformation de Riez pour les lois gaussiennes. In Séminaire de Probabilités XVIII. Lecture Notes in Math. 1059, pp 179-193. Springer Verlag (1984).

[8] Nualart D. Malliavin Calculus and related topics. Springer Verlag (1995).

[9] Ocone D. Malliaivin Calculus and Stochastic Integral Representation of Diffusion Processes. Stochastics and Stochastic Repports 12, pp 161-185 (1984).

[10] Sanz-Solé M. Malliavin cCalculus with applications to Stochastic Differential Equations EPFL Press, Fundamental Sciences, Mathematics, (2006).

[11] Shigekawa I. Stochastic Analysis. Translations of Mathematical Monographs, Volume 224, American Math? Soc. (2004).

[12] Watanabe S. Lectures on stochastic differential equations and Malliavin calculus. Tata institute of fundamental research, Springer Verlag, Berlin (1984).