Internship report

Jérémy Zurcher

Convergence in law in Wiener chaos

under the direction of Ciprian Tudor

2022

Introduction

This report summarizes all the knowledge I had the chance to learn during my internship in Lille, with Ciprian Tudor, during the spring 2022. C. Tudor is known for its breakthroughs about Malliavin calculus combined with Stein's method. Before introducing the subject, I would like to thank him, and his PhD students Charles-Phillipe Diez and Julie Gamain for their mathematical, and human interactions I had with them.

This report contains two parts. The first one introduce the notion of Malliavin derivative, used by the French mathematician Paul Malliavin (1925 - 2010). He used it for studying the regularities of the law of stochastic (partial) differential equations, and applications for mathematical finance. The second part is about Stein's method, introduced by Charles Stein (1920 - 2016) in 1972. The main goal of this method is to give smooth conditions to prove a central limit theorem.

The main idea of this report is to know how to mix those two notions to have conditions to have central limit theorems, better than simply using Stein's method. This idea came from Nourdin and Peccati, and this report is very inspired by their book, *Normal approximations with Malliavin calculus*, published in 2012. Our aim is the very last theorem of this report, Breuer-Major theorem.

The logic of this report is the following one. The Malliavin calculus, which is very inspired by the notion of Gateaux derivative on general Hilbert space, is about differentiating some random variables *with respect to the chance* ω. The difficulty is to differentiate on general probability space $(\Omega, \mathcal{F}, \mathbb{P})$, since we do not have any topological structure on it. Here comes the Gateaux derivative. A map on two real Hilbert spaces is Gateaux differentiable in a point of the Hilbert space if it is in the real differential calculus sense, in every direction. To deal with Hilbert space with our case of random variables, we introduce the notion of *isonormal Gaussian process*, which send every element of a set real separable Hilbert space into the set of centered Gaussian random variables. Inspired by the Gateaux derivative, we can define the notion of derivative with the help of those process. We generalize it by taking the closure of a family of polynomials taken in the isonormal process, which yields to the notion of *Wiener chaos* (named after Norbert Wiener, 1896-1964).

We begin the first part by an example in the case where $(\Omega, \mathcal{F}, \mathbb{P})$ is the space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \gamma)$, where γ is the standard Gaussian measure on $\mathbb{R}$. We will refer it as "*the one dimensional case*". We will see with it why the polynomial family we will use is the Hermite polynomials. Then, we will define in general case what is the derivative in a Malliavin sense. We will introduce its *adjoint*, the divergence operator (in the sense of the operator theory), and we will show that this operator extends the notion of Itô's stochastic integral. Those notions will be linked by the notion of *Ornstein-Ulhenbeck semi group*, which appears naturally when we want to study the law of the solution of some stochastic differential equations. The interpretation of the notions using this semi group is a good key for solving some computations, as we will see in the last section of this part by some applications. This semi-group allows us to prove the *Nelson's hypercontractibity*, stating that the L^p convergence in a set Wiener chaos are all *equivalent*. This part is greatly inspired by Ivan Nourdin and Giovanni Peccati and David Nualart, *The Malliavin calculus and its related topics*, 1995.

The second part presents the Stein's method. The idea is the following : we have a characterization with test functions of the standard Gaussian law, that looks like the distribution theory and the variational representation of some partial differential equations. Here, we want to have a quantitative information about *how far a law is with respect to a Gaussian measure*. To do this, we introduce different distances on probability measure that respect the law convergence. The Stein's method consists of solving the *Stein's equation*, and using it to have some bounds (we refer it as *Stein's bound*) on distances between a Gaussian law and a random variable. We will begin, like in part I, by treating the so-called one dimensional case, when the random variables are real. We will (quickly) generalize it in the multi dimensional case, when we deal with random vectors. We will mix it with some notions of Malliavin calculus to prove one important and surprising theorem : the fourth moment theorem, stating that for a sequence in a set Wiener chaos, the convergence in law of this sequence to a Gaussian measure is *equivalent* with convergence of the expectation of the square and the fourth power of this sequence to the expectation of the Gaussian. This theorem has a great consequence. The convergence in law of a vector of entries belonging to set Wiener chaos to a Gaussian law is *equivalent* to the component by component one to a Gaussian law.

The second part concludes with an application of those two parts and can be seen as a synthesis of this report, the Breuer-Major theorem. This one states that a function of centered stationary Gaussian process, integrable with respect to the standard Gaussian measure, satisfies a central limit theorem if the covariance function at a certain power (depending on the function) is summable.

Part I - Malliavin calculus

Table of contents

I One dimensional case

We consider the density measure γ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$\mathrm{d}\gamma(x) \overset{\text{déf.}}{=} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \, \mathrm{d}x.$$

I.1 Derivative operator

Definition I.1

We call *smooth functions* all function $f : \mathbb{R} \longrightarrow \mathbb{R}$ which are C^∞ such that :

$$\forall p \in \mathbb{N}, \exists n \in \mathbb{N}, \exists C \geqslant 0, \forall x \in \mathbb{R}, \left| f^{(p)}(x) \right| \leqslant C|x|^n$$

We note $\mathcal{S}$ the set of all smooth functions.

We will generalize this in next section with *smooth functionals*. To justify the relevance of those smooth functions, let introduce the following lemma :

Lemma I.1

Let M be a subspace of $L^q(X, \mathcal{A}, \mu)$, where μ is a σ-finite measure, and $(X, \mathcal{A})$ is a measurable space, for $q \in [1, +\infty[$. Let $q' \in]1, +\infty]$, such that $\frac{1}{q'} + \frac{1}{q} = 1$. Then, M is dense in $L^q(\mu)$ if and only if for all $g \in L^{q'}(\mu)$:

$$\left[\forall f \in M, \int_X f(x)g(x) \, \mathrm{d}\mu(x) = 0 \right] \Longrightarrow \mu(g \neq 0) = 0$$

Proof of the lemma : • Before beginning, we suppose known the two following facts, the second one coming from Hahn-Banach theorem :

i. *Riesz duality theorem.* The map

$$J : \begin{pmatrix} L^{q'}(\mu) & \longrightarrow & (L^q(\mu))' \\ g & \longmapsto & \begin{pmatrix} L^q(\mu) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_X f(x)g(x) \, \mathrm{d}\mu(x) \end{pmatrix} \end{pmatrix}$$

is an isometry for the norm associated with L^q and $(L^q)'$.

ii. If M is linear subspace of $L^q(\mu)$ and $f_0 \in M$, then $f_0 \in \bar{M}$ if and only if there is no $T \in (L^q(\mu))'$ such that $T_{|M} = 0$ and $T(f_0) \neq 0$.

[$\Longleftarrow$] Suppose that we have this implication. Let $f_0 \in L^q(\mu)$, and make the assumption that there exists $T \in L^q(\mu)'$ such that $T_{|M} = 0$ and $Tf_0 \neq 0$. Then, by the isometry J, there exists $g \in L^{q'}(\mu)$ such that

$$\forall f \in L^q(\mu), Tf = \int_X fg \, \mathrm{d}\mu$$

By hypothesis, $T_{|M} = 0$. By the implication, it follows that $g = 0$ μ-almost everywhere, and so that T is the null form, which contradicts the fact that $Tf_0 \neq 0$. Hence, there is no such T, and by **ii.**, it follows that $f_0 \in \bar{M}$, and so that $\bar{M} = L^q(\mu)$.

[$\Longrightarrow$] Suppose that M is dense in $L^q(\mu)$. Let $g \in L^{q'}(\mu)$ such that

$$\forall f \in M, Tf \overset{\text{déf.}}{=} \int_X fg \, \mathrm{d}\mu = 0$$

Let us show that $g = 0$. Since $T \in (L^q(\mu))'$, *ii.* implies that for all $f_0 \in \bar{M}$, $Tf_0 = 0$, and since M is dense in $L^q(\mu)$, it follows that $T = 0$. By the isometry J, since $T = J(g)$, we conclude that $g = 0$ on $L^q(\mu)$, hence $g = 0$ μ-almost everywhere. □

Proposition I.1 : Density of smooth functions

The set of monomial functions $\{x \longmapsto x^n, n \in \mathbb{N}\}$ generates a dense subspace of $L^q(\gamma)$, for all $q \in [1, +\infty[$. In particular, the set $\mathcal{S}$ of smooth functions is dense in $L^q(\mathbb{R})$.

Proof : We use the lemma.

• We begin with the case where $q \in]1, +\infty[$ and its conjugate $q' \in]1, +\infty[$ cannot be infinite : $\frac{1}{q} + \frac{1}{q'} = 1$. Let $g \in L^{q'}(\gamma)$ such that

$$\forall n \in \mathbb{N}, \int_{\mathbb{R}} g(x)x^n \, \mathrm{d}\gamma(x) = 0.$$

We want to prove that $g = 0$ almost everywhere (in the sens of Lebesgue measure). We will show that $\hat{g} = 0$. We want to prove that pour all $\xi \in \mathbb{R}$:

$$\int_{\mathbb{R}} g(x)e^{\mathrm{i}\xi x} \, \mathrm{d}\gamma(x) = \sum_{n=0}^{+\infty} \frac{(\mathrm{i}\xi)^n}{n!} \int_{\mathbb{R}} g(x)x^n \, \mathrm{d}\gamma(x) = 0$$

To prove this switching between sum and integral, we have to show that, when $\xi \in \mathbb{R}$:

$$\sum_{n=0}^{+\infty} \int_{\mathbb{R}} \left| \frac{(\mathrm{i}x\xi)^n}{n!} g(x) e^{\frac{-x^2}{2}} \right| \, \mathrm{d}x < +\infty$$

To prove it, we switch by Fubini-Tonelli theorem :

$$\sum_{n=0}^{+\infty} \int_{\mathbb{R}} \left| \frac{(\mathrm{i}x\xi)^n}{n!} g(x) e^{\frac{-x^2}{2}} \right| \, \mathrm{d}x = \int_{\mathbb{R}} |g(x)| e^{|x\xi|} e^{\frac{-x^2}{2}} \, \mathrm{d}x.$$

And we use the Hölder inequality :

$$\sum_{n=0}^{+\infty} \int_{\mathbb{R}} \left| \frac{(\mathrm{i}x\xi)^n}{n!} g(x) e^{\frac{-x^2}{2}} \right| \, \mathrm{d}x \leqslant \sqrt{2\pi} \|g\|_{L^q(\gamma)} \int_{\mathbb{R}} e^{|x\xi|} e^{\frac{-x^2}{2}} \, \mathrm{d}x.$$

And as a consequence, we have shown that $\hat{g} = 0$, so $g = 0$ almost everywhere. The generated set of monomials is then dense in $L^q(\gamma)$, and since this space is included in $\mathcal{S}$, it follows that $\mathcal{S}$ is dense in $L^q(\gamma)$, for all $q > 1$.

• The case $q = 1$ is almost the same. Let $g \in L^\infty(\gamma)$. We use the Fourier transform to show that $g = 0$ almost everywhere. The only thing that changes is how we justify the switching, which is easier since :

$$\sum_{n=0}^{+\infty}\int_{\mathbb{R}}\left|\frac{(\mathrm{i}x\xi)^n}{n!}g(x)e^{\frac{-x^2}{2}}\right|\ \mathrm{d}x \leqslant \|g\|_{L^\infty}\int_{\mathbb{R}} e^{|\xi x|}e^{\frac{-x^2}{2}}\ \mathrm{d}x$$

which is finite. □

Remind the definition of a closable operator, and let us use it for the derivative operator.

Proposition I.2 : Closable operator

Let $(A, D(A))$ an operator with domain $D(A)$, that is a linear map $A : D(A) \longrightarrow F$ defined on a subspace $D(A)$ of a Banach space E into a Banach space F. Then, the following propositions are equivalents :

(i) There is an operator $(B, D(B))$ such that $D(A) \subset D(B)$, $B_{|_{D(A)}} = A$ and $D(B)$ is closed in E ;

(ii) If $\Gamma(A) \overset{\text{déf.}}{=} \{(f, Af), f \in D(A)\} \subset E \times F$, then there exists an operator $(B, D(B))$ such that $\overline{\Gamma(A)} = \Gamma(B)$, where the closure is taken for the product norm of $E \times F$;

(iii) For all $(f_n)_n \in D(A)^{\mathbb{N}}$ such that $(f_n)_n$ converges in E to 0 and $(Af_n)_n$ converges in F to g then $g = 0$.

In this case, we say that A is *closable*, and we note $B = \bar{A}$ the *closure* of A.

Definition I.2

We consider the *derivative operator* $(D, \mathcal{S})$ on $L^q(\gamma)$ defined by

$$\forall f \in \mathcal{S}, Df = f' \in \mathcal{S} \subset L^q(\gamma).$$

We note D^p the p-th iteration of D on $\mathcal{S}$.

Proposition I.3 : Extension of the derivative operator

For all $p \in \mathbb{N}^*$, the operator $(D^p, \mathcal{S})$ is closable in $L^q(\gamma)$.

In the following, we will still note D the extension of $(D, \mathcal{S})$.

Proof : We will use the sequence characterization of the proposition **I.2.**. Let first make a observation.

- First, we will observe that for $f \in \mathcal{S}$:

$$\int_{\mathbb{R}} f'(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} xf(x)\ \mathrm{d}\gamma(x).$$

Since f is C^∞, an integration by parts can conclude this equality. We wrote for $f \in \mathcal{S}$:

$$[\delta f](x) \overset{\text{déf.}}{=} f'(x) - xf(x).$$

Then δ sends $\mathcal{S}$ to itself and satisfies

$$\forall f \in \mathcal{S}, \int_{\mathbb{R}} [\delta f](x)\ \mathrm{d}\gamma(x) = 0.$$

- Let $p \in \mathbb{N}^*$ and $(f_n)_n \in \mathcal{S}^{\mathbb{N}}$ such that

$$\int_{\mathbb{R}} |f_n(x)|^q\ \mathrm{d}\gamma(x) \xrightarrow[n\to+\infty]{} 0$$

and such that there is $\eta \in L^q(\gamma)$ satisfying :

$$\int_{\mathbb{R}} \left|f_n^{(p)}(x) - \eta(x)\right|^q\ \mathrm{d}\gamma(x) \xrightarrow[n\to+\infty]{} 0$$

Let us prove that $\eta = 0$. To do that, we show that for every $g \in \mathcal{S}$:

$$\int_{\mathbb{R}} \eta(x)g(x)\ \mathrm{d}\gamma(x) = 0$$

Since $\mathcal{S}$ is dense in $L^q(\gamma)$, that would implies that $\eta = 0$ almost everywhere, by the lemme **I.1**, so $\eta = 0$ in $L^q(\gamma)$.

To show the integral equality, we proceed by integration by parts. We observe indeed that

$$\int_{\mathbb{R}} f_n^{(p)}(x)g(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} f_n^{(p-1)}(x)[\delta g](x)\ \mathrm{d}\gamma(x).$$

By induction, we conclude that

$$\int_{\mathbb{R}} f_n^{(p)}(x)g(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} f_n(x)[\delta^p g](x)\ \mathrm{d}\gamma(x).$$

Moreover, by Hölder inequality :

$$\left|\int_{\mathbb{R}} f_n(x)[\delta^p g](x)\ \mathrm{d}\gamma(x)\right| \leqslant \|f_n\|_{L^q(\gamma)}\|\delta^p g\|_{L^{q'}(\gamma)} \xrightarrow[n\to+\infty]{} 0,$$

since $\delta^p g \in \mathcal{S} \subset L^{q'}(\gamma)$. Finally, by continuity of the map

$$J : \left(\begin{array}{ccc} L^q(\gamma) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_{\mathbb{R}} f(x)g(x)\ \mathrm{d}\gamma(x) \end{array}\right),$$

we have :

$$\int_{\mathbb{R}} \eta(x)g(x)\ \mathrm{d}\gamma(x) = \lim_{n\to+\infty} f_n(x)g(x)\ \mathrm{d}\gamma(x) = 0.$$

That proves that $\eta = 0$, hence $(D^p, \mathcal{S})$ is closable in $L^q(\gamma)$. □

By this proposition, we can write $D^p f = f^{(p)}$ for every $f \in L^q(\gamma)$ and $p \in \mathbb{N}^*$, as follows : since $\mathcal{S}$ is dense into $L^q(\gamma)$, we can consider a sequence $(f_n)_n \in \mathcal{S}^{\mathbb{N}}$ converging to f in $L^q(\gamma)$. We can by consequence write

$$D^p f \overset{\text{déf.}}{=} \lim_{n\to+\infty} D^{(p)} f_n,$$

where the limit is taken in $L^q(\gamma)$. That being said, we can enter the subject by defining the space $\mathbb{D}$.

Definition I.3

Let $p \in \mathbb{N}$ and $q \in [1, +\infty[$. We denote by $\mathbb{D}^{p,q}$ the closure of $\mathcal{S}$ the following norm :

$$\forall f \in L^q(\gamma),\ \|f\|_{\mathbb{D}^{p,q}} \overset{\text{déf.}}{=} \left[\sum_{k=0}^{p} \left\|f^{(k)}\right\|_{L^q(\gamma)}^q\right]^{\frac{1}{q}} = \left[\sum_{k=0}^{p} \int_{\mathbb{R}} \left|f^{(k)}(x)\right|^q \,\mathrm{d}\gamma(x)\right]^{\frac{1}{q}}$$

where $f^{(k)}$ is taken in the sens of $D^k f$, defined previously. We also note

$$\mathbb{D}^{\infty,q} \overset{\text{déf.}}{=} \bigcap_{p=1}^{+\infty} \mathbb{D}^{p,q}$$

We check that the last definition is correct by observing that $\mathbb{D}^{p+1,q} \subset \mathbb{D}^{p,q}$.
An element $f \in L^q(\gamma)$ belongs to $\mathbb{D}^{p,q}$ if and only if there exists a sequence $(f_n)_n$ of elements of $\mathcal{S}$ such that

$$\forall k \in [\![0,p]\!],\ \int_{\mathbb{R}} \left|f_n^{(k)}(x) - f^{(k)}(x)\right|^q \,\mathrm{d}\gamma(x) \xrightarrow[n\to+\infty]{} 0,$$

if and only if, since D is closed :

$$\int_{\mathbb{R}} |f_n(x) - f(x)|^q \,\mathrm{d}\gamma(x) \xrightarrow[n\to+\infty]{} 0 \text{ and } \forall k \in [\![1,p]\!],\ \left(f_n^k\right)_n \text{ is a Cauchy sequence.}$$

Definition I.4

Let $p \in \mathbb{N}$ and $q \in [1, +\infty[$. Then, the operator $(D^p, \mathbb{D}^{p,q})$ on $\left(L^q(\gamma), \|\cdot\|_{L^q(\gamma)}\right)$ is called the *p-th derivation operator*.

A consequence of the closability of D^p is the following.

Proposition I.4 : Domain of the derivative operator

The set $\mathbb{D}^{p,q}$ is the domain of the closed operator D^p on $L^q(\gamma)$.

I.2 Case of square integrability

Proposition I.5 : Divergence

Let

$$D \overset{\text{déf.}}{=} \left\{g \in L^2(\gamma), \exists C > 0, \forall f \in \mathcal{S}, \left|\int_{\mathbb{R}} f^{(p)}(x) g(x) \,\mathrm{d}\gamma(x)\right| \leqslant C\|f\|_{L^2(\gamma)}\right\}.$$

For all $g \in D$, there exists an unique element $\delta^p g \in L^2(\gamma)$ such that

$$\forall f \in \mathcal{S},\ \int_{\mathbb{R}} f^{(p)}(x) g(x) \,\mathrm{d}\gamma(x) = \int_{\mathbb{R}} f(x)[\delta^p g](x) \,\mathrm{d}\gamma(x)$$

We call the operator (δ, D) the *p-th divergence operator*.

Proof : Let $g \in D$. Then, by hypothesis,

$$\begin{pmatrix} \mathcal{S} & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_{\mathbb{R}} f^{(p)}(x) g(x) \,\mathrm{d}\gamma(x) \end{pmatrix}$$

is continuous for $\|\cdot\|_{L^2(\gamma)}$. By consequence, we can extend it to $L^2(\gamma)$, and apply Riesz representation theorem. □

The map

$$\left(\begin{array}{ccc} (\mathcal{S}, \|\cdot\|_{\mathbb{D}^{p,2}}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_{\mathbb{R}} g(x) f^{(p)}(x) \ \mathrm{d}\gamma(x) \end{array}\right)$$

is continuous, hence the equality defining $D(\delta^p)$ the same as the equality with $f \in \mathbb{D}^{p,2}$. The set $\mathcal{S}$ is included in $D(\delta^p)$. By the same continuity, it follow that $\mathbb{D}^{p,2} \subset D(\delta^p)$.

Definition I.5

For all $f \in \mathcal{S}$, we note P_t the *Ornstein-Uhlenbeck operator* defined by :

$$\forall t \geqslant 0, \forall x \in \mathbb{R}, [P_t f](x) \overset{\text{def.}}{=} \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) \ \mathrm{d}\gamma(y).$$

An interesting interpretation of P_t is the following : if $N \sim \mathcal{N}(0,1)$, then

$$[P_t f](x) = \mathbb{E}\left[f\left(e^{-t}x + \sqrt{1-e^{-2t}}N\right)\right].$$

Proposition I.6 : Ornstein-Ulenbeck properties

(i) For all $f \in L^1(\gamma)$ and for all $s, t \geqslant 0$, $P_{t+s}f = P_t P_s f$;

(ii) For all $f \in L^q(\gamma)$, for all $t \geqslant 0$,

$$\|P_t f\|_{L^q(\gamma)} \leqslant \|f\|_{L^q(\gamma)}.$$

In other words, P_t is a contraction of $L^q(\gamma)$.

(iii) For all $f \in \mathbb{D}^{1,2}$, $t \geqslant 0$, $P_t \in \mathbb{D}^{1,2}$ with

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

Proof : **(i)** Let $s, t \geqslant 0$. Then

$$\begin{aligned}[P_t P_s f](x) = & \int_{\mathbb{R}}\int_{\mathbb{R}} f\Big(e^{-(s+t)}x \\ & + ye^{-s}\sqrt{1-e^{-2t}} + z\sqrt{1-e^{-2s}}\Big) \\ & \mathrm{d}\gamma(z) \ \mathrm{d}\gamma(y).\end{aligned}$$

Hence, if $Y, Z \sim \mathcal{N}(0,1)$ are independents, then

$$\begin{aligned}[P_t P_s f](x) = & \mathbb{E}\Big[f\Big(e^{-(s+t)}x \\ & + Ye^{-s}\sqrt{1-e^{-2t}} + Z\sqrt{1-e^{-2s}}\Big)\Big].\end{aligned}$$

We can compute the characteristic function, by independence :

$$\begin{aligned} & \mathbb{E}\left[\exp\left(\mathrm{i}u\left[Ye^{-s}\sqrt{1-e^{-2t}} + Z\sqrt{1-e^{-2s}}\right]\right)\right] \\ = & \mathbb{E}\left[\exp\left(\mathrm{i}uYe^{-s}\sqrt{1-e^{-2t}}\right)\right]. \\ & \mathbb{E}\left[\exp\left(\mathrm{i}uZ\sqrt{1-e^{-2s}}\right)\right].\end{aligned}$$

And since $Y, Z \sim \mathcal{N}(0,1)$:

$$\begin{aligned} & \mathbb{E}\left[\exp\left(\mathrm{i}u\left[Ye^{-s}\sqrt{1-e^{-2t}} + Z\sqrt{1-e^{-2s}}\right]\right)\right] \\ = & \exp\left(\frac{-u^2 e^{-2s}\left(1-e^{-2t}\right)}{2}\right). \\ & \exp\left(\frac{-u^2(1-e^{-2s})}{2}\right).\end{aligned}$$

Then,

$$\begin{aligned} & \mathbb{E}\left[\exp\left(\mathrm{i}u\left[Ye^{-s}\sqrt{1-e^{-2t}} + Z\sqrt{1-e^{-2s}}\right]\right)\right] \\ = & \exp\left(\frac{-u^2\left(1-e^{-2(t+s)}\right)}{2}\right).\end{aligned}$$

We can conclude that :

$$\frac{1}{\sqrt{1-e^{-2(t+s)}}}\left[Ye^{-s}\sqrt{1-e^{-2t}} + Z\sqrt{1-e^{-2s}}\right] \sim \mathcal{N}(0,1).$$

If $N \sim \mathcal{N}(0,1)$, this means for $P_t P_s$ that :

$$[P_t P_s f](x) = \mathbb{E}\left[f\left(e^{-(s+t)}x + \sqrt{1-e^{-2(t+s)}}N\right)\right].$$

Finally, we obtain our semi group relation :

$$[P_t P_s f](x) = P_{t+s} f(x).$$

(ii) Let $f \in \mathcal{S}$. By Jensen inequality :

$$\begin{aligned} & \int_{\mathbb{R}} \left|\mathbb{E}\left[f\left(e^{-t}x + \sqrt{1-e^{-2t}}N\right)\right]\right|^q \ \mathrm{d}\gamma(x) \\ \leqslant & \int_{\mathbb{R}} \mathbb{E}\left[\left|f\left(e^{-t}x + \sqrt{1-e^{-2t}}N\right)\right|^q\right] \ \mathrm{d}\gamma(x)\end{aligned}$$

Let $X \sim \mathcal{N}(0,1)$ independent of N. Then :

$$\begin{aligned} & \int_{\mathbb{R}} \left|\mathbb{E}\left[f\left(e^{-t}x + \sqrt{1-e^{-2t}}N\right)\right]\right|^q \ \mathrm{d}\gamma(x) \\ \leqslant & \mathbb{E}\left[\left|f\left(e^{-t}X + \sqrt{1-e^{-2t}}N\right)\right|^q\right]\end{aligned}$$

We can once again compute the characteristic function of what is inside the expectation, by using independence, the expression of the characteristic function of a Gaussian random variable, and the transfer theorem.

$$\begin{aligned}&\mathbb{E}\left[\exp\left(\mathrm{i}u\left(e^{-t}X+\sqrt{1-e^{-2t}}N\right)\right)\right]\\ =\ &\exp\left(\frac{-u^2e^{-2t}}{2}\right)\exp\left(\frac{-u^2\left(1-e^{-2t}\right)}{2}\right)\end{aligned}$$

So,

$$\mathbb{E}\left[\exp\left(\mathrm{i}u\left(e^{-t}X+\sqrt{1-e^{-2t}}N\right)\right)\right]=e^{\frac{-u^2}{2}}$$

and $e^{-t}X+\sqrt{1-e^{-2t}}N\sim\mathcal{N}(0,1)$. Finally :

$$\begin{aligned}&\int_{\mathbb{R}}\left|\mathbb{E}\left[f\left(e^{-t}x+\sqrt{1-e^{-2t}}N\right)\right]\right|^q\ \mathrm{d}\gamma(x)\\ \leqslant\ &\mathbb{E}\left[|f(Z)|^q\right]\end{aligned}$$

where $Z\sim\mathcal{N}(0,1)$. That is :

$$\|P_tf\|_{L^q(\gamma)}\leqslant\|f\|_{L^q(\gamma)}.$$

Since both size are well-defined for $f\in L^q(\gamma)$, and by the continuity we just checked on the dense space $\mathcal{S}$, we conclude that the inequality remains true on $L^q(\gamma)$.

(iii) By contraction property of P_t, if $f\in\mathbb{D}^{1,2}$, then $P_tf\in\mathbb{D}^{1,2}$. Let us show the relation between D and P_t. By continuity of P_t, it is sufficient to check this for $f\in\mathcal{S}$, where we can freely switch integrals and derivatives. We have :

$$[DP_t\ f](x)=e^{-t}\int_{\mathbb{R}}f'\left(e^{-t}x+\sqrt{1-e^{-2t}}y\right)\ \mathrm{d}\gamma(y).$$

So,

$$[DP_t\ f](x)=e^{-t}[P_tf'](x)=e^{-t}[P_tDf](x).$$

This equality stays true by continuity for all $f\in\mathbb{D}^{1,2}$.

□

We can see $(P_t)_t$ as a semi-group associated with a stochastic process.

Proposition I.7 : Ornstein-Uhlenbeck stochastic process

Let $(X_t^x)_t$ the stochastic process defined by the solution of the following stochastic differential equation starting from $X_0=x$:

$$\mathrm{d}X_t=\sqrt{2}\ \mathrm{d}B_t-X_t\ \mathrm{d}t$$

Hence, for all $f\in\mathcal{S}$:

$$\forall x\in\mathbb{R},[P_tf](x)=\mathbb{E}\left[f\left(X_t^x\right)\right].$$

Proof : • Let us write the explicit solution of this SDE. If we remove the Brownian part, then the solution would be $X_t=e^{-t}X_0$. We consider as a consequence

$$f(t,x)\overset{\text{def.}}{=}e^tx$$

and we apply the Itô formula to $Y_t\overset{\text{def.}}{=}f(t,X_t)$.

$$\mathrm{d}Y_t=e^tX_t\ \mathrm{d}t+e^t\ \mathrm{d}X_t.$$

We substitute by using the SDE that X_t solves :

$$\mathrm{d}Y_t=e^tX_t\ \mathrm{d}t+\sqrt{2}e^t\ \mathrm{d}B_t-e^tX_t\ \mathrm{d}t.$$

That is

$$\mathrm{d}Y_t=\sqrt{2}e^t\mathrm{d}B_t,$$

meaning that :

$$Y_t=Y_0+\sqrt{2}\int_0^te^s\ \mathrm{d}B_s$$

And then, we finally have, since $Y_0=X_0$:

$$X_t=e^{-t}X_0+\sqrt{2}\int_0^te^{-(t-s)}\ \mathrm{d}B_s$$

• Let $f\in\mathcal{S}$. Then, for all $x\in\mathbb{R}$,

$$\mathbb{E}\left[f\left(X_t^x\right)\right]=\mathbb{E}\left[f\left(e^tx+\sqrt{2}\int_0^te^{-(t-s)}\ \mathrm{d}B_s\right)\right].$$

Let us show that

$$A_t\overset{\text{def.}}{=}\sqrt{t}\frac{\sqrt{2}}{\sqrt{1-e^{-2t}}}\int_0^te^{-(t-s)}\ \mathrm{d}B_s$$

is a Brownian motion. To do so, we use the Lévy characterisation theorem. In other words, we compute $\langle A,A\rangle_t$:

$$\langle A,A\rangle_t=\frac{2te^{-2t}}{1-e^{-2t}}\int_0^te^{2s}\ \mathrm{d}s.$$

It yields to :

$$\langle A,A\rangle_t=t.$$

Since the process A is continuous, we conclude that A is a Brownian motion. So $A_t\sim\mathcal{N}(0,t)$, and so

$$\frac{\sqrt{2}}{\sqrt{1-e^{-2t}}}\int_0^te^{-(t-s)}\ \mathrm{d}B_s\sim\mathcal{N}(0,1).$$

If $N\sim\mathcal{N}(0,1)$ we get :

$$\mathbb{E}\left[f\left(X_t^x\right)\right]=\mathbb{E}\left[f\left(e^tx+\sqrt{1-e^{-2t}}N\right)\right].$$

So $\mathbb{E}\left[f\left(X_t^x\right)\right]=[P_tf](x)$. □

Since $(P_t)_t$, we can associate it an infinitesimal generator L.

Proposition I.8 : Infinitesimal generator

Let $(L, D(L))$ the infinitesimal generator of P_t on $L^2(\gamma)$, that is the unbounded operator defined on

$$D(L) \overset{\text{def.}}{=} \left\{ f \in L^2(\gamma) \middle| \exists g \in L^2(\gamma), \int_{\mathbb{R}} \left| \frac{P_t f - f}{t} - g \right|^2 \, \mathrm{d}\gamma \xrightarrow[t \to 0^+]{} 0 \right\},$$

by

$$Lf = \lim_{t \to 0^+} \frac{P_t f - f}{t},$$

where the limit takes place in $L^2(\gamma)$. Then $\mathcal{S} \subset D(L)$. Moreover, for all $f \in \mathcal{S}$,

$$Lf = -\delta D f$$

Proof : • Let $f \in \mathcal{S}$. Let us show that $f \in D(L)$, that is $\frac{P_t f - f}{t}$ admits a limit when $[t \to 0]$ in $L^2(\gamma)$. Consider first the simple convergence, taking a look at the derivative with respect to t of $P_t f$. Since $f \in \mathcal{S}$, we can switch derivative and integral :

$$\begin{aligned} & \frac{\mathrm{d}}{\mathrm{dt}} \left[P_t f \right](x) \\ = \quad & -e^{-t} x [P_t D f](x) \\ + \quad & \frac{e^{-2t}}{\sqrt{1-e^{-2t}}} \int_{\mathbb{R}} f' \left(e^{-t} x + \sqrt{1-e^{-2t}} y \right) y \, \mathrm{d}\gamma(y). \end{aligned}$$

And we integrate by parts, knowing that "$y \, \mathrm{d}\gamma(y)$" becomes "$\mathrm{d}\gamma(y)$" with a derivative on f :

$$\begin{aligned} & \frac{\mathrm{d}}{\mathrm{dt}} \left[P_t f \right](x) \\ = \quad & -e^{-t} x [P_t D f](x) \\ + \quad & e^{-2t} \int_{\mathbb{R}} f'' \left(e^{-t} x + \sqrt{1-e^{-2t}} y \right) \, \mathrm{d}\gamma(y). \end{aligned}$$

Finally :

$$\frac{\mathrm{d}}{\mathrm{dt}} \left[P_t f \right](x) = -e^{-t} x \left[P_t D f \right](x) + e^{-2t} \left[P_t D^2 f \right](x).$$

But, we also have for all $g \in D(L)$:

$$Lg = \frac{\mathrm{d}}{\mathrm{dt}} [P_t g] \Big|_{t=0}$$

Hence, we would have :

$$[Lf](x) = -x f'(x) + f''(x)$$

So, to prove that $f \in D(L)$, we just need to conclude by dominated convergence. Indeed, since $f \in \mathcal{S}$, we consider k such that f and all its derivative are bounded by $O(|x|^k)$ at $+\infty$. Then, by mean value theorem, we have the existence of a constant C such that

$$\left| \frac{P_t - f}{t} \right| \leqslant \frac{1 - e^{-t}}{t} |x| + C \frac{\sqrt{1-e^{-2t}}}{t}.$$

Then, we have :

$$\left| \frac{P_t - f}{t} \right| \leqslant |x| + C\sqrt{2}.$$

Hence, we proved the convergence in $L^2(\gamma)$. This proved that $\mathcal{S} \subset D(L)$ and the expected equality, since $Lf = -\delta f' = -\delta D f$. □

Proposition I.9 : Heisenberg relation of D ans δ

We have $[D, \delta] = \mathrm{id}_{\mathcal{S}}$. By induction :

$$\forall p \in \mathbb{N}^*, \forall f \in \mathcal{S}, [D, \delta^p] f = p \delta^{p-1} f$$

Proof : It is a straightforward computation. If $f \in \mathcal{S}$,

$$[\delta, D] f = \delta f' - D \delta f,$$

with $\delta f(x) = x f(x) - f'(x)$. Hence,

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

And so

$$[\delta, D] f(x) = x f'(x) - f''(x) - x f'(x) - f(x) + f''(x) = -f(x).$$

□

The following corollary will show us a new way to use P_t thanks to those relations.

Corollary I.1 : A relation with P_t

Let $f \in \mathcal{S}$. Then, for all $x \in \mathbb{R}$,

$$\int_0^{+\infty} e^{-2t} \left[P_t f'' \right](x) \, \mathrm{dt} - x \int_0^{+\infty} e^{-t} \left[P_t f' \right](x) \, \mathrm{dt} = f(x) - \int_{\mathbb{R}} f(y) \, \mathrm{d}\gamma(y).$$

Proof : Recall that if $f \in \mathcal{S}$:

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

We use the relation $DP_t = e^{-t}P_tD$ to the left hand side of the wanted equality, we call it A :

$$A \overset{\text{def.}}{=} \int_0^{+\infty} e^{-2t}\left[P_tf''\right](x)\,\mathrm{d}t - x\int_0^{+\infty} e^{-t}\left[P_tf'\right](x)\,\mathrm{d}t.$$

Then

$$A = \int_0^{+\infty}\left(\left[D^2P_tf\right](x) - x\left[DP_tf\right](x)\right)\,\mathrm{d}t.$$

We recover δ :

$$A = \int_0^{+\infty} [\delta DP_t]f(x)\,\mathrm{d}t.$$

Since $L = -\delta D$, we get :

$$A = -\int_0^{+\infty} [LP_t]f(x)\,\mathrm{d}t.$$

And by property of infinitesimal generator, $LP_t = \frac{\mathrm{d}P_t}{\mathrm{d}t}$, so good as :

$$A = -\int_0^{+\infty} \frac{\mathrm{d}P_tf}{\mathrm{d}t}(x)\,\mathrm{d}t.$$

By dominated convergence,

$$P_\infty f(x) = \int_{\mathbb{R}} f(y)\,\mathrm{d}\gamma(y).$$

We finally get what we want :

$$A = P_0f(x) - P_\infty f(x) = f(x) - \int_{\mathbb{R}} f(y)\,\mathrm{d}\gamma(y).$$

□

I.3 Hermite polynomials

Definition I.6

Let $p \in \mathbb{N}$. We define the p-th *Hermite polynomial* as $H_0 = 1$ if $p = 0$ and for $p \geqslant 1$:

$$H_p \overset{\text{def.}}{=} \delta^p 1,$$

where $1 = \mathbf{1}_{\mathbb{R}}$ is the constant function equal to 1. We make the convention that $H_{-1} = 0$ is the null polynomial.

Consequently, we define H_p as the unique function of $L^2(\gamma)$ such that

$$\forall g \in \mathbb{D}^{p,2}, \int_{\mathbb{R}} H_p(x)g(x)\,\mathrm{d}\gamma(x) = \int_{\mathbb{R}} g^{(p)}(x)\,\mathrm{d}\gamma(x).$$

Here are the fundamental properties of the Hermite polynomials, beginning by the fact that H_p is indeed a polynomial function by using a induction relation.

Proposition I.10 : Recurrent relation for H_p

For any $p \geqslant 0$ and $t \geqslant 0$:

(i) $H_p' = pH_{p-1}$;

(ii) $H_{p+1}(x) = xH_p(x) - pH_{p-1}(x)$;

(iii) $LH_p = -pH_p$;

(iv) $P_tH_p = e^{-pt}H_p$.

Proof : **(i)** We use the Heisenberg relation for D and δ :

$$DH_p = D\delta^p 1 = \delta^p D1 + p\delta^{p-1}1.$$

So,

$$H_p' = pH_{p-1}.$$

We conclude that H_p is indeed a polynomial by this relation.

(ii) Let us see that $H_{p+1} = \delta H_p$, meaning that, with the definition of δ :

$$H_{p+1}(x) = xH_p(x) - H_p'(x) = xH_p(x) - pH_{p-1}(x)$$

by using **(i)**.

(iii) We have $L = -\delta D$ so :

$$LH_p = \delta DH_p = p\delta H_{p-1} = pH_p.$$

(iv) Let us define $y(t) = P_tH_p(x)$. We have $y(0) = H_p(x)$ and

$$y'(t) = P_tLH_p(x) = -pP_tH_p(x) = -py(t)$$

so $y(t) = e^{-pt}y(0)$, meaning that $P_tH_p = e^{-pt}H_p$. □

We can claim that H_p as the same parity as p, we can prove it by induction and using **(ii)**.
Here's an important property that justifies the use of those polynomials.

Proposition I.11 : Hilbert basis of $L^2(\gamma)$

The family $\left\{\frac{1}{\sqrt{p!}}H_p, p \in \mathbb{N}\right\}$ generates an orthonormal basis of $L^2(\gamma)$, in the sens of Hilbert space.

Proof : **1.** We begin by checking that $(H_p)_p$ is an orthogonal family. This a consequence from the following recurrent relation : for $p \in \mathbb{N}$ and $r \in \mathbb{N}$:

$$\int_{\mathbb{R}} H_p(x)H_{p+r}(x)\ \mathrm{d}\gamma(x) = (p+r)\int_{\mathbb{R}} H_{p-1}(x)H_{(p-1)+r}(x)\ \mathrm{d}\gamma(x).$$

This relation is true by using the fact that $H_p = \delta^p 1$:

$$\int_{\mathbb{R}} H_p(x)H_{p+r}(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} [\delta H_{p-1}](x)H_{p+r}(x)\ \mathrm{d}\gamma(x).$$

By definition of δ :

$$\int_{\mathbb{R}} H_p(x)H_{p+r}(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} H_{p-1}(x)H'_{p+r}(x)\ \mathrm{d}\gamma(x).$$

And by the expression of the derivative of H_p given by **(i)**, we finally have :

$$\int_{\mathbb{R}} H_p(x)H_{p+r}(x)\ \mathrm{d}\gamma(x) = (p+r)\int_{\mathbb{R}} H_{p-1}(x)H_{(p-1)+r}(x)\ \mathrm{d}\gamma(x).$$

2. Let us prove by induction on $p \geqslant 0$ that

$$\forall r \geqslant 1, \int_{\mathbb{R}} H_p(x)H_{p+r}(x)\ \mathrm{d}\gamma(x) = 0$$

For $p = 0$, we check that

$$\int_{\mathbb{R}} H_r(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} \delta^r 1(x)\ \mathrm{d}\gamma(x) = 0$$

Then, if we suppose the property true for $p \geqslant 0$, we use the relation to prove it at the rank $p+1$. That concludes the induction, we have shown that $(H_p)_p$ is a orthonormal family.

3. Let us compute $\int_{\mathbb{R}} H_p(x)^2\ \mathrm{d}\gamma(x)$. To do that, we use the same recurrent relation. For $p = 0$, we have

$$\int_{\mathbb{R}} H_0(x)^2\ \mathrm{d}\gamma(x) = 1.$$

Then, for $p \geqslant 1$,

$$\int_{\mathbb{R}} H_p(x)^2\ \mathrm{d}\gamma(x) = p\int_{\mathbb{R}} H_{p-1}(x)^2\ \mathrm{d}\gamma(x).$$

Finally, we obtain that

$$\|H_p\|_{L^2(\gamma)} = \sqrt{p!}$$

And so $\left(\frac{H_p}{\sqrt{p!}}\right)_p$ is an orthonormal family of $L^2(\gamma)$.

4. It remains to show that the family is total. By **(ii)**, we can show that H_p is polynomial whose degree is p. Hence, the subspace spanned by the H_p is the same the one spanned by the monomials. The proposition **I.1** shows that this subspace is dense in $L^2(\gamma)$. Consequently, the family $(H_p)_p$ generates a dense subspace of $L^2(\gamma)$. □

We can decompose any function L^2 in this basis. The following proposition is more specific about $\mathbb{D}^{\infty,2}$ functions.

Proposition I.12 : Decomposition of elements of $\mathbb{D}^{\infty,2}$

Let $f \in \mathbb{D}^{\infty,2}$. Then,

$$f = \sum_{p=0}^{+\infty}\left(\frac{1}{p!}\int_{\mathbb{R}} f^{(p)}(x)\ \mathrm{d}\gamma(x)\right)H_p$$

where the convergence of the series takes place in $L^2(\gamma)$.

Proof : Since the family $\left(\frac{1}{\sqrt{p!}}H_p\right)_p$ is a orthonormal basis, we obtain the following decomposition for all $f \in L^2(\gamma)$:

$$f = \sum_{p=0}^{+\infty}\left(\frac{1}{p!}\int_{\mathbb{R}} f(x)H_p(x)\ \mathrm{d}\gamma(x)\right)H_p,$$

where the convergence is in $L^2(\gamma)$. Since $f \in \mathbb{D}^{\infty,2}$, then for all $p \in \mathbb{N}$:

$$\int_{\mathbb{R}} f(x)H_p(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} f^p(x)\ \mathrm{d}\gamma(x).$$

We obtain the wanted decomposition. □

An another way to express it uses the expectation of each derivative of f. More precisely, if $N \sim \mathcal{N}(0,1)$ then in $L^2(\gamma)$:

$$f = \sum_{p=0}^{+\infty}\frac{\mathbb{E}\left[f^{(p)}(N)\right]}{p!}H_p.$$

Corollary I.2 : Exponential formula and Rodrigues formula

1. If $c > 0$ then we have in $L^2(\gamma)$:

$$e^{cx-\frac{c^2}{2}} = \sum_{p=0}^{+\infty} \frac{c^p}{p!} H_p(x).$$

2. *Rodrigues formula.* We have the expression for H_p :

$$\forall p \geqslant 1, H_p(x) = (-1)^p e^{\frac{x^2}{2}} \frac{\mathrm{d}^p}{\mathrm{d}x^p} \left[e^{\frac{-x^2}{2}} \right].$$

Proof : **1.** We use the decomposition of the previous proposition for $f(x) = e^{cx}$

$$e^{cx} = \sum_{p=0}^{+\infty} \frac{c^p}{p!} \left(\int_{\mathbb{R}} e^{cx} \, \mathrm{d}\gamma(x) \right) H_p.$$

We just have to compute this integral to conclude. To do this, we just use the canonical form of the trinomial :

$$\frac{x^2}{2} - cx = \frac{(x-c)^2}{2} - \frac{c^2}{2}.$$

We get :

$$\int_{\mathbb{R}} e^{cx} \, \mathrm{d}\gamma(x) = e^{\frac{c^2}{2}}.$$

And finally :

$$e^{cx-\frac{c^2}{2}} = \sum_{p=0}^{+\infty} \frac{c^p}{p!} H_p.$$

2. We use the previous reduction of the trinomial

$$e^{cx-\frac{c^2}{2}} = e^{\frac{x^2}{2}} e^{\frac{-(x-c)^2}{2}}.$$

We expend in power series the second exponential term around c.

$$e^{cx-\frac{c^2}{2}} = \sum_{p=0}^{+\infty} e^{\frac{x^2}{2}} \frac{(-c)^p}{p!} \frac{\mathrm{d}^p}{\mathrm{d}x^p} \left[e^{\frac{-x^2}{2}} \right].$$

By unicity of the expend in power series around c (equality in L^2 implies equality almost everywhere), we conclude that almost everywhere

$$H_p(x) = (-1)^p e^{\frac{x^2}{2}} \frac{\mathrm{d}^p}{\mathrm{d}x^p} \left[e^{\frac{-x^2}{2}} \right],$$

equality remaining true by continuity of both sides. □

We can deduce an another relation between Hermite polynomials thanks to those relations.

Corollary I.3 : A new relation between Hermite polynomials

For all $p, j \in \mathbb{N}$:

$$H_p H_{p+j} = \sum_{r=0}^{p} r! \binom{p}{r} \binom{p+j}{r} H_{2(p-r)+j}$$

Proof : This proof is quite technical and only comes from the expansion :

$$H_p H_{p+j} = \sum_{s=0}^{+\infty} \left[\int_{\mathbb{R}} (H_p H_{p+j})^{(s)} \, \mathrm{d}\gamma \right] \frac{H_s}{s!}.$$

We compute the integral thanks to the Leibniz formula :

$$(H_p H_{p+j})^{(s)} = \sum_{k=0}^{s} \binom{s}{k} H_p^{(k)} H_{p+j}^{(s-k)}.$$

By iterating the equality $H_p' = pH_{p-1}$, we have for all $p, k \in \mathbb{N}$:

$$H_p^{(k)} = \frac{p!}{(p-k)!} H_{p-k} \mathbf{1}_{\{k \leqslant p\}}.$$

Hence, we have the following expression of the s-th derivative of $H_p H_{p+j}$:

$$(H_p H_{p+j})^{(s)} = \sum_{k=0}^{s} \binom{s}{k} \frac{p!}{(p-k)!} \frac{(p+j)!}{(p+j-s+k)!} \cdot H_{p-k} H_{p+j-s+k} \mathbf{1}_{\{k \leqslant p\} \cap \{s-k \leqslant p+j\}}.$$

By orthogonality of the family $(H_p)_p$, we have

$$\int_{\mathbb{R}} H_{p-k} H_{p+j-s+k} \, \mathrm{d}\gamma = 0,$$

except if $2k = s - j$. In this case, we have

$$\int_{\mathbb{R}} H_{p-k}^2 \, \mathrm{d}\gamma = (p-k)!.$$

Hence, we have the following heavy expression of the expansion of $H_p H_{p+j}$:

$$H_p H_{p+j} = \sum_{\substack{s=0 \\ s-j \in 2\mathbb{N}}}^{+\infty} \binom{s}{\frac{s-j}{2}} \frac{p!}{\left(p - \frac{j-s}{2}\right)!} \frac{(p+j)!}{\left(p - \frac{j-s}{2}\right)!} \left(p - \frac{j-s}{2}\right)! \cdot \mathbf{1}_{\left\{s-(p+j) \leqslant \frac{s-j}{2} \leqslant p\right\}} \frac{H_s}{s!}.$$

One of the $\left(p - \frac{j-s}{2}\right)!$ simplifies. We do the change $2r = s - j$ to clear the expression.

$$H_p H_{p+j} = \sum_{r=0}^{+\infty} \binom{2r+j}{r} \frac{p!(p+j)!}{(j+2r)!(p-r)!} \cdot \mathbf{1}_{\{2r-p \leqslant r \leqslant p\}} H_{j+2r}.$$

But, $2r - p \leqslant r$ if and only if $r \leqslant p$. So, we have here the expression :

$$H_p H_{p+j} = \sum_{r=0}^{p} \binom{2r+j}{r} \frac{p!(p+j)!}{(j+2r)!(p-r)!} H_{j+2r}.$$

And we are done : we have

$$\binom{2r+j}{r} \frac{p!(p+j)!}{(j+2r)!(p-r)!} = \frac{(2r+j)!p!(p+j)!}{r!(2r+j)!(r+j)!(p-r)!}.$$

By arranging :

$$\binom{2r+j}{r} \frac{p!(p+j)!}{(j+2r)!(p-r)!} = \binom{p}{r} \frac{(p+j)!}{(r+j)!}.$$

So, we have

$$H_p H_{p+j} = \sum_{r=0}^{p} (p-r)! \binom{p}{r} \binom{p+j}{p-r} H_{j+2r}.$$

We conclude by setting the variable change $r \leftarrow p - r$. □

We can explicit the expansion of H_p.

Corollary I.4 : Expansion of the Hermite polynomials

For all $p \in \mathbb{N}$ and $x \in \mathbb{R}$:

$$H_p(x) = \sum_{k=0}^{\lfloor \frac{p}{2} \rfloor} \frac{(-1)^k p!}{k!(p-2k)!2^k} x^{p-2k}$$

Proof : We use the relation $H_p' = pH_{p-1}$ and we proceed by induction. But, first, we need to compute $H_p(0)$. To do this, we use the exponential formula, applied in $x = 0$:

$$e^{\frac{-c^2}{2}} = \sum_{p=0}^{+\infty} \frac{c^p}{p!} H_p(0).$$

We expend in power series this function :

$$e^{\frac{-c^2}{2}} = \sum_{r=0}^{+\infty} \frac{(-1)^r c^{2r}}{2^r r!}.$$

By unicity of this expansion, we finally have that $H_p(0) = 0$ if p is odd, and if $p = 2r$ is even, then

$$H_p(0) = \frac{(-1)^{\frac{p}{2}}}{2^{\frac{p}{2}} (\frac{p}{2})!}.$$

Hence, we got this recurrent expression for H_p :

$$H_p(x) = p \int_0^x H_{p-1}(y) \, \mathrm{d}y + H_p(0).$$

We conclude by induction. □

I.4 Applications

I.4.1 Variance expansions

Proposition I.13 : Variance expansion

Let $N \sim \mathcal{N}(0,1)$.

1. Let $f \in \mathbb{D}^{\infty,2}$. We have the following expansion for the variance of $f(N)$:

$$\mathrm{Var}(f(N)) = \sum_{n=1}^{+\infty} \frac{1}{n!} \mathbb{E}\left[f^{(n)}(N)\right]^2.$$

2. Let $f \in \mathcal{S}$. If

$$\frac{\mathbb{E}\left[f^{(n)}(N)^2\right]}{n!} \xrightarrow[n \to +\infty]{} 0,$$

then we have the following expansion for the variance of $f(N)$:

$$\mathrm{Var}(f(N)) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n!} \mathbb{E}\left[f^{(n)}(N)^2\right].$$

Proof : 1. We have the following decomposition for f in $L^2(\gamma)$:

$$f = \sum_{p=0}^{+\infty} \frac{\mathbb{E}\left[f^{(p)}(N)\right]}{p!} H_p$$

By consequence, their $L^2(\gamma)$ norm are equals, which gives :

$$\int_{\mathbb{R}} f(x)^2 \, \mathrm{d}\gamma(x) = \int_{\mathbb{R}} \left(\sum_{p=0}^{+\infty} \frac{\mathbb{E}\left[f^{(p)}(N)\right]}{p!} H_p(x) \right)^2 \mathrm{d}\gamma(x).$$

Since $\left(\frac{H_p}{\sqrt{p!}}\right)_p$ is a orthonormal basis of $L^2(\gamma)$, we can use the Pythagoras theorem :

$$\int_{\mathbb{R}} f(x)^2 \, \mathrm{d}\gamma(x) = \int_{\mathbb{R}} \sum_{p=0}^{+\infty} \left(\frac{\mathbb{E}\left[f^{(p)}(N)\right]}{\sqrt{p!}} \right)^2 \, \mathrm{d}\gamma(x).$$

The content inside the integral does not longer depends on x. We have consequently obtain the wished expansion :

$$\mathbb{E}\left[f(N)^2\right] = \mathbb{E}[f(N)]^2 + \sum_{p=1}^{+\infty} \frac{\mathbb{E}\left[f^{(p)}(N)\right]^2}{\sqrt{p!}}.$$

2. We have chosen $f \in \mathcal{S}$. The idea is to represent the variance as an increment of a function g depending of P_t, using the facts that $P_0 f = f$ and $P_\infty f = \mathbb{E}[f(N)]$. Here's the plan : we introduce for all $t \in]0,1[$:

$$g(t) \overset{\text{def.}}{=} \mathbb{E}\left[\left(\left[P_{\frac{-\ln(t)}{2}} f\right](N)\right)^2\right]$$

a. We calculate first all the derivatives of g :

$$g^{(p)}(t) = \mathbb{E}\left[\left(\left[P_{\frac{-\ln(t)}{2}} f^{(p)}\right](N)\right)^2\right].$$

b. Then, we use the Taylor formula with integral reminder to show that we have the expansion around 1 :

$$g(1) - g(0) = -\sum_{p=1}^{+\infty} \frac{(-1)^p g^{(p)}(1)}{p!},$$

where we have extended all the derivatives of g to $[0,1]$. This extension gives that $\mathrm{Var}(f(N)) = g(1) - g(0)$ and allows us to conclude.

Here we go. A little word on g : the logarithm will be useful for the relation $DP_t = e^{-t} P_t D$, the "$\frac{1}{2}$" to balance the square inside the expectation, and the "$-$" to consider g on a finite interval $]0,1[$, on which we could extend g on a segment $[0,1]$. Let's go with the proof.

a. Since $f \in \mathcal{S}$, we can freely interchange expectations and derivatives. Doing this, we got (derivative of different compositions, the square, the P_t operator whose derivative is L) :

$$g'(t) = \frac{-1}{t} \mathbb{E}\left[\left(P_{\frac{-\ln(t)}{2}} f(N)\right)\left(L P_{\frac{-\ln(t)}{2}} f(N)\right)\right].$$

We use the previous propositions to conclude. Since $L = -\delta D$, we have :

$$g'(t) = \frac{1}{t} \mathbb{E}\left[\left(P_{\frac{-\ln(t)}{2}} f(N)\right)\left(\delta D P_{\frac{-\ln(t)}{2}} f(N)\right)\right].$$

By definition of δ as the adjoint of the derivative :

$$g'(t) = \frac{1}{t} \mathbb{E}\left[\left(D P_{\frac{-\ln(t)}{2}} f(N)\right)^2\right].$$

By using $DP_t = e^{-t} P_t D$:

$$g'(t) = \frac{1}{t}\left(e^{\frac{-\ln(t)}{2}}\right)^2 \mathbb{E}\left[\left(P_{\frac{-\ln(t)}{2}} f'(N)\right)^2\right].$$

We finally have :

$$g'(t) = \mathbb{E}\left[\left(P_{\frac{-\ln(t)}{2}} f'(N)\right)^2\right].$$

We have by replacing f by $f^{(p)}$ the same expression for all the derivatives of g.

b. By definition of P, we have

$$g(1) = \mathbb{E}\left[P_0 f(N)^2\right] = \mathbb{E}\left[f(N)^2\right].$$

Moreover by dominated convergence, we have

$$g(t) \xrightarrow[t \to 0^+]{} \mathbb{E}\left[P_\infty f(N)^2\right] = \mathbb{E}[f(N)]^2.$$

We can extend g by continuity on $[0,1]$ by those two little computations. We can do the same for all the derivatives of g, since all the $f^{(p)}$ belongs to $\mathcal{S}$. We can as a consequence use the Taylor formula with integral reminder on $[0,1]$:

$$g(0) = g(1) + \sum_{p=1}^{N} \frac{(-1)^p g^{(p)}(1)}{p!} - \int_0^1 \frac{t^N}{N!} g^{(N+1)}(t) \, \mathrm{d}t.$$

Let us show that the integral goes to 0 as $[N \to +\infty]$. To do this, we use the contraction property of P_t :

$$0 \leqslant g^{(N+1)}(t) \leqslant \mathbb{E}\left[f^{(p+1)}(N)^2\right].$$

Hence, by integration, we get :

$$0 \leqslant \int_0^1 \frac{t^N}{N!} g^{(N+1)}(t) \, \mathrm{d}t \leqslant \left(\int_0^1 \frac{t^N}{N!} \, \mathrm{d}t\right) \mathbb{E}\left[f^{(p+1)}(N)^2\right].$$

That is :

$$0 \leqslant \int_0^1 \frac{t^N}{N!} g^{(N+1)}(t) \, \mathrm{d}t \leqslant \frac{\mathbb{E}\left[f^{(p+1)}(N)^2\right]}{(N+1)!} \xrightarrow[N \to +\infty]{} 0.$$

Finally, the integral goes to 0, and we get :

$$g(1) - g(0) = -\sum_{p=1}^{+\infty} \frac{(-1)^p g^{(p)}(1)}{p!}.$$

In terms of probabilistic objects, it means that

$$\mathrm{Var}(f(N)) = \mathbb{E}[f(N)^2] - \mathbb{E}[f(N)]^2 = \sum_{p=1}^{+\infty} \frac{(-1)^{p+1}}{p!} \mathbb{E}\left[f^{(p)}(N)^2\right].$$

Note : We can find **1.** by this proof, with the same Taylor expansion, but around 0 this time :

$$g(1) - g(0) = \sum_{p=1}^{+\infty} \frac{g^{(p)}(0)}{p!},$$

which is exactly what we want :

$$\mathrm{Var}(f(N)) = \sum_{p=1}^{+\infty} \frac{\mathbb{E}\left[f^{(p)}(N)\right]^2}{p!}.$$

Warning, to prove that the integral remainder goes to 0, this proof uses the hypothesis of the convergence of $\frac{\mathbb{E}\left[f^{(n)}(N)^2\right]}{n!}$ to 0, hypothesis we don't make to prove **1.**. □

We can show a third expansion by using the same argument as this proof.

Corollary I.5 : A third expansion of the variance

Let $N \sim \mathcal{N}(0,1)$ and $f \in \mathbb{D}^{\infty,2}$. We suppose that

$$\frac{\mathbb{E}\left[f^{(n)}(N)^2\right]}{2^n n!} \xrightarrow[n\to+\infty]{} 0.$$

Then we have the following expansion :

$$\operatorname{Var}(f(N)) = \sum_{n=1}^{+\infty} \frac{1}{2^n n!} \left(\mathbb{E}\left[f^{(n)}(N)\right]^2 + (-1)^{n+1}\mathbb{E}\left[f^{(n)}(N)^2\right] \right).$$

Proof : Let us introduce g as the previous proof. Then, by a Taylor formula with integral reminder around $\frac{1}{2}$, we got for all $n \in \mathbb{N}^*$:

$$g\left(\frac{1}{2}\right) = g(0) + \sum_{p=1}^{n} \frac{g^{(p)}(0)}{2^p p!} + \int_0^{\frac{1}{2}} \frac{g^{(n+1)}(t) t^n}{n!}\,\mathrm{d}t.$$

By the computation of $g^{(n)}$ made on the previous proof, we have the following estimation for the integral :

$$0 \leqslant \int_0^{\frac{1}{2}} \frac{g^{(n+1)}(t)t^n}{n!}\,\mathrm{d}t \leqslant \frac{\mathbb{E}\left[f^{(n+1)}\right]}{2^{n+1}(n+1)!}$$

Which converges to 0 when $[n \to +\infty]$. We do the same expend, but this time at 1 :

$$g\left(\frac{1}{2}\right) = g(1) + \sum_{p=1}^{n} \frac{(-1)^p g^{(p)}(1)}{2^p p!} - \int_{\frac{1}{2}}^{1} \frac{g^{(n+1)}(t)(1-t)^n}{n!}\,\mathrm{d}t.$$

Since

$$\int_{\frac{1}{2}}^{1} \frac{(1-t)^n}{n!}\,\mathrm{d}t = \int_0^{\frac{1}{2}} \frac{t^n}{n!}\,\mathrm{d}t.$$

So, we also have the convergence to 0 of this integral. We finally have the two expansions :

$$g\left(\frac{1}{2}\right) = \sum_{p=0}^{+\infty} \frac{g^{(p)}(0)}{2^p p!} = \sum_{p=0}^{+\infty} \frac{(-1)^p g^{(p)}(1)}{2^p p!}.$$

We subtract to have $g(1) - g(0) = \operatorname{Var}(f(N))$:

$$\operatorname{Var}(f(N)) = \sum_{p=1}^{+\infty} \frac{g^{(p)}(0)}{2^p p!} + \frac{(-1)^{p+1} g^{(p)}(1)}{2^p p!}.$$

By the expression of $g^{(p)}$ in terms of $f^{(p)}$, this is exactly the claimed equality. □

I.4.2 Poincaré inequalities

We present here two Poincaré inequalities, here's the one in the first order.

Proposition I.14 : First order Poincaré inequality

Let $N \sim \mathcal{N}(0,1)$ and $f \in \mathbb{D}^{1,2}$. Then,

$$\operatorname{Var}(f(N)) \leqslant \mathbb{E}\left[f'(N)^2\right]$$

Proof : By continuity with the norm associated with $\mathbb{D}^{1,2}$, we just have to check this relation for $f \in \mathcal{S}$. By definition of the variance :

$$\operatorname{Var}(f(N)) = \mathbb{E}\left[f(N)^2\right] - \mathbb{E}[f(N)]^2.$$

We write it as an expectation of a product.

$$\operatorname{Var}(f(N)) = \mathbb{E}\left[f(N)\left(f(N) - \mathbb{E}[f(N)]\right)\right].$$

And we notice that $P_0 f = f$ and $P_\infty f = \mathbb{E}[f(N)]$, by dominated convergence. Then, we have :

$$\operatorname{Var}(f(N)) = \mathbb{E}\left[f(N)\left(P_0 f(N) - P_\infty f(N)\right)\right].$$

And we write this as in integral :

$$\operatorname{Var}(f(N)) = -\mathbb{E}\left[f(N) \int_0^{+\infty} \frac{\mathrm{d}P_t f}{\mathrm{d}t}(N)\,\mathrm{d}t\right].$$

We switch integral and expectation (possible since $f \in \mathcal{S}$) :

$$\operatorname{Var}(f(N)) = -\int_0^{+\infty} \mathbb{E}\left[f(N)\frac{\mathrm{d}P_t f}{\mathrm{d}t}(N)\right]\,\mathrm{d}t.$$

By the relation $\frac{\mathrm{d}P_t f}{\mathrm{d}t} = LP_t f$:

$$\operatorname{Var}(f(N)) = -\int_0^{+\infty} \mathbb{E}\left[f(N)[LP_t\ f](N)\right]\,\mathrm{d}t.$$

By the expression of L proved in proposition **I.7** :

$$\operatorname{Var}(f(N)) = \int_0^{+\infty} \mathbb{E}\left[f(N)[\delta DP_t\ f](N)\right]\,\mathrm{d}t.$$

By the expression of DP_t proved in proposition **I.5** :

$$\operatorname{Var}(f(N)) = \int_0^{+\infty} e^{-t}\mathbb{E}\left[f(N)[\delta P_t\ f'](N)\right]\,\mathrm{d}t.$$

By definition of δ as the adjoint of the derivative :

$$\operatorname{Var}(f(N)) = \int_0^{+\infty} e^{-t}\mathbb{E}\left[f'(N)[P_t f'](N)\right]\,\mathrm{d}t.$$

By Cauchy-Schwarz :

$$\operatorname{Var}(f(N)) \leqslant \int_0^{+\infty} e^{-t}\mathbb{E}\left[f'(N)^2\right]^{\frac{1}{2}}\mathbb{E}\left[[P_t f'](N)^2\right]^{\frac{1}{2}}\,\mathrm{d}t.$$

By contraction property of $(P_t)_t$:

$$\mathrm{Var}(f(N)) \leqslant \mathbb{E}\left[f'(N)^2\right] \int_0^{+\infty} e^{-t} \, \mathrm{d}t.$$

And we successfully proved that :

$$\mathrm{Var}(f(N)) \leqslant \mathbb{E}\left[f'(N)^2\right].$$

□

This first order inequality means that if f' is "small", then $f(N)$ is concentrated around its mean. The second order will precise how : if f'' is small too, then $f(N)$ is quite close to a Gaussian distribution. Let us precise this.

Definition I.7

Let F, N two integrable random variables. Then we introduce the distance between the laws of F and N like this : the *Wasserstein distance* is defined by

$$d_{\mathrm{W}}(F,N) \overset{\text{def.}}{=} \sup_{\phi \in \mathsf{Lip}(1)} \left|\mathbb{E}[\phi(F)] - \mathbb{E}[\phi(N)]\right|,$$

where $\mathsf{Lip}(1)$ stands for the set of 1-Lipschitz functions on $\mathbb{R}$.

This distance is still finite, since we have supposed F, N integrable.

Proposition I.15 : Second order Poincaré inequality

Let $N \sim \mathcal{N}(0,1)$ and $f \in \mathbb{D}^{2,4}$ (we recall that means that f is limit of elements of $\mathcal{S}$ for the norm $\left(\|\cdot\|_{L^4(\gamma)}^4 + \|D\cdot\|_{L^4(\gamma)}^4 + \|D^2\cdot\|_{L^4(\gamma)}^4\right)^{\frac{1}{4}}$. Then, if we assume that $\mathbb{E}[f(N)] = 0$ and $\mathbb{E}\left[f(N)^2\right] = 1$. Then,

$$d_{\mathrm{W}}(N, f(N)) \leqslant \frac{3}{\sqrt{2\pi}} \left(\mathbb{E}\left[f'(N)^4\right] \cdot \mathbb{E}\left[f''(N)^4\right]\right)^{\frac{1}{4}}.$$

To prove it, we will use a main result about this distance, we will prove it in the second part, about the Stein's method.

Lemma I.2 : Stein's bound

Let $N \sim \mathcal{N}(0,1)$ and F an integrable random variable. Then

$$d_{\mathrm{W}}(F,N) \leqslant \sup_{\phi \in \mathcal{C}^1 \cap \mathsf{Lip}\left(\sqrt{\frac{2}{\pi}}\right)} \left|\mathbb{E}\left[F\phi(F)\right] - \mathbb{E}\left[\phi'(F)\right]\right|.$$

Proof : Here's the steps. Let $\phi \in \mathcal{C}^1 \cap \mathsf{Lip}\left(\frac{2}{\pi}\right)$ and $f \in \mathcal{S}$.

a. We interpret the first term with P_t :

$$\mathbb{E}[f(N)\phi(f(N))] = \mathbb{E}\left[\phi'(f(N))f'(N) \int_0^{+\infty} e^{-t} P_t f'(N) \, \mathrm{d}t\right].$$

b. We get the first inequality :

$$\begin{aligned} &\left|\mathbb{E}\left[F\phi(F)\right] - \mathbb{E}\left[\phi'(F)\right]\right| \\ \leqslant \;& \sqrt{\frac{2}{\pi}} \sqrt{\mathrm{Var}\left(f'(N) \int_0^{+\infty} e^{-t} P_t f'(N) \, \mathrm{d}t\right)}. \end{aligned}$$

c. Using the first order Poincaré inequality, the contraction property and classical inequalities, we conclude the estimation of the square root of this variance.

Let us begin the proof.

a. We use the usual trick about $P_t f$, with $P_0 f = f$ and $P_\infty f = \mathbb{E}[f(N)]$. Since $\mathbb{E}[f(N)] = 0$, it gives here

$$\mathbb{E}[f(N)\phi(f(N))] = \mathbb{E}\left[\left(P_0 f(N) - P_\infty f(N)\right)\phi(f(N))\right].$$

By the integral interpretation, and since $\frac{\mathrm{d}}{\mathrm{d}t} P_t = L$:

$$\mathbb{E}[f(N)\phi(f(N))] = -\mathbb{E}\left[\left(\int_0^{+\infty} L P_t f(N) \, \mathrm{d}t\right)\phi(f(N))\right].$$

Since $L = -\delta D$, and by switching integral and expectation :

$$\mathbb{E}[f(N)\phi(f(N))] = \int_0^{+\infty} \mathbb{E}\left[[\delta D P_t] f(N)\phi(f(N))\right] \, \mathrm{d}t.$$

By definition of δ :

$$\mathbb{E}[f(N)\phi(f(N))] = \int_0^{+\infty} \mathbb{E}\left[[D P_t] f(N) f'(N) \phi'(f(N))\right] \, \mathrm{d}t.$$

And by the relation $DP_t = e^{-t} P_t D$:

$$\mathbb{E}[f(N)\phi(f(N))] = \int_0^{+\infty} e^{-t} \mathbb{E}\left[P_t f'(N) f'(N) \phi'(f(N))\right] \, \mathrm{d}t.$$

Finally, we switch again to have the following expression :

$$\mathbb{E}[f(N)\phi(f(N))] = \mathbb{E}\left[f'(N)\phi'(f(N)) \int_0^{+\infty} e^{-t} P_t f'(N) \, \mathrm{d}t\right].$$

b. By **a.**, we can factorize by $\phi'(f(N))$ in the computation of what is inside the supremum :

$$\begin{aligned} &\left|\mathbb{E}\left[F\phi(F)\right] - \mathbb{E}\left[\phi'(F)\right]\right| \\ \leqslant \;& \left|\mathbb{E}\left[\phi'(f(N)) \left(f'(N) \int_0^{+\infty} e^{-t} P_t f'(N) \, \mathrm{d}t - 1\right)\right]\right|. \end{aligned}$$

By triangular inequality, and since ϕ is $\sqrt{\frac{2}{\pi}}$-lipschitzian,

$$\begin{aligned}&\left|\mathbb{E}\left[F\phi(F)\right]-\mathbb{E}\left[\phi'(F)\right]\right|\\ \leqslant\ &\sqrt{\frac{2}{\pi}}\mathbb{E}\left[\left|f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t-1\right|\right].\end{aligned}$$

However,

$$\mathbb{E}\left[f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t\right]=1.$$

Indeed, if we apply **a.** with $\phi=\mathrm{id}_{\mathbb{R}}$, then

$$\mathbb{E}[f(N)^2]=\mathbb{E}\left[f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t\right]$$

and since we suppose that $\mathbb{E}[f(N)^2]=1$, we can conclude. Back to our inequality, we use the concavity of the square root, like this :

$$\begin{aligned}&\left|\mathbb{E}\left[F\phi(F)\right]-\mathbb{E}\left[\phi'(F)\right]\right|\\ \leqslant\ &\sqrt{\frac{2}{\pi}}\mathbb{E}\left[\sqrt{\left|f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t-1\right|^2}\right].\end{aligned}$$

By Jensen inequality for concavity, we get :

$$\begin{aligned}&\left|\mathbb{E}\left[F\phi(F)\right]-\mathbb{E}\left[\phi'(F)\right]\right|\\ \leqslant\ &\sqrt{\frac{2}{\pi}}\sqrt{\mathbb{E}\left[\left|f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t-1\right|^2\right]}.\end{aligned}$$

This gives us the wanted inequality :

$$\begin{aligned}&\left|\mathbb{E}\left[F\phi(F)\right]-\mathbb{E}\left[\phi'(F)\right]\right|\\ \leqslant\ &\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}\left[f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t\right]}.\end{aligned}$$

c. Let us note

$$F(x)=f'(x)\int_0^{+\infty}e^{-t}P_tf'(x)\ \mathrm{d}t$$

Then, by first order Poincaré inequality :

$$\mathrm{Var}(F(N))\leqslant\mathbb{E}\left[F'(N)^2\right].$$

We compute F', using the D operator and the relation $DP_t=e^{-t}P_tD$:

$$\begin{aligned}F'(x)=\ &f''(x)\int_0^{+\infty}e^{-t}P_ff'(x)\ \mathrm{d}t\\ +\ &f'(x)\int_0^{+\infty}e^{-2t}P_tf''(x)\ \mathrm{d}t\end{aligned}$$

In our case, it writes

$$\begin{aligned}&\sqrt{\mathrm{Var}\left[f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t\right]}\\ \leqslant\ &\mathbb{E}\left[\left(f''(N)\int_0^{+\infty}e^{-t}P_ff'(N)\ \mathrm{d}t\right.\right.\\ &\left.\left.+f'(x)\int_0^{+\infty}e^{-2t}P_tf''(x)\ \mathrm{d}t\right)^2\right]^{\frac{1}{2}}.\end{aligned}$$

By triangular inequality on L^2 :

$$\begin{aligned}&\sqrt{\mathrm{Var}\left[f'(N)\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t\right]}\\ \leqslant\ &\mathbb{E}\left[\left(f''(N)\int_0^{+\infty}e^{-t}P_ff'(N)\ \mathrm{d}t\right)^2\right]^{\frac{1}{2}}\\ &+\mathbb{E}\left[\left(f'(x)\int_0^{+\infty}e^{-2t}P_tf''(x)\ \mathrm{d}t\right)^2\right]^{\frac{1}{2}}.\end{aligned}$$

Let's call the first term (A), and the second (B). Then, by Cauchy-Schwarz inequality :

$$(A)\leqslant\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\mathbb{E}\left[\left(\int_0^{+\infty}e^{-t}P_tf'(N)\ \mathrm{d}t\right)^4\right]^{\frac{1}{4}}.$$

By Jensen inequality :

$$(A)\leqslant\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\mathbb{E}\left[\int_0^{+\infty}e^{-4t}P_tf'(N)^4\ \mathrm{d}t\right]^{\frac{1}{4}}.$$

We switch the second expectation and integral :

$$(A)\leqslant\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\left(\int_0^{+\infty}e^{-4t}\mathbb{E}\left[P_tf'(N)^4\right]\ \mathrm{d}t\right)^{\frac{1}{4}}.$$

By contraction property of P_t in $L^4(\gamma)$:

$$(A)\leqslant\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\mathbb{E}\left[f'(N)^4\right]^{\frac{1}{4}}\left(\int_0^{+\infty}e^{-4t}\ \mathrm{d}t\right)^{\frac{1}{4}}.$$

The explicit computation of the integral finally yields to

$$(A)\leqslant\frac{\sqrt{2}}{2}\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\mathbb{E}\left[f'(N)^4\right]^{\frac{1}{4}}.$$

By the same argument,

$$(B)\leqslant\frac{1}{2^{\frac{3}{4}}}\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\mathbb{E}\left[f'(N)^4\right]^{\frac{1}{4}}.$$

And finally

$$\begin{aligned}&\left|\mathbb{E}\left[f(N)\phi(f(N))\right]-\mathbb{E}\left[\phi'(f(N))\right]\right|\\ \leqslant\ &\frac{\sqrt{2}+2^{\frac{1}{4}}}{\sqrt{2\pi}}\mathbb{E}\left[f''(N)^4\right]^{\frac{1}{4}}\mathbb{E}\left[f'(N)^4\right]^{\frac{1}{4}}.\end{aligned}$$

Which gives us a better constant than the one claimed by the proposition. We proved the inequality for $f\in\mathcal{S}$. Both size are continuous with respect to f, for the $\|\cdot\|_{\mathbb{D}^{2,4}}$, so we can extend this for all $f\in\mathbb{D}^{2,4}$. Since the right side of the inequality does not depends on ϕ, we conclude for the inequality. □

In the same way, we can find using the $(P_t)_t$ semi-group an estimation of the covariance of two functions of random Gaussian variables.

Proposition I.16 : Estimation of a covariance

Let $(N, \tilde{N})$ a Gaussian couple centered, with $N, \tilde{N} \sim \mathcal{N}(0,1)$ and let ρ their covariance. Then, for all $f, g \in \mathbb{D}^{1,2}$:

$$|\mathrm{Cov}(f(N), g(\tilde{N}))| \leqslant |\rho| \ \mathbb{E}\left[f'(N)^2\right]^{\frac{1}{2}} \mathbb{E}\left[g'(N)^2\right]^{\frac{1}{2}}$$

Proof : Here's the plan. If $\rho = 0$, N and $\tilde{N}$ are independent and so are $f(N)$ and $g(\tilde{N})$, the equality is true. We suppose $|\rho| > 0$. Let $f, g \in \mathcal{S}$, the inequality will remain true by density.

1. We show the following expression for the covariance :

$$\mathrm{Cov}(f(N), g(\tilde{N})) = \mathbb{E}\left[f(N)\left(P_{-\ln|\rho|}g(\tilde{N}) - P_{\infty}g(\tilde{N})\right)\right]$$

2. We conclude using the usual relations between P_t, D, L, δ.

Let us begin.

1. We go from the right hand side (even if this make the proof less intuitive). We call it (B). By the interpretation of P_t using the expectation, we have

$$(B) = \mathbb{E}\left[f(N)g\left(e^{\ln|\rho|}N + \sqrt{1 - e^{2\ln|\rho|}}Z\right)\right] - \mathbb{E}[f(N)]\mathbb{E}[g(\tilde{N})],$$

where $Z \sim \mathcal{N}(0,1)$ is independent of N. (B) rewrites :

$$(B) = \mathbb{E}\left[f(N)g\left(|\rho|N + \sqrt{1-\rho^2}Z\right)\right] - \mathbb{E}[f(N)]\mathbb{E}[g(\tilde{N})].$$

We can compute the characteristic function of what is inside g to show that

$$|\rho|N + \sqrt{1-\rho^2}Z \sim \mathcal{N}(0,1).$$

Hence,

$$(B) = \mathbb{E}\left[f(N)g\left(\tilde{N}\right)\right] - \mathbb{E}[f(N)]\mathbb{E}[g(\tilde{N})].$$

So (B) is the covariance of $f(N)$ and $g(\tilde{N})$.

2. We write this differences of P_t with an integral. We keep noting (B) the quantity of **1.**.

$$(B) = -\mathbb{E}\left[f(N)\int_{-\ln|\rho|}^{+\infty} LP_t g(\tilde{N}) \ \mathrm{d}t\right].$$

By switching expectation and integral :

$$(B) = -\int_{-\ln|\rho|}^{+\infty} \mathbb{E}\left[f(N)LP_t g(\tilde{N})\right] \ \mathrm{d}t$$

By the expression $L = -\delta D$:

$$(B) = \int_{-\ln|\rho|}^{+\infty} \mathbb{E}\left[f(N)[\delta D P_t g](\tilde{N})\right] \ \mathrm{d}t$$

By definition of δ :

$$(B) = \int_{-\ln|\rho|}^{+\infty} \mathbb{E}\left[f'(N)[D P_t g](\tilde{N})\right] \ \mathrm{d}t$$

By $DP_t = e^{-t}P_t D$:

$$(B) = \int_{-\ln|\rho|}^{+\infty} e^{-t}\mathbb{E}\left[f'(N)[P_t g'](\tilde{N})\right] \ \mathrm{d}t$$

By Cauchy-Schwarz and by contraction property of P_t on $L^2(\gamma)$:

$$|(B)| \leqslant \mathbb{E}\left[f'(N)^2\right]^{\frac{1}{2}} \mathbb{E}\left[g'(N)^2\right]^{\frac{1}{2}} \int_{-\ln|\rho|}^{+\infty} e^{-t} \ \mathrm{d}t$$

Which is the wanted conclusion. □

I.4.3 Expansion of the heat kernel

Proposition I.17 : Expansion of the heat kernel

For all $\varepsilon > 0$, we consider

$$\forall x \in \mathbb{R}, p_\varepsilon(x) \stackrel{\text{def.}}{=} \frac{1}{\sqrt{2\pi\varepsilon}} e^{\frac{-x^2}{2\varepsilon}}.$$

Then we have the following expansion of p_ε in $L^2(\gamma)$:

$$p_\varepsilon = \frac{1}{\sqrt{2\pi(1+\varepsilon)}} \sum_{n=0}^{+\infty} \left(\frac{(-1)^n}{n!(2n)!2^n(1+\varepsilon)^n}\right) H_{2n}$$

Proof : We begin by the expansion :

$$p_\varepsilon = \sum_{n=0}^{+\infty}\left(\int_{\mathbb{R}} p_\varepsilon^{(n)}(x) \ \mathrm{d}\gamma(x)\right)\frac{H_n}{n!}.$$

The goal is to compute this integral.

- Let us see first that we have

$$p_\varepsilon(x) = f\left(\frac{x}{\sqrt{\varepsilon}}\right),$$

where $f(x) = e^{\frac{-x^2}{2}}$. So, we can compute its derivatives thanks to the Rodrigues formula :

$$p_\varepsilon^{(n)}(x) = \frac{1}{\sqrt{\varepsilon}^n}(-1)^n H_n\left(\frac{x}{\sqrt{\varepsilon}}\right)\frac{1}{\sqrt{2\pi\varepsilon}} f\left(\frac{x}{\sqrt{\varepsilon}}\right).$$

Giving :

$$p_\varepsilon^{(n)}(x) = \frac{1}{\sqrt{\varepsilon}^n}(-1)^n H_n\left(\frac{x}{\sqrt{\varepsilon}}\right) p_\varepsilon(x).$$

- To compute the integral, we cannot use directly this expression. The idea is to add a parameter and differentiate it. Here we deal with the convolution :

$$g(u) \stackrel{\text{def.}}{=} \int_{\mathbb{R}} p_\varepsilon(x-u) \ \mathrm{d}\gamma(x).$$

Then g is C^∞ and we can differentiable under the integration symbol :

$$g^{(n)}(u) = (-1)^n \int_{\mathbb{R}} p_\varepsilon^{(n)}(x-u)\, \mathrm{d}\gamma(x).$$

So we just need to input $u = 0$ to have our expression.

- Let us compute explicitly $g(u)$. We have :

$$g(u) = \frac{1}{\sqrt{2\pi\varepsilon}} \int_{\mathbb{R}} \exp\left(\frac{-(x-u)^2}{2\varepsilon}\right) \exp\left(\frac{-x^2}{2}\right) \frac{\mathrm{d}x}{\sqrt{2\pi}}.$$

We simply reduce the trinomial in x :

$$\left(1+\tfrac{1}{\varepsilon}\right)x^2 - \tfrac{2xu}{\varepsilon} + \tfrac{u^2}{\varepsilon} = \tfrac{u^2}{\varepsilon+1} + \left(\sqrt{\tfrac{\varepsilon+1}{\varepsilon}}x - \tfrac{u}{\sqrt{\varepsilon(\varepsilon+1)}}\right)^2.$$

Hence,

$$g(u) = \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(\frac{-u^2}{2(\varepsilon+1)}\right) \cdot \int_{\mathbb{R}} \exp\left(\frac{-1}{2\varepsilon}\left(\sqrt{\varepsilon+1}x - \frac{u}{\sqrt{\varepsilon+1}}\right)\right) \frac{\mathrm{d}x}{\sqrt{2\pi}}.$$

We do $z = \sqrt{\varepsilon+1}x - \frac{u}{\sqrt{\varepsilon+1}}$:

$$g(u) = \frac{1}{\sqrt{2\pi\varepsilon(\varepsilon+1)}} \exp\left(\frac{-u^2}{2(\varepsilon+1)}\right) \int_{\mathbb{R}} \exp\left(\frac{-z^2}{2\varepsilon}\right) \frac{\mathrm{d}z}{\sqrt{2\pi}}.$$

The integral is equal to $\sqrt{\varepsilon}$, so we finally have :

$$g(u) = \frac{1}{\sqrt{2\pi(\varepsilon+1)}} \exp\left(\frac{-u^2}{2(\varepsilon+1)}\right) = p_{1+\varepsilon}(u).$$

- Hence, we have :

$$\int_{\mathbb{R}} p_\varepsilon^{(n)}(x)\, \mathrm{d}\gamma(x) = (-1)^n p_{1+\varepsilon}^{(n)}(x).$$

By the expression we derived of the derivative of p_ε, we get :

$$\int_{\mathbb{R}} p_\varepsilon^{(n)}(x)\, \mathrm{d}\gamma(x) = \frac{H_n(0)}{\left(\sqrt{1+\varepsilon}\right)^n} p_{1+\varepsilon}(0).$$

By the expression of $H_n(0)$ we proved for the expansion of the Hermite polynomials, this integral is zero for n odd. Hence, we have :

$$p_\varepsilon = \frac{1}{\sqrt{2\pi(1+\varepsilon)}} \sum_{n=0}^{+\infty} \frac{(-1)^n}{(\varepsilon+1)^n 2^n n!} \frac{H_{2n}}{(2n)!}.$$

This shows the expansion. □

II Isonormal Gaussian processes and Wiener chaos

II.1 Isonormal Gaussian processes

Definition II.1

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and $\mathcal{H}_1$ a subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\mathcal{H}_1$ is a *Gaussian subspace* if $\mathcal{H}_1$ is closed and contains only zero-mean Gaussian random variables.

Definition II.2

Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ a real separable Hilbert space that we set for all this section. We say that a process $X = \{X(h), h \in \mathcal{H}\}$, indexed by elements of $\mathcal{H}$, is an *isonormal Gaussian process* if :

i. X is a Gaussian process ;

ii. For all $h \in \mathcal{H}$, $\mathbb{E}[X(h)] = 0$;

iii. For all $h, g \in \mathcal{H}$:

$$\mathbb{E}[X(h)X(g)] = \langle h, g\rangle.$$

In the following, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space where $\mathcal{F}$ is the σ-algebra generated by $(X(h))_{h\in\mathcal{H}}$.

Proposition II.1 : An isonormal process is linear

Let X an isonormal Gaussian process on $\mathcal{H}$. Then the map

$$\begin{pmatrix} \mathcal{H} & \longrightarrow & L^2(\Omega) \\ h & \longmapsto & X(h) \end{pmatrix}$$

is a linear isometry. As a consequence, its image $\{X(h), h \in \mathcal{H}\}$ is closed in $L^2(\Omega)$, it is a Gaussian subspace, denoted in the following $\mathfrak{H}_1$.

Proof : • To show the linearity, we just need to expand when $\lambda \in \mathbb{R}$ and $h, g \in \mathcal{H}$:

$$\mathbb{E}\left[|X(\lambda h+g)-\lambda X(h)-X(g)|^2\right]$$

And we use the equality satisfied by the covariance of $X(g)$ and $X(h)$. We conclude as a consequence that $X(\lambda h+g)=\lambda X(h)+X(g)$ on L^2.

• The isometry is true by definition. To show that the image is closed, let us consider a sequence $(Y_n)_n$ of elements of $\mathfrak{H}_1$ converging in L^2 to $Y \in L^2$, and we show that $Y \in \mathfrak{H}_1$. By definition, it exists a sequence $(h_n)_n \in \mathcal{H}^{\mathbb{N}}$ such that $Y_n = X(h_n)$. Then, the sequence $(Y_n)_n$ is a Cauchy sequence. Let $n \in \mathbb{N}$ and $p \in \mathbb{N}$. We have by isometry :

$$\mathbb{E}\left[|Y_{n+p}-Y_n|^2\right] = \|h_{n+p}-h_n\|^2$$

so the sequence $(h_n)_n$ is also a Cauchy sequence in $\mathcal{H}$, so converges. Let consider $Z = X(h)$. Then

$$\mathbb{E}\left[|Z-Y|^2\right] = \lim_{n\to+\infty} \mathbb{E}\left[|Z-Y_n|^2\right]$$

By isometry :

$$\mathbb{E}\left[|Z-Y|^2\right] = \lim_{n\to+\infty} \|h_n-h\|^2 = 0$$

so $Y = X(h)$ on L^2, so $Y \in \mathfrak{H}_1$. □

Corollary II.1 : Corresponding definition

A process $X = (X(h), h\in\mathcal{H})$ is a Gaussian isonormal process if and only if $h \longmapsto X(h)$ is linear, and for all $h \in \mathcal{H}$, X_h is Gaussian and $\mathbb{E}[X_h]=0$.

A little lemma to conclude this part which could be useful for computations in the future.

Lemma II.1 : Independence of a orthogonal family

Let $(e_i)_{i\in I}$ an orthogonal family of $\mathcal{H}$. Then $(X(e_i))_{i\in I}$ is independent.
Moreover, if $(e_i)_{i\in I}$ is an orthonormal family of $\mathcal{H}$, then $(X(e_i))_{i\in I}$ is a family of independent and equally distribued random variables, of law $\mathcal{N}(0,1)$.

Proof of the lemma : By this definition, if $(e_i)_{i\in I}$ is an orthonormal family of $\mathcal{H}$ then $(X(e_i))_{i\in I}$ are equally distributed (of law $\mathcal{N}(0,1)$) and are two by two independent. But, since $(X(h))_{h\in\mathcal{H}}$ is Gaussian process, this is equivalent to say that the family $(X(e_i))_{i\in I}$ is (mutually) independent and equally distributed. □

II.2 Wiener chaos

Recall that H_p stands for the p-th Hermite polynomial.

Lemma II.2 : Relation between Gaussian variables and Hermite polynomials

Let (X,Y) a centered Gaussian couple such that $\mathbb{E}[X^2]=\mathbb{E}[Y^2]=1$. Then

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

Proof of the lemma : • By the exponential formula for Hermite polynomials, we will compute for $s,t\in\mathbb{R}$:

$$\mathbb{E}\left[e^{sX-\frac{s^2}{2}}e^{tY-\frac{t^2}{2}}\right] = e^{\frac{-s^2}{2}}e^{\frac{-t^2}{2}}\mathbb{E}\left[e^{sX+tY}\right].$$

To compute this expectation, we expend in power series and us the following fact : if $N\sim\mathcal{N}(0,1)$ then $\mathbb{E}[N^k]=0$ if k is odd, and

$$\mathbb{E}\left[N^{2k}\right] = \frac{(2k)!}{2^k k!}.$$

Since

$$\mathrm{Var}(sX+tY) = (s+t)^2 + 2st(\rho-1),$$

where $\rho = \mathbb{E}[XY]$, we have

$$\mathbb{E}\left[(sX+tY)^{2k}\right] = \frac{(2k)!}{k!2^k}\left(\frac{(s+t)^2}{2}+st(\rho-1)\right)^k.$$

We finally have :

$$\mathbb{E}\left[e^{sX-\frac{s^2}{2}}e^{tY-\frac{t^2}{2}}\right] = e^{\frac{-s^2}{2}}e^{\frac{-t^2}{2}}\sum_{k=0}^{+\infty}\frac{(2k)!}{(2k)!k!2^k}\left(\frac{(s+t)^2}{2}+st(\rho-1)\right)^k.$$

We can rewrite it as :

$$\mathbb{E}\left[e^{sX-\frac{s^2}{2}}e^{tY-\frac{t^2}{2}}\right] = e^{\frac{-s^2}{2}}e^{\frac{-t^2}{2}}e^{st(\rho-1)}e^{\frac{(s+t)^2}{2}}.$$

And finally :

$$\mathbb{E}\left[e^{sX-\frac{s^2}{2}}e^{tY-\frac{t^2}{2}}\right] = e^{st\rho}.$$

• To conclude, we use the exponential formula, and we need to differentiate m times in s, n times in t, and take $s=t=0$. The left hand side becomes after differentiation :

$$\frac{\partial^{n+m}}{\partial^m s \partial^n t}\text{LHS} = \mathbb{E}\left[\left(\sum_{p=0}^{+\infty}\frac{(p+m)!}{(p+m)!p!}s^p H_{p+m}(X)\right) \cdot \left(\sum_{q=0}^{+\infty}\frac{(q+n)!}{(q+n)!q!}t^q H_{q+n}(Y)\right)\right]$$

Taking $s=t=0$ gives :

$$\left.\frac{\partial^{n+m}}{\partial^m s\partial^n t}\text{LHS}\right|_{s=t=0} = \mathbb{E}\left[H_m(X)H_n(Y)\right].$$

For the right hand side, we use Leibniz formula. Suppose without loss of generality that $m \geqslant n$. We have

$$\frac{\partial^m}{\partial s^m}e^{ts\rho} = \rho^m t^m e^{ts\rho}.$$

Then,

$$\frac{\partial^m}{\partial s^m}e^{ts\rho} = \sum_{k=0}^{n}\binom{n}{k}\frac{m!}{(m-k)!}t^{m-k}s^{n-k}\rho^{m+n-k}.$$

We take $s=0$ first :

$$\left.\frac{\partial^m}{\partial s^m}e^{ts\rho}\right|_{s=0} = \frac{m!}{(m-n)!}t^{m-n}\rho^m.$$

Taking $t=0$ and we have two options : it turns 0 if $n \neq m$ and 1 if $n=m$:

$$\left.\frac{\partial^m}{\partial s^m}e^{ts\rho}\right|_{t=s=0} = m!\rho^m\delta_{m,n}.$$

Finally, we get :

$$\mathbb{E}[H_n(X)H_m(X)] = m!\rho^m\delta_{m,n}.$$

Which is what we expect. □

Lemma II.3 : Density of the image par Hermite polynomials of isonormal Gaussian process

Let X an isonormal Gaussian process. Then

$$\{H_p(X(h)), p\in\mathbb{N}, h\in\mathcal{H}, \|h\|=1\}$$

is dense in L^q, for all $q \in [1,+\infty[$.

Proof of the lemma : We try to copy the proof of proposition **I.1**. Let q' the conjugate of q. We consider $Z \in L^{q'}$ such that for all $p\in\mathbb{N}$, for all $h\in\mathcal{H}$ with $\|h\|=1$:

$$\mathbb{E}\left[ZH_p(X(h))\right] = 0$$

Let us prove that $Z=0$ on $L^{q'}$. This fact is equivalent to for all $p\in\mathbb{N}$, for all $h\in\mathcal{H}$ with $\|h\|=1$:

$$\mathbb{E}\left[ZX(h)^p\right] = 0$$

We use the Fourier transform. By Holder inequality, we can justify with Fubini theorem that we have for all $u\in\mathbb{R}$:

$$\mathbb{E}\left[Ze^{\mathrm{i}uX(h)}\right] = 0$$

Since $\mathcal{H}$ is real and separable, the Hilbert theory implies that $\mathcal{H}$ admits a countable orthonormal basis. We denote it by $(e_i)_i$:

$$\mathcal{H} = \bigoplus_{i\in\mathbb{N}^*}^{\perp}\mathbb{R}e_i.$$

We note also $\mathcal{F}_m$ the σ-algebra generated by $(X(e_1),\cdots,X(e_m))$. Let us show that

$$\forall m\in\mathbb{N}^*, \mathbb{E}[Z|\mathcal{F}_m] = 0.$$

We have for all $\lambda_1,\cdots,\lambda_m\in\mathbb{R}$:

$$\mathbb{E}\left[Z\exp\left(\mathrm{i}\sum_{j=1}^{n}\lambda_j X(e_j)\right)\right] = 0.$$

Since the exponential is $\mathcal{F}_m$-measurable, this means that :

$$\mathbb{E}\left[\mathbb{E}[Z|\mathcal{F}_m]\exp\left(\mathrm{i}\sum_{j=1}^{n}\lambda_j X(e_j)\right)\right] = 0.$$

Moreover, by $\mathcal{F}_m$-measurability of $\mathbb{E}[Z|\mathcal{F}_m]$, there exists a function $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R}$ measurable such that

$$\mathbb{E}[Z|\mathcal{F}_m] = \varphi(X(e_1),\cdots,X(e_m)).$$

Hence, the equality gives that for all $\lambda_1,\cdots,\lambda_m\in\mathbb{R}$:

$$\int_{\mathbb{R}^m}\varphi(x_1,\cdots,x_m)e^{\frac{-1}{2}\sum_{j=1}^m x_j^2}e^{\mathrm{i}\sum_{j=1}^m \lambda_j x_j} = 0.$$

We used the fact that $\mathbb{E}[X(e_i)X(e_j)] = \delta_{i,j}$. This equality implies that the Fourier transform of

$$(x_1,\cdots,x_m)\longmapsto \varphi(x_1,\cdots,x_m)e^{\frac{-1}{2}\sum_{j=1}^m x_j^2}$$

is the null function. By injectivity, it means that φ is the null function, and so that

$$\forall m\in\mathbb{N}^*, \mathbb{E}[Z|\mathcal{F}_m] = 0.$$

This means that

$$\mathbb{E}[Z|\mathcal{F}] = 0$$

But since Z is $\mathcal{F}$-measurable, we conclude that $Z=0$ in $L^{q'}$. Hence, by the lemma **I.1.**, the family generated by $H_p(X(h))$ is dense in L^q. □

Theorem II.1 : Wiener-Itô decomposition

Let X an isonormal Gaussian process on an real separable Hilbert space $\mathcal{H}$. For all $n\in\mathbb{N}$, we note

$$\mathfrak{H}_n \overset{\text{def.}}{=} \overline{\text{Vect}\left\{H_n(X(h)), h\in\mathcal{H}, \|h\|=1\right\}}$$

the n-th *chaos of Wiener*. Then, we have the following decomposition :

$$L^2(\Omega,\mathcal{F},\mathbb{P}) = \bigoplus_{n\in\mathbb{N}}^{\perp} \mathfrak{H}_n.$$

Proof : • Let us show first that $(\mathfrak{H}_m)_m$ are an orthogonal family. Let $n \neq m$ and $F \in \mathfrak{H}_m^0 \cap \mathfrak{H}_n^0$, where

$$\mathfrak{H}_p^0 = \text{Vect}\left\{H_n(X(h)), h \in \mathcal{H}, \|h\| = 1\right\}.$$

Then F admits two expansions of the form $\sum_i \alpha_i H_n(h_i)$ and $\sum_j \beta_j H_m(h_j)$. Hence, it suffices to show that for every $h, g \in \mathcal{H}$:

$$\mathbb{E}[H_n(X(h))H_m(X(g))] = 0$$

And this property is true by lemma **II.2**. Hence, the $\mathfrak{H}_n^0$ are in direct sums, and by closure, the spaces $\mathfrak{H}_n$ are in direct sums..

• We will show the density of $\mathfrak{G} \overset{\text{def.}}{=} \bigoplus_{n\in\mathbb{N}}^{\perp} \mathfrak{H}_n$. Let $Z \in \mathfrak{G}^{\perp}$. Then for all $h \in \mathcal{H}$ with $\|h\| = 1$ and for all $n \in \mathbb{N}$:

$$\mathbb{E}[ZH_n(X(h))] = 0.$$

By lemma **II.3**, the set of all $H_n(X(h))$ is dense in L^2. It means that $Z = 0$. This proves our theorem,. □

As a consequence, we can decompose all L^2 random variables in terms of $\mathfrak{H}_n$.

Definition II.3

Let Z a random variable belonging to L^2. We note $J_n(Z)$ the projection of Z over $\mathfrak{H}_n$. In other words, we note in L^2 :

$$Z = \sum_{n=0}^{+\infty} J_n(Z) = \sum_{n=0}^{+\infty} \text{Proj}\left(Z \,|\mathfrak{H}_n\right).$$

It could be interesting to have a basis of each space $\mathfrak{H}_n$ to have a global basis of L^2 where we can decompose each $J_n(Z)$. We can do it thanks to the Hermite polynomials once again. Before going to it, let us introduce a few notations on multi-indexes.

Definition II.4

We define Λ as set of the sequences taking values into $\mathbb{N}$ almost null :

$$\Lambda \overset{\text{def.}}{=} \left\{a \in \mathbb{N}^{\mathbb{N}} \,\middle|\, \exists n_0 \in \mathbb{N}, \forall n \geqslant n_0, a_n = 0\right\}.$$

For $a \in \Lambda$, we define the *length* of a and the *factorial* of a by :

$$|a| \overset{\text{def.}}{=} \sum_{i=1}^{+\infty} a_i \text{ and } a! \overset{\text{def.}}{=} \prod_{i=1}^{+\infty} a_i!$$

We note Λ_n the set of elements of Λ with length n.

Those two objects are well-defined since there is a finite number of non null elements for a.

The following proposition generalizes the proposition **I.10** about an Hilbert basis of $L^2(\gamma)$ thanks to the Hermite polynomials.

Proposition II.2 : A basis of $\mathfrak{H}_n$

Let $(e_i)_{i\geqslant 1}$ an orthonormal basis of $\mathcal{H}$. For all $a \in \Lambda$, we define

$$\Phi_a \overset{\text{def.}}{=} \frac{1}{\sqrt{a!}} \prod_{i=1}^{+\infty} H_{a_i}(X(e_i)).$$

Then the set $\bigcup_{k\in[\![1,n]\!]}\{\Phi_a, a \in \Lambda_n\}$ is an orthonormal basis (in the hilbertian sense) of $\oplus_{k=1}^n \mathfrak{H}_k$.

Proof : • Let $a, b \in \Lambda$. We show that $\mathbb{E}[\Phi_a\Phi_b] = 0$ when $a \neq b$. To do that, we use the independence of $(X(e_i))_i$ we discussed at the end of the subsection about isonormal Gaussian processes, in the lemma **II.1**. We have :

$$\mathbb{E}[\Phi_a\Phi_b] = \frac{1}{\sqrt{a!}\sqrt{b!}} \prod_{i=1}^{+\infty} \mathbb{E}\left[H_{a_i}(X(e_i))H_{b_i}(X(e_i))\right].$$

By the lemma **II.1**,

$$\mathbb{E}\left[H_{a_i}(X(e_i))H_{b_i}(X(e_i))\right] = a_i!\delta_{a_i,b_i}$$

And so :

$$\mathbb{E}[\Phi_a\Phi_b] = \delta_{a,b}.$$

Hence, $(\Phi_a)_{a\in\Lambda}$ is an orthogonal family of $L^2(\mathbb{P})$.

• For the density, just think that if $a \in \Lambda_n$, Φ_a is a multivariate polynomial with degree n, so can be express as linear combination of Hermite polynomials. By definition, the set of Hermite polynomials up to degree n taken in element of $\mathfrak{H}$ with norm 1 is dense into $\oplus_{p=1}^n \mathfrak{H}_p$. It follows that $(\Phi_a)_{a\in\Lambda_n}$ is itself dense into $\oplus_{p=1}^n \mathfrak{H}_p$. □

Corollary II.2 : A basis of L^2

With the same notations as the previous proposition. The family $\{\Phi_a, a \in \Lambda\}$ is an orthonormal basis of $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Recall that $\mathcal{F}$ is here the σ-algebra generated by the $X(h)$, with $h \in \mathcal{H}$.

II.3 Construction of Itô-Wiener multiple integrals

We take place in $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$, where μ contains no atoms and is σ-finite. In this case, $\mathcal{H}$ is indeed a real separable Hilbert space. We will precise $J_n(Z)$ in this case thanks to *multiple Itô-Wiener* integrals. We will denote $\mathcal{H}$ sometimes by $L^2(\mu)$, and $L^2(\mathbb{P})$ will refer to the square integrable random variables.

Lemma II.4 : Gaussian measure

Let X an isonormal process on $L^2(\mu)$. Then if we define

$$\forall A \in \mathcal{B}, W(A) \overset{\text{def.}}{=} X(\mathbf{1}_A),$$

then W is a *Gaussian measure* on $(T, \mathcal{B}, \mu)$, meaning that $W : \mathcal{B} \longrightarrow L^2(\mathbb{P})$ satisfies :

(i) $W(\varnothing) = 0$;

(ii) If $(A_n)_n \in \mathcal{B}^{\mathbb{N}}$ are disjoints, and satisfies $\sum_n \mu(A_n) < +\infty$ then

$$W\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} W(A_n)$$

in $L^2(\mathbb{P})$;

(iii) If $\mu(A) < +\infty$ then $W(A) \sim \mathcal{N}(0, \mu(A))$.

We set those notations for the rest of this subsection.

To define multiples integral, we define it first on a simple set, which we will be dense in $L^2(\mu^{\otimes m})$. We set $m \geqslant 1$ for the rest of this section.

Definition II.5

Let $f \in L^2(T^m, \mathcal{B}^{\otimes m}, \mu^{\otimes m})$. We say that f is an *elementary function* if there exists $N \in \mathbb{N}$ such that for all $t_1, \cdots, t_m \in T$:

$$f(t_1, \cdots, t_m) = \sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m} \mathbf{1}_{A_{i_1}\times\cdots\times A_{i_m}}(t_1, \cdots, t_m),$$

where :

(i) $a_{i_1,\cdots,i_m} \in \mathbb{R}$ are coefficients, with indexes $i_j \in [\![1, N]\!]$, such that they are equal to zero if there is two equal indexes;

(ii) $A_{i_1}, \cdots, A_{i_m} \in \mathcal{B}$ are two by two disjoint sets with $\mu(A_i) < +\infty$, for all $i \in \{i_m\}_m$.

We note $\mathcal{E}_m$ the set of all elementary functions with m variables. In this case, we note its *Itô-Wiener integral* :

$$I_m(f) = \int_{T^m} f(t_1, \cdots, t_m)\, \mathrm{d}W_{t_1} \cdots \mathrm{d}W_{t_m}$$

defined by :

$$I_m(f) \overset{\text{def.}}{=} \sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m} W(A_{i_1})\cdots W(A_{i_m}).$$

Proposition II.3 : Properties of integrals on elementary functions

Let $f \in \mathcal{E}_m$, defined by the equality of the previous definition.

(i) The definition of $I_m(f)$ does not depend on the representation of f ;

(ii) $I_m : \mathcal{E}_m \longrightarrow L^2(\mathbb{P})$ is linear ;

(iii) If we note, for all $t_1, \cdots, t_m \in T$:

$$\tilde{f}(t_1,\cdots,t_m) \overset{\text{def.}}{=} \frac{1}{m!}\sum_{\sigma\in\mathfrak{S}_n} f\left(t_{\sigma(1)},\cdots,t_{\sigma(m)}\right)$$

the *symmetrization* of f, then $I_m(f) = I_m(\tilde{f})$;

(iv) If $g \in \mathcal{E}_q$, then

$$\mathbb{E}\left[I_m(f)I_q(g)\right] = m!\langle\tilde{f},\tilde{g}\rangle_{L^2(T^m)}\delta_{m,q}.$$

Proof : **(i)** If f admits two representations

$$\begin{aligned} & f(t_1,\cdots,t_m) \\ = & \textstyle\sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m}\mathbf{1}_{A_{i_1}\times\cdots A_{i_m}}(t_1,\cdots,t_m) \\ = & \textstyle\sum_{j_1,\cdots,j_m=1}^{M} b_{j_1,\cdots,j_m}\mathbf{1}_{B_{j_1}\times\cdots B_{j_m}}(t_1,\cdots,t_m) \end{aligned}$$

then we can consider its decomposition on $C_{\phi(i,j)} = A_i\cap B_j$, where ϕ is bijection between $[\![1,N]\!]\times[\![1,M]\!]$ to $[\![1,MN]\!]$. Those sets can be indeed used since the $C_{\phi(i,j)}$ are two by two disjoints and $C_{\phi(i,j)}$ is μ-finite. We note $c_{k_1,\cdots,k_m}$ the values of the coefficient, with $k_p = \phi(i_p,j_p)$:

$$c_{k_1,\cdots,k_m} = a_{i_1,\cdots,i_m} = b_{j_1,\cdots,j_m}.$$

In this case, if we note I_m^A the integral with the A representation, same with I_m^B and I_m^C, we have

$$I_m^C(f) = \sum_{k_1,\cdots,k_m=1}^{MN} c_{k_1,\cdots,k_m} W(C_{k_1})\cdots W(C_{k_m}).$$

With the help of ϕ, it means that :

$$\begin{aligned} I_m^C(f) = & \textstyle\sum_{i_1,\cdots,i_m=1}^{N}\Bigg(\sum_{j_1,\cdots,j_m=1}^{M} c_{\phi(i_1,j_1),\cdots,\phi(i_m,j_m)} \\ & \cdot W(A_{i_1}\cap B_{j_1})\cdots W(A_{i_m}\cap B_{j_m})\Bigg). \end{aligned}$$

By definition of the coefficients c, we obtain :

$$\begin{aligned} I_m^C(f) = & \textstyle\sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m}\Bigg(\sum_{j_1,\cdots,j_m=1}^{M} \\ & \cdot W(A_{i_1}\cap B_{j_1})\cdots W(A_{i_m}\cap B_{j_m})\Bigg). \end{aligned}$$

The last sums decouples :

$$\begin{aligned} I_m^C(f) = & \textstyle\sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m} \\ & \cdot\left(\prod_{k=1}^{m}\sum_{j=1}^{M} W(A_{i_k}\cap B_j)\right). \end{aligned}$$

By additivity of W, since the B_j's are disjoints :

$$\begin{aligned} I_m^C(f) = & \textstyle\sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m} \\ & \cdot\left(\prod_{k=1}^{m} W\left(A_{i_k}\cap\bigsqcup_{j=1}^{M} B_j\right)\right). \end{aligned}$$

And since f is reprenseted by A_i and the B_j, we have μ-almost everywhere

$$\bigsqcup_{j=1}^{M} B_j = \bigsqcup_{i=1}^{N} A_i.$$

Finally, we have :

$$I_m^C(f) = \sum_{i_1,\cdots,i_m=1}^{N} a_{i_1,\cdots,i_m} W(A_{i_1})\cdots W(A_{i_m}).$$

So good as $I_m^C(f) = I_m^A(f)$. By symmetry of the computation, $I_m^C(f) = I_m^B(f)$. This shows that $I_m(f)$ is well-defined.

(ii) Let $f,g \in \mathcal{E}_m$. We can suppose that f,g admits a decomposition on the same sets of $\mathcal{B}$, if not, we just have to consider their intersection. Then, in this case, the linearity just comes from the linearity of the sum.

(iii) By linearity of $f \longmapsto \tilde{f}$, we just need to consider the case where f is a product of indicative functions :

$$f(t_1,\cdots,t_m) = \mathbf{1}_{A_1\times\cdots A_m}(t_1,\cdots,t_m).$$

Then, by linearity of I_m :

$$I_m(\tilde{f}) = \frac{1}{m!}\sum_{\sigma\in\mathfrak{S}_n} I_m(f).$$

Since the inside of the sum does not depends on σ, we finally have $I_m(\tilde{f}) = I_m(f)$.

(iv) By *(iii)*, we just need to consider the case where f,g are symmetric. We can consider that f,g are associated with the same partition $A_1,\cdots,A_n$. It means that

$$f = \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1,\cdots,i_m}\mathbf{1}_{A_{i_1}}\cdots\mathbf{1}_{A_{i_m}}$$

and

$$g = \sum_{j_1,\cdots,j_q=1}^{n} b_{j_1,\cdots,j_q} \mathbf{1}_{A_{j_1}} \cdots \mathbf{1}_{A_{j_q}}.$$

Moreover, the symmetry implies that for all permutation σ, the coefficients indeed by i_k and by $i_{\sigma(k)}$ are the same. It implies that

$$I_m(f) = m! \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} a_{i_1,\cdots,i_m} W(A_{i_1})\cdots W(A_{i_m})$$

Same thing for g. Hence, we can compute $\mathbb{E}[I_m(f)I_q(g)]$:

$$\begin{aligned} & \mathbb{E}[I_m(f)I_q(g)] \\ = \; & m!q! \textstyle\sum_{i\in I}\sum_{j\in J} a_i b_j \mathbb{E}\left[\prod_{k=1}^m \prod_{l=1}^q W(A_{i_k})W(A_{j_l})\right] \end{aligned}$$

where I is the ranged set of $\{i_k\}$ and J the ranged set of $\{j_l\}$. If we suppose that $m > q$, then there exists i_{k_0} such that i_k is not equal to any j_l. By orthogonality of the $\mathbf{1}_{A_j}$ for $L^2(\mu)$, we use the independence of the $W(A_j)$ to obtain that

$$\begin{aligned} & \mathbb{E}\left[\prod_{k=1}^m \prod_{l=1}^q W(A_{i_k})W(A_{j_l})\right] \\ = \; & \mathbb{E}\left[\prod_{\substack{k=1\\k\neq k_0}}^m \prod_{l=1}^q W(A_{i_k})W(A_{j_l})\right] \mathbb{E}[W(A_{i_{k_0}})] \\ = \; & 0 \end{aligned}$$

We got the result for $m \neq q$. If $m = q$, then we will show that the sum in j is null except if the sequence $\{j_l\}$ is exactly $\{i_k\}$. If we suppose that there exists k_0 such that for all k, $i_k \neq j_{k_0}$ than the expectation goes to zero by independence. It remains all the terms such that : for all k, there exists $n(k)$ such that $i_k = j_{n(k)}$. But, since the sequence $\{i_k\}$ is strictly increasing, $k \longmapsto n(k)$ is also strictly increasing. But, the only strictly increasing function from $[\![1,m]\!]$ to itself is the identity. Hence, $i_k = j_k$ for all k, and then

$$\begin{aligned} & \mathbb{E}[I_m(f)I_m(g)] \\ = \; & (m!)^2 \textstyle\sum_{i_1<\cdots<i_m} a_{i_1,\cdots,i_m} b_{i_1,\cdots,i_m} \mathbb{E}\left[\prod_{k=1}^m W(A_{i_k})^2\right]. \end{aligned}$$

This yields to

$$\begin{aligned} & \mathbb{E}[I_m(f)I_m(g)] \\ = \; & (m!)^2 \textstyle\sum_{i_1<\cdots<i_m} a_{i_1,\cdots,i_m} b_{i_1,\cdots,i_m} \mu(A_{i_1})\cdots\mu(A_{i_m}). \end{aligned}$$

So, if we write it for every indexes :

$$\begin{aligned} & \mathbb{E}[I_m(f)I_m(g)] \\ = \; & m! \textstyle\sum_{i_1,\cdots,i_m=1}^n a_{i_1,\cdots,i_m} b_{i_1,\cdots,i_m} \mu(A_{i_1})\cdots\mu(A_{i_m}). \end{aligned}$$

And by definition of the integral for $\mu^{\otimes m}$:

$$\mathbb{E}[I_m(f)I_m(g)] = m!\langle f,g\rangle_{L^2(T^m)}.$$

□

Proposition II.4 : Density of elementary functions

The space $\mathcal{E}_m$ is dense in $L^2(T^m, \mathcal{B}^{\otimes m}, \mu^{\otimes})$.

Proof : • We suppose known the fact that

$$\mathrm{Vect}\left(\mathbf{1}_{A_1\cdots A_m}, A_i \in \mathcal{B}, \mu(A_i) < +\infty\right)$$

is dense in $L^2(\mu^{\otimes m})$. To complete the proof, we consequently show that every function of the form $\mathbf{1}_{A_1\cdots A_m}$ can be approximated by elements from $\mathcal{E}_m$.

• Let $A = A_1\cdots A_m$, with every $A_i \in \mathcal{B}$ satisfying $\mu(A_i) < +\infty$. We note

$$\alpha \overset{\text{def.}}{=} \mu\left(\bigcup_{i=1}^m A_i\right).$$

Let $\varepsilon > 0$. Let us show that there exists $f \in \mathcal{E}_m$ such that

$$\|\mathbf{1}_A - f\|_{L^2(\mu^{\otimes m})} \leqslant \varepsilon.$$

Since μ contains no atoms, there exists $B_1, \cdots, B_n \in \mathcal{B}$ such that

$$\mu(B_i) \leqslant \frac{\varepsilon}{\binom{m}{2}\alpha^{m-1}},$$

and such that for all $i \in [\![1,m]\!]$, the set A_i is partitioned by some B_i. Hence, we write the indicator function of A can be written as :

$$\mathbf{1}_A = \sum_{(i_1,\cdots,i_m)\in[\![1,n]\!]^m} \varepsilon_{i_1,\cdots,i_m} \mathbf{1}_{B_{i_1}\times\cdots\times B_{i_m}},$$

with every ε belonging to $\{0,1\}$. With this decomposition, we can define

$$f \overset{\text{def.}}{=} \sum_{(i_1,\cdots,i_m)\in\Delta} \varepsilon_{i_1,\cdots,i_m} \mathbf{1}_{B_{i_1}\times\cdots\times B_{i_m}}$$

where Δ is the set of the indexes $(i_1,\cdots,i_m)$ containing only different entries :

$$\Delta \overset{\text{def.}}{=} \{(i_1,\cdots,i_m) \in [\![1,n]\!]^m, \forall k \neq l, i_k \neq i_l\}.$$

• We note J its complementary. We have on $L^2(\mu^{\otimes})$:

$$\|\mathbf{1}_A - f\|_{L^2}^2 = \left\| \sum_{(i_1,\cdots,i_m)\in J} \varepsilon_{i_1,\cdots,i_m} \mathbf{1}_{B_{i_1}\times\cdots\times B_{i_m}} \right\|_{L^2}^2$$

By integration of stair functions :

$$\|\mathbf{1}_A - f\|_{L^2}^2 = \sum_{(i_1,\cdots,i_m)\in J} \varepsilon_{i_1,\cdots,i_m}^2 \mu(B_{i_1})\cdots\mu(B_{i_m}).$$

All the ε are lower or equal to 1.

$$\|\mathbf{1}_A - f\|_{L^2}^2 \leqslant \sum_{(i_1,\cdots,i_m)\in J} \mu(B_{i_1})\cdots\mu(B_{i_m}).$$

By definition,

$$J = \{(i_1,\cdots,i_m) \in [\![1,n]\!]^m, \exists k \neq l, i_k = i_l\}.$$

Hence, we can make a partition of J in function of the values taken by the equal indexes :

$$J = \bigsqcup_{i=1}^m \{(i_1,\cdots,i_m) \in [\![1,n]\!]^m, \exists k \neq l, i_k = i_l = i\}.$$

We note J_i the set inside the union. We get until now :

$$\|\mathbf{1}_A - f\|_{L^2}^2 \leqslant \sum_{i=1}^m \sum_{(i_1,\cdots,i_m)\in J_i} \mu(B_{i_1})\cdots\mu(B_{i_m}).$$

- Let $(i_1, \cdots, i_m) \in J_i$. Then, there are $k \neq l$ such that $i_l = i_k = i$. We have the following estimation, by majoring the unknown measure by the sum of all measures :

$$\prod_{\sigma=1}^{m} \mu\left(B_{i_\sigma}\right) \leqslant \mu(B_i)^2 \left(\sum_{j=1}^{m} \mu(B_j)\right)^{m-2}.$$

Since $(B_j)_j$ is a partition of A, we have :

$$\prod_{\sigma=1}^{m} \mu\left(B_{i_\sigma}\right) \leqslant \mu(B_i)^2 \alpha^{m-2}.$$

And since everything is positive, if we estimate one of the two factors in $\mu(B_i)^2$, we got :

$$\prod_{\sigma=1}^{m} \mu\left(B_{i_\sigma}\right) \leqslant \frac{\varepsilon}{\alpha\binom{m}{2}} \mu(B_i).$$

- So, the estimation of the reminder yields to :

$$\|\mathbf{1}_A - f\|_{L^2}^2 \leqslant \frac{\varepsilon}{\alpha\binom{m}{2}} \sum_{i=1}^{m} \mu(B_i) \sharp J_i.$$

Since $\sharp J_i = \binom{m}{2}$, we get :

$$\|\mathbf{1}_A - f\|_{L^2}^2 \leqslant \frac{\varepsilon}{\alpha} \sum_{i=1}^{m} \mu(B_i) = \varepsilon.$$

That shows how the space $\mathcal{E}_m$ is dense in $L^2(T^m)$. □

By this density, we can conclude the properties of I_m on every L^2 function.

Proposition II.5 : Properties of Itô-Wiener integrals on L^2 functions

Let $f \in L^2(T^m)$.

(i) $I_m : L^2(T^m) \longrightarrow L^2(\mathbb{P})$ is a linear continuous map ;

(ii) If $\tilde{f}$ is the symmetrization of f, then $I_m(f) = I_m(\tilde{f})$;

(iii) If $g \in L^2(T^q)$, then

$$\mathbb{E}\left[I_m(f) I_q(g)\right] = m! \langle \tilde{f}, \tilde{g} \rangle_{L^2(T^m)} \delta_{m,q}.$$

Proof : All those properties comes from the continuity of I_m. Its continuity can be shown by taking $f = g$ in **(iv)** of the proposition **II.3**, for all $f \in \mathcal{E}_m$:

$$\mathbb{E}\left[I_m(f)^2\right] = m! \left\|\tilde{f}\right\|_{L^2(T^m)}^2$$

We obtain :

$$\mathbb{E}\left[I_m(f)^2\right] \leqslant m! \|f\|_{L^2(T^m)}^2$$

Hence, I_m is continuous linear on a dense subset of the Hilbert space $L^2(T^m)$. □

Corollary II.3 : Interpretation of an isonormal process in terms of Itô-Wiener integral

Let $h \in L^2(T)$. Then,

$$X(h) = \int_T h(t) \, \mathrm{d}W_t.$$

Proof : The equality is true for every $h \in \mathcal{E}_1$, by definition of W :

$$X(\mathbf{1}_A) = W(A) = \int_T \mathbf{1}_A(t) \, \mathrm{d}W_t.$$

$h \longmapsto X(h)$ is continuous, I_1 is also continuous. We conclude in the equality for every $h \in L^2(T)$. □

As an example, let us consider the case where μ is the product measure product between the Lebesgue measure on $\mathbb{R}_+$ and the count measure on $[\![1, d]\!]$, if we set on $\mathcal{H} = L^2(\mathbb{R}_+ \times [\![1, d]\!], \mu)$, embedded with the inner product :

$$\langle f, g \rangle = \sum_{i=1}^{d} \int_0^{+\infty} f(i,t) g(i,t) \, \mathrm{d}t.$$

Then, if

$$B^i(t) = W([0,t] \times \{i\})$$

then $B^i \sim \mathcal{N}(0,t)$ (since $\mu([0,t] \times \{i\}) = t$). If $i \neq j$, $\mathbb{E}[B^i(t) B^j(t)] = 0$. If $s \leqslant t$:

$$\mathbb{E}\left[B^i(t) B^i(s)\right] = \int_0^{+\infty} \mathbf{1}_{[0,t]}(u) \mathbf{1}_{[0,s]}(u) \, \mathrm{d}u = s.$$

We get the Brownian motion in dimension d.

II.4 Decomposition of square integrable random variables with multiple integrals

We will prove the following theorem in this section.

Theorem II.2 : Decomposition in Wiener chaos

Let X an isonormal Gaussian process on $L^2(T,\mathcal{B},\mu)$. Let $F \in L^2(\Omega,\mathcal{F},\mathbb{P})$. Then, there exists an unique sequence of symmetric elements $f_n \in L^2(T^n,\mathcal{B}^{\otimes n},\mu^{\otimes n})$ such that in L^2 :

$$F = \sum_{n=0}^{+\infty} I_n(f_n),$$

with $I_n(f_n) \in \mathfrak{H}_n$. In other words,

$$\forall n \in \mathbb{N}^*, I_n(f_n) = J_n(F).$$

We will give later an expression of the f_n with the Malliavin derivatives.
We prove now a relation which is close to the relation satisfied by the Hermite polynomials.

Lemma II.5 : A recurrent relation for I_p

Let $f \in L^2(T^p)$ and $g \in L^2(T)$. Then if we note for all $(t_1,\cdots,t_{p-1}) \in T^{p-1}$:

$$[f \otimes_1 g](t_1,\cdots,t_{p-1}) \overset{\text{def.}}{=} \int_T f(t_1,\cdots,t_{p-1},s)g(s)\ \mathrm{d}\mu(s),$$

then I_p satisfies the following relation :

$$I_p(f)I_1(g) = I_{p+1}(f\otimes g) + pI_{p-1}(f\otimes_1 g).$$

Proof of the lemma : By density of elementary functions, and by linearity of I_p, let us consider f as a symmetrization of an indicator function :

$$f = \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p} \mathbf{1}_{A_{\sigma(1)}\times\cdots\times A_{\sigma(p)}}$$

where A_i are two by two disjoints, and $\mu(A_i) < +\infty$. For g, by density of elementary function and by linearity of I_1, we only have two cases to check : if $g = \mathbf{1}_{A_1}$, or if $g = \mathbf{1}_{A_0}$, with A_0 disjoint of the A_i.

• The easier case is when $g = \mathbf{1}_{A_0}$. Indeed, we can compute all the quantities which are in the lemme. First,

$$I_p(f)I_1(g) = W(A_1)\cdots W(A_p)W(A_0).$$

Then,

$$f\otimes_1 g(t_1,\ldots,t_{p-1}) = \int_T f(t_1,\ldots,t_{p-1},s)g(s)\ \mathrm{d}\mu(s),$$

and so $f\otimes_1 g = 0$. Finally,

$$f\otimes g = \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p} \mathbf{1}_{A_{\sigma(1)}\times\cdots\times A_{\sigma(p)}\times A_0}.$$

And so

$$I_{p+1}(f\otimes g) = W(A_0)W(A_1)\ldots W(A_p).$$

We have shown the equality.

• Let us consider $g = \mathbf{1}_{A_1}$. Here's the plan.

1. $I_{p-1}(f\otimes_1 g)$ is still easy to compute.
2. Let $\varepsilon > 0$. Since μ is non-atomic, we consider a partition of A_1 whose elements have measure less than $\frac{\varepsilon}{6\beta}$, where $\beta = \prod_{k=1}^p \mu(A_k)$. We note this partition $(B_1,\cdots,B_n)$. We introduce the function

$$h_\varepsilon = \sum_{i\neq j} \mathbf{1}_{B_i\times B_j\times A_2\times\cdots\times A_p}.$$

Then $h_\varepsilon \in \mathcal{E}_{p+1}$ satisfies

$$\mathbb{E}\left[|I_{p+1}(h_\varepsilon) - I_{p+1}(f\otimes g)|^2\right] \leqslant \frac{\varepsilon}{6}.$$

3. We also have

$$I_p(f)I_1(g) = I_{p+1}(h_\varepsilon) + R_\varepsilon + pI_{p-1}(f\otimes_1 g).$$

Then,

$$\mathbb{E}\left[R_\varepsilon^2\right] \leqslant \frac{\varepsilon}{3}.$$

4. We can conclude on our equality.

Let's prove it.

1. Let us call for $t_1,\cdots,t_{p-1} \in T$:

$$\text{(A)} \overset{\text{def.}}{=} f\otimes_1 g(t_1,\cdots,t_{p-1}).$$

Then, by definition :

$$\text{(A)} = \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}\left(\int_T \mathbf{1}_{A_{\sigma(p)}\cap A_1}(s)\ \mathrm{d}\mu(s)\right) \cdot\mathbf{1}_{A_{\sigma(1)}\times\cdots\times A_{\sigma(p-1)}}(t_1,\cdots,t_{p-1}).$$

The integral is equal to zero except if $\sigma(p) = 1$, where it is equal to $\mu(A_1)$. By changing variables, we get :

$$\text{(A)} = \frac{\mu(A_1)}{p!}\sum_{\sigma\in\mathfrak{S}([\![2,p]\!])} \mathbf{1}_{A_{\sigma(2)}\times\cdots\times A_{\sigma(p)}}(t_1,\cdots,t_{p-1}).$$

We found the symmetrization of an indicator function :

$$\text{(A)} = \frac{\mu(A_1)}{p}\tilde{\mathbf{1}}_{A_2\times\cdots\times A_p}.$$

Hence, we can compute $I_{p-1}(f\otimes_1 g)$:

$$I_{p-1}(f\otimes_1 g) = \frac{\mu(A_1)}{p}W(A_2)\cdots W(A_p).$$

2. By the proposition **II.5**, the relation **(iii)** gives :

$$\mathbb{E}\left[|I_{p+1}(h_\varepsilon) - I_{p+1}(f\otimes g)|^2\right] = \left\|\tilde{h}_\varepsilon - \widetilde{f\otimes g}\right\|^2_{L^2(T^{m+1})}.$$

Let's call it (B). Then, by property of symmetrization :

$$(\mathsf{B}) = \left\|\tilde{h}_\varepsilon - \tilde{\mathbf{1}}_{A_1\times A_1\times A_2\times\cdots\times A_p}\right\|^2_{L^2(T^{m+1})}.$$

By triangular inequality, we remove the tildes :

$$(\mathsf{B}) \leqslant \left\|h_\varepsilon - \mathbf{1}_{A_1\times A_1\times A_2\times\cdots\times A_p}\right\|^2_{L^2(T^{m+1})}.$$

We compute the right hand side. We have :

$$\begin{aligned}(\mathsf{B}) \leqslant\ & \mu(A_2)\cdots\mu(A_p) \\ & \cdot\int_{T^2}\left[\sum_{i\neq j}\mathbf{1}_{B_i\times B_j}(t,s) - \mathbf{1}_{A_1\times A_1}(t,s)\right]^2 \mathrm{d}\mu(t)\mathrm{d}\mu(s).\end{aligned}$$

The sum can be easier expressed :

$$\begin{aligned}(\mathsf{B}) \leqslant\ & \mu(A_2)\cdots\mu(A_p) \\ & \cdot\int_{T^2}\left[\sum_{i=1}^n\mathbf{1}_{B_i\times B_i}(t,s)\right]^2 \mathrm{d}\mu(t)\mathrm{d}\mu(s).\end{aligned}$$

Since the sums is over indicators with disjoints support :

$$\begin{aligned}(\mathsf{B}) \leqslant\ & \mu(A_2)\cdots\mu(A_p) \\ & \cdot\int_{T^2}\left[\sum_{i=1}^n\mathbf{1}_{B_i}(t)\mathbf{1}_{B_i}(s)\right] \mathrm{d}\mu(t)\mathrm{d}\mu(s).\end{aligned}$$

We have finally :

$$(\mathsf{B}) \leqslant \left(\sum_{i=1}^n \mu(B_i)\right)^2 \mu(A_2)\cdots\mu(A_p).$$

In terms of ε, and using the fact that $(B_i)_i$ is partition of A_1, we have :

$$(\mathsf{B}) \leqslant \frac{\varepsilon\beta}{6}.$$

3. ▷ Let us give an expression on R_ε thanks to the two first points. To do this, we expend $I_p(f)I_1(g)$, and we artificially appear the term $I_{p+1}(h_\varepsilon)$. We have :

$$W(A_1)^2W(A_2)\cdots W(A_p).$$

With the partition $(B_i)_i$, we have by expanding the square of the sum :

$$\begin{aligned}I_p(f)I_1(g) =\ & \sum_{i\neq j} W(B_i)W(B_j)W(A_2)\ldots W(A_p) \\ & + \sum_{i=1}^n W(B_i)^2W(A_2)\cdots W(A_p).\end{aligned}$$

We have $I_{p+1}(h_\varepsilon)$ in the first term. We introduce $\mu(A_1)$ by introducing $\mu(B_i)$ in the second sum :

$$\begin{aligned}I_p(f)I_1(g) =\ & I_{p+1}(h_\varepsilon) \\ & + \sum_{i=1}^n \mu(B_i)W(A_2)\cdots W(A_p) \\ & + \sum_{i=1}^n \left(W(B_i)^2 - \mu(B_i)\right)W(A_2)\cdots W(A_p).\end{aligned}$$

Since $\sum_i \mu(B_i) = \mu(A_1)$, we get hat we computed in **1.** :

$$\begin{aligned}I_p(f)I_1(g) =\ & I_{p+1}(h_\varepsilon) + I_{p-1}(f\otimes_1 g) \\ & + \sum_{i=1}^n \left(W(B_i)^2 - \mu(B_i)\right)W(A_2)\cdots W(A_p).\end{aligned}$$

We call R_ε the last term :

$$R_\varepsilon \overset{\text{def.}}{=} \sum_{i=1}^n \left(W(B_i)^2 - \mu(B_i)\right)W(A_2)\cdots W(A_p).$$

▷ We have by independence of $(W(B_i))_i$:

$$\mathbb{E}\left[R_\varepsilon^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^n\left(W(B_i)^2 - \mu(B_i)\right)\right)^2\right]\mu(A_2)\cdots\mu(A_p).$$

Let's call the expectation term (C). We expend the square :

$$(\mathsf{C}) = \sum_{i,j=1}^n \mathbb{E}\left[\left(W(B_i)^2 - \mu(B_i)\right)\left(W(B_j)^2 - \mu(B_j)\right)\right].$$

Since $\mu(B_i)$ is the variance of $W(B_i)$, and by independence, the sum in j is null except if $j = i$:

$$(\mathsf{C}) = \sum_{i=1}^n \mathbb{E}\left[\left(W(B_i)^2 - \mu(B_i)\right)^2\right].$$

We just need to expend the square, and use the fact that $\mathbb{E}[N^4] = 3$ when $N\sim\mathcal{N}(0,1)$ to obtain that :

$$(\mathsf{C}) = 2\sum_{i=1}^n \mu(B_i)^2.$$

Finally, since $(B_1,\cdots,B_n)$ is a partition of A_1, we get :

$$\mathbb{E}\left[R_\varepsilon^2\right] \leqslant \frac{2\varepsilon\beta}{6}$$

4. We can finally conclude in our equality. Let us show that the square of the difference has a null expectation. Let us call

$$(\mathsf{D}) \overset{\text{def.}}{=} \mathbb{E}\left[|I_p(f)I_1(g) - I_{p+1}(f\otimes g) - pI_{p-1}(f\otimes_1 g)|^2\right].$$

By using the relation **3.**, and by convexity of the square, we have :

$$(\mathsf{D}) \leqslant 2\mathbb{E}\left[R_\varepsilon^2\right] + 2\mathbb{E}\left[I_{p+1}(h_\varepsilon - f\otimes g)^2\right].$$

By **2.** and **3.**, we get :

$$(\mathsf{D}) \leqslant \frac{2\varepsilon}{3} + \frac{2\varepsilon}{6} = \varepsilon.$$

We show this inequality for all $\varepsilon > 0$. This means we have finally the wanted equality. □

The main consequence is the following.

Proposition II.6 : Hermite polynomial and Itô-Wiener integral

Let $h \in L^2(T)$ with norm 1. Then,

$$\forall m \in \mathbb{N}^*, I_m\left(h^{\otimes m}\right) = H_m(X(h)).$$

Proof : We prove it by induction on m. For $m = 1$, this the corollary **II.3** we have proven. Let us suppose this property true for every integer lower or equal to m. We use the previous lemma.

$$I_{m+1}\left(h^{\otimes(m+1)}\right) = I_m(h^{\otimes m})I_1(h) - mI_{m-1}\left(h^{\otimes m} \otimes_1 h\right).$$

We have $I_1(h) = X(h)$, and $I_m(h) = H_m(X(h))$ by induction hypothesis. We compute $h^{\otimes m} \otimes_1 h$:

$$h^{\otimes m} \otimes_1 h(t_1, \cdots, t_{m-1}) = \int_T h(t_1) \cdots h(t_{m-1})h(u)^2 \, \mathrm{d}\mu(u).$$

Which gives :

$$h^{\otimes m} \otimes_1 h(t_1, \cdots, t_{m-1}) = h^{m-1}(t_1, \cdots, t_{m-1})\|h\|^2_{L^2(\mu)}.$$

And $\|h\|_{L^2(\mu)} = 1$. Finally, we have

$$I_{m+1}\left(h^{\otimes(m+1)}\right) = X(h)H_m(X(h)) - mH_{m-1}(X(h)).$$

By the relation **(i)** satisfied by the Hermite polynomials, the induction is complete. □

Corollary II.4 : Wiener chaos seen as an image of the Itô-Wiener integral

We note $L^2_{\mathsf{S}}(T^m)$ the set of symmetric functions of $L^2(T^m)$. Then, the map

$$I_m\Big|_{L^2_{\mathsf{S}}} : L^2_{\mathsf{S}}(T^m) \longrightarrow \mathfrak{H}_m$$

is an isomorphism.

Proof : By the previous proposition,

$$\{H_m(X(h)), \|h\| = 1\} \subset I_m(L^2_{\mathsf{S}}(T^m)).$$

Moreover, if $f \in L^2_{\mathsf{S}}(T^m)$, we have

$$\mathbb{E}\left[I_m(f)^2\right] = m!\|f\|^2_{L^2(\mu)}.$$

So $I_m(L^2_{\mathsf{S}}(T^m))$ is closed in $L^2(\mathbb{P})$, and by taking the closure, we have the inclusion :

$$\mathfrak{H}_m \subset I_m(L^2_{\mathsf{S}}(T^m)).$$

Moreover, this equality shows that I_m reduced to L^2_{S} is injective. To conclude, we just have to show the other inclusion. Let $h \in L^2(T)$ with $\|h\| = 1$. Then for all $n \neq m$ and $g \in L^2_{\mathsf{S}}(T^m)$, by **(iii)** in proposition **II.5** :

$$\langle H_n(X(h)), I_m(g)\rangle = \langle I_n(h^{\otimes n}), I_m(g)\rangle = 0.$$

Hence, for all $n \neq m$:

$$\{H_n(X(h)), \|h\| = 1\} \subset I_m(L^2_{\mathsf{S}}(T^m))^\perp$$

So, by taking the closure :

$$\mathfrak{H}_n \subset I_m(L^2_{\mathsf{S}}(T^m))^\perp$$

So :

$$I_m(L^2_{\mathsf{S}}(T^m)) \subset \overset{\perp}{\bigoplus_{n \neq m}} \mathfrak{H}_n = \mathfrak{H}_m.$$

This concludes the surjectivity of this map. □

This corollary concludes the proof of the theorem, since $L^2(\mathbb{P})$ is the direct sum of the Wiener chaos $\mathfrak{H}_m$.
Remark : By symmetry of the $(f_n)_n$, we have the following expression for $I_n(f_n)$:

$$I_n(f_n) = n! \int_0^{+\infty} \int_0^{t_n} \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n) \, \mathrm{d}W_{t_1} \cdots \mathrm{d}W_{t_n}.$$

II.5 Decomposition in Wiener chaos for variables with values in Hilbert space

We want to generalize the previous theorem, and we do it using it. We still set X an isonormal process on $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F}$ is the σ-algebra generated by X. We begin by defining what is a Wiener chaos for a random variable with values in $\mathcal{V}$.

Definition II.6

Let $\mathcal{V}$ a real separable Hilbert space. Then, we define the n-th *Wiener chaos* in $\mathcal{V}$ by :

$$\mathfrak{H}_n(\mathcal{V}) \overset{\text{def.}}{=} \overline{\mathrm{Vect}(Fv, F \in \mathfrak{H}_n, v \in \mathcal{V})}.$$

Then, we have the following theorem :

Theorem II.3 : Decomposition in Wiener chaos for a Hilbert

Let $\mathcal{V}$ a real-separable Hilbert space. Then :

$$L^2(\Omega \to \mathcal{V}) = \bigoplus_{n\in\mathbb{N}}^{\perp} \mathfrak{H}_n(\mathcal{V}).$$

for the norm :

$$\forall U \in L^2(\Omega \to \mathcal{V}), \|U\|^2_{L^2(\Omega\to\mathcal{V})} \overset{\text{def.}}{=} \mathbb{E}\left[\|U\|^2_{\mathcal{V}}\right]$$

Proof : Each space $\mathfrak{H}_n(\mathcal{V})$ are orthogonal since the space $\mathfrak{H}_n$ are too. We decompose $\mathcal{V}$ with a Hilbert basis :

$$\mathcal{V} = \bigoplus_{i\in\mathbb{N}}^{\perp} \mathbb{R}v_i.$$

Let $U \in L^2(\Omega \to \mathcal{V})$. We can decompose it in L^2 by :

$$U = \sum_{i\in\mathbb{N}} \langle U, v_i\rangle v_i.$$

But, since

$$\mathbb{E}\left[\|U\|^2_{\mathcal{V}}\right] = \sum_{i\in\mathbb{N}} \mathbb{E}\left[\langle U, v_i\rangle^2\right],$$

we have $\langle U, v_i\rangle \in L^2(\mathbb{P})$, and so it can be decomposed by Wiener-Itô decomposition we proves previously :

$$\langle U, v_i\rangle = \sum_{n\in\mathbb{N}} I_n(u_n^i).$$

As a consequence, by Fubini theorem, since both series converging absolutely in $L^2(\Omega \to \mathcal{V})$:

$$U = \sum_{n\in\mathbb{N}} \left(\sum_{i\in\mathbb{N}} I_n(u_n^i)v_i\right) \in \bigoplus_{n\in\mathbb{N}}^{\perp} \mathfrak{H}_n(\mathcal{V}).$$

This concludes the theorem. □

We still note J_n the projection of n-th Wiener chaos (we don't precise that J_n takes $\mathcal{V}$-valued random variables).

Set $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$, with μ a non-atomic measure. We can still construct Wiener integral on Hilbert space, and leading to this theorem. We introduce for that the tensor product between two Hilbert spaces.

Theorem II.4 : Isometry between Wiener chaos and symmetric functions

Let $\mathcal{V}$ a real separable Hilbert space, and $n \in \mathbb{N}^*$. There exists a map

$$I_n^{\mathcal{V}} : L^2_{\mathsf{S}}(T^n) \longrightarrow \mathfrak{H}_n$$

such that for all $m, q \in \mathbb{N}^*$, $f \in L^2_{\mathsf{S}}(T^m) \otimes \mathcal{V}$ and $g \in L^2_{\mathsf{S}}(T^q) \otimes \mathcal{V}$:

$$\mathbb{E}\left[\left\langle I_m^{\mathcal{V}}(f), I_q^{\mathcal{V}}(g)\right\rangle_{\mathcal{V}}\right] = m!\langle f, g\rangle_{L^2(T^m)\otimes\mathcal{V}}\delta_{m,q}.$$

Hence, for all all $U \in L^2(\Omega \to \mathcal{V})$, there exists an unique sequence $(u_n)_n \in L^2_{\mathsf{S}}(T^m) \otimes \mathcal{V}$ such that

$$J_n(U) = I_n^{\mathcal{V}}(u_n).$$

Proof : We define for $f \in L^2(T^m)$ and $v \in \mathcal{V}$:

$$I_m^{\mathcal{V}}(f \otimes v) \overset{\text{def.}}{=} I_m(f)v \in \mathfrak{H}_m(V).$$

We extend it by linearity on the span of $f \otimes v$, for $f \in L^2(T^m)$, $v \in \mathcal{V}$. Set $f \in L^2_{\mathsf{S}}(T^m)$, $g \in L^2_{\mathsf{S}}(T^q)$ symmetric, $v, w \in \mathcal{V}$. Then,

$$\mathbb{E}\left[\left\langle I_m^{\mathcal{V}}(f\otimes v), I_q^{\mathcal{V}}(g\otimes w)\right\rangle_{\mathcal{V}}\right] = \langle v, w\rangle_{\mathcal{V}}\mathbb{E}[I_m(f)I_q(g)].$$

By the isometry equality of the previous subsection, we have

$$\mathbb{E}\left[\left\langle I_m^{\mathcal{V}}(f\otimes v), I_q^{\mathcal{V}}(g\otimes w)\right\rangle_{\mathcal{V}}\right] = m!\langle v, w\rangle_{\mathcal{V}}\langle f, g\rangle_{L^2}\delta_{q,m}.$$

It yields to our result by the definition of the scalar product on $L^2(T) \otimes \mathcal{V}$. □

III The derivative operator

III.1 Definition in a general Hilbert space

Like in the one-dimensional case, we will introduce the derivative of a random variable in L^q, first on simple elements and then by density on more general random variables. The idea of this derivative is to differentiate with respect to ω, the random parameter.

In the following, we consider an isonormal Gaussian process $X = \{X(h), h \in \mathcal{H}\}$, on a real separable Hilbert space $\mathcal{H}$, defined on $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ est generated by X.

Definition III.1

Let F a real random variable. We say that F is *smooth* if there exists $n \in \mathbb{N}^*$ and :

(i) $f \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^n)$, that is :

$$\exists C \geqslant 0, \exists k \in \mathbb{N}, \forall l \in \mathbb{N}, \forall (i_1, \cdots, i_l) \in [\![1, n]\!]^l, \forall x \in \mathbb{R}^n, \left| \frac{\partial^l f}{\partial x_{i_1} \cdots \partial x_{i_l}}(x) \right| \leqslant C|x|^k;$$

(ii) $h_1, \cdots, h_n \in \mathcal{H}$;

such that

$$F = f(X(h_1), \cdots, X(h_n)).$$

We note $\mathcal{S}$ the set of all smooth random variables.

Lemma III.1 : Density of smooth functions

The set $\mathcal{S}$ is dense into $L^q(\mathbb{R})$ for the usual norm, for every finite $q \in [1, +\infty[$.

Proof of the lemma : This a consequence of the fact that $\mathcal{S}$ contains all the $H_p(X(h))$, with $h \in \mathcal{H}$ with norm 1. Since the set of those type of functions id dense by the lemma **II.3**, the set $\mathcal{S}$ is also dense in $L^q(\mathbb{R})$. □

Definition III.2

Let $F \in \mathcal{S}$, given by the definition **III.1**. We call the *Malliavin derivative of F*, denoted by $\mathrm{D}F$, the $\mathcal{H}$-valued random variable given by :

$$\mathrm{D}F \overset{\text{def.}}{=} \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))h_i.$$

In particular, for all $h \in \mathcal{H}$:

$$\mathrm{D}(X(h)) = h.$$

This operator depends on the choice of the orthonormal Gaussian process X.
Here's a first property which justifies the choice to define the Malliavin derivative.

Lemma III.2 : Anti derivative in $\mathcal{H}$

Let $F \in \mathcal{S}$. Then, for all $h \in \mathcal{H}$:

$$\mathbb{E}\left[\langle \mathrm{D}F, h\rangle_{\mathcal{H}}\right] = \mathbb{E}[FX(h)].$$

Proof of the lemma : The result is true for $h = 0$. Else, we can consider that $\|h\| = 1$. Since F is smooth, there exist $h_1, \cdots, h_m \in \mathcal{H}$ such that

$$F = f_1(W(h), W(h_1), \cdots, W(h_n))$$

with $f_1 \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^n)$. However, there exists an orthonormal basis $(e_1, \cdots, e_n)$ of $\mathrm{Vect}(h, h_1, \cdots, h_n)$ with $h = e_1$. By changing basis, and by linearity of W, we can suppose that

$$F = f(W(h), W(e_2), \cdots, W(e_n)).$$

Hence, we can conclude that

$$\langle \mathrm{D}F, h\rangle_{\mathcal{H}} = \frac{\partial f}{\partial x_1}(X(h), X(e_2), \cdots, X(h_n))$$

Computing the expectation of this is now an easy game. If we note for $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$:

$$\phi(x) \overset{\text{def.}}{=} \exp\left(\frac{-1}{2}\sum_{j=1}^n x_j^2\right)\frac{1}{\left(\sqrt{2\pi}\right)^n},$$

then,

$$\mathbb{E}\left[\langle \mathrm{D}F, h\rangle_{\mathcal{H}}\right] = \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_1}(x)\phi(x)\,\mathrm{d}x.$$

We integrate by parts with respect to x_1 :

$$\mathbb{E}\left[\langle \mathrm{D}F, h\rangle_{\mathcal{H}}\right] = \int_{\mathbb{R}^n} f(x)x_1\phi(x)\,\mathrm{d}x.$$

We finally get :

$$\mathbb{E}\left[\langle \mathrm{D}F, h\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[f(X(e_1), \cdots, X(e_n))X(e_1)\right].$$

In terms of h and F, it means that :

$$\mathbb{E}\left[\langle \mathrm{D}F, h\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[FX(h)\right].$$

That is the wanted formula. □

Thanks to this lemma, we can derive an integration by parts formula.

Lemma III.3 : Integration by parts 2

Let F, G smooth random variables, and $h \in \mathcal{H}$. Then,

$$\mathbb{E}\left[G\langle \mathrm{D}F, h\rangle_{\mathcal{H}}\right] = \mathbb{E}[FG \cdot X(h)] - \mathbb{E}\left[F\langle \mathrm{D}G, h\rangle_{\mathcal{H}}\right].$$

Here, what plays the role of the integration is the inner product with respect to h, and a anti derivative is given by the first lemma.

Proof of the lemma : We apply the first lemma to the product FG, where we can use the usual derivative rules for real functions :

$$\mathbb{E}\left[FG \cdot X(h)\right] = \mathbb{E}\left[\langle \mathrm{D}(FG), h\rangle_{\mathcal{H}}\right].$$

And we have by differentiating real functions :

$$\mathrm{D}(FG) = G\mathrm{D}F + F\mathrm{D}G.$$

□

We can as a consequence show the following property which generalize what we saw on first section.

Proposition III.1 : The derivative operator is closable

Let $q \geqslant 1$. Then,

$$\mathrm{D} : \left(\mathcal{S}, \mathbb{E}\left[|\cdot|^q\right]^{\frac{1}{q}}\right) \longrightarrow \left(L^q(\Omega \to \mathcal{H}), \mathbb{E}\left[\|\cdot\|_{\mathcal{H}}^q\right]^{\frac{1}{q}}\right)$$

is closable on $L^q(\mathbb{P})$. We still note D the extension of this map, and $\mathbb{D}^{q,1}$ the domain of the closed operator.

Proof : • First, let us see that it is indeed taking values in $L^p(\Omega \to \mathcal{H})$. Let $F \in \mathcal{S}$. Then, by convexity :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^p\right] \leqslant n^{p-1}\sum_{i=1}^n \mathbb{E}\left[\left|\frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))\right|^p\right]\|h_i\|_{\mathcal{H}}^p.$$

Since $f \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^n)$, we have :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^p\right] \leqslant n^{p-1}C^p\sum_{i=1}^n \mathbb{E}\left[|X(h_1)\cdots X(h_n)|^p\right]\|h_i\|_{\mathcal{H}}^p.$$

Which is finite, since $\{X(h), h \in \mathcal{H}\}$ is a Gaussian process.

• We use the sequence characterization. Let $(F_N)_N$ a sequence of $\mathcal{S}$ such that

$$\mathbb{E}[F_N^q] \xrightarrow[N\to+\infty]{} 0$$

and such that there exists $\eta \in L^q(\Omega \to \mathcal{H})$ such that

$$\mathbb{E}\left[\|\mathrm{D}F_N - \eta\|_{\mathcal{H}}^q\right] \xrightarrow[N\to+\infty]{} 0.$$

Let us show that $\eta = 0$. We do this by proving that $\eta \in \mathcal{H}^\perp$, and to do this, we show that :

$$\forall G \in \mathcal{S}, \forall h \in \mathcal{H}, \mathbb{E}\left[G\langle \eta, h\rangle_{\mathcal{H}}\right] = 0.$$

Let $G \in \mathcal{S}$:

$$G = g(X(h_1), \cdots, X(h_n));$$

and for all $\delta > 0$:

$$Z_\delta \overset{\text{def.}}{=} Ge^{-\delta X(h)^2}e^{-\delta\sum_{i=1}^n X(h_i)^2}.$$

Then $Z_\delta \in \mathcal{S}$ since by corollary **I.2** (Rodrigues formula) :

$$\forall k \in \mathbb{N}^*, \frac{\mathrm{d}^k}{\mathrm{d}x^k}\left[e^{\frac{-x^2}{2}}\right] = (-1)^p e^{\frac{-x^2}{2}}H_k(x).$$

We note k such that all the partial derivatives of g satisfy :

$$|\partial g(x_1, \cdots, x_n)| \leqslant C|x_1|^k \cdots |x_n|^k,$$

and

$$M_{\delta,k} \overset{\text{def.}}{=} \sup_{x\in\mathbb{R}}\left|x^k e^{-\delta x^2}\right|.$$

Hence, $\mathbb{P}$-almost surely :

$$|Z_\delta| \leqslant CM_{\delta,k}^n e^{-\delta X(h)^2}.$$

▷ We have for all $h \in \mathcal{H}$:

$$\mathbb{E}\left[Z_\delta\langle \eta, h\rangle_{\mathcal{H}}\right] \xrightarrow[\delta\to0^+]{} \mathbb{E}\left[G\langle \eta, h\rangle_{\mathcal{H}}\right].$$

Indeed, by using $1 - e^{-x} \leqslant x$ for $x \geqslant 0$, we have :

$$\mathbb{E}\left[|Z_\delta - G|^2\right] \leqslant \delta\mathbb{E}\left[G^2\right]^{\frac{1}{2}} \cdot \mathbb{E}\left[\left(X(h)^2 + \sum_{k=1}^n X(h_k)^2\right)^2\right]^{\frac{1}{2}},$$

which goes to zero when $[\delta \to 0^+]$. By continuity of the inner product on L^2, we conclude on the expected limit. As a consequence, we just need to show that for all $\delta > 0$:

$$\mathbb{E}\left[Z_\delta\langle \eta, h\rangle_{\mathcal{H}}\right] = 0.$$

▷ We also have :

$$\mathbb{E}\left[Z_\delta\langle \mathrm{D}F_N, h\rangle_{\mathcal{H}}\right] \xrightarrow[N\to+\infty]{} \mathbb{E}\left[Z_\delta\langle \eta, h\rangle_{\mathcal{H}}\right].$$

But, by integration by parts formula :

$$\mathbb{E}\left[Z_\delta \langle \mathrm{D}F_N, h\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[F_N Z_\delta\ X(h)\right] - \mathbb{E}\left[F_n \langle \mathrm{D}Z_\delta, h\rangle_{\mathcal{H}}\right].$$

We just need to show that both terms go to zero when $[N \to +\infty]$.

▷ For the first term, we have by definition of M_δ :

$$|\mathbb{E}\left[F_N Z_\delta\ X(h)\right]| \leqslant C M_{\delta,k}^{n+1} \mathbb{E}[|F_N|].$$

By Hölder inequality, this term goes to 0 when $[N \to +\infty]$.

▷ For the second term, we need to find a deterministic bound for $\|\mathrm{D}Z_\delta\|$ to use the same argument as the first term. The (awful) computation gives :

$$\begin{aligned} & \mathrm{D}Z_\delta \\ = \quad & \sum_{j=1}^n \frac{\partial g}{\partial x_j}(X(h_1), \cdots, X(h_n))h_j \\ & \cdot e^{-\delta\left(X(h)^2 + \sum_{i=1}^n X(h_i)^2\right)} \\ - \quad & \delta \sum_{j=1}^n X(h_j) Z_\delta h_j \\ - \quad & \delta X(h) Z_\delta h. \end{aligned}$$

We can then estimate its norm :

$$\begin{aligned} & \|\mathrm{D}Z_\delta\|_{\mathcal{H}} \\ \leqslant \quad & C M_{\delta,k}^n e^{-\delta X(h)^2} \sum_{j=1}^n \|h_j\|_{\mathcal{H}} \\ + \quad & C\delta M_{\delta,k}^n M_{\delta,k+1} \left(\|h\|_{\mathcal{H}} + \sum_{j=1}^n \|h_j\|_{\mathcal{H}} \right). \end{aligned}$$

Here is our deterministic bound. We can consequently estimate the second term of the integration by parts, using Cauchy-Schwarz inequality :

$$\begin{aligned} & |\mathbb{E}\left[F_n \langle \mathrm{D}Z_\delta, h\rangle_{\mathcal{H}}\right]| \\ \leqslant \quad & \mathbb{E}[|F_N|] \cdot C\|h\|_{\mathcal{H}} M_{\delta,k}^n e^{-\delta X(h)^2} \sum_{j=1}^n \|h_j\|_{\mathcal{H}} \\ + \quad & \mathbb{E}[|F_N|] \cdot C\delta M_{\delta,k}^n M_{\delta,k+1} \left(\|h\|_{\mathcal{H}}^2 + \sum_{j=1}^n \|h_j\|_{\mathcal{H}} \|h\|_{\mathcal{H}} \right). \end{aligned}$$

It follows that the second term also goes to zero.

▷ So, we have

$$\mathbb{E}\left[Z_\delta \langle \mathrm{D}F_N, h\rangle_{\mathcal{H}}\right] \xrightarrow[N \to +\infty]{} 0.$$

Meaning that for every $h \in \mathcal{H}$ and $\delta > 0$:

$$\mathbb{E}\left[Z_\delta \langle \eta, h\rangle_{\mathcal{H}}\right] = 0.$$

By letting $[\delta \to 0^+]$, we get, for all $G \in \mathcal{S}$ and $h \in \mathcal{H}$:

$$\mathbb{E}\left[G \langle \eta, h\rangle_{\mathcal{H}}\right] = 0.$$

It means that $\langle \eta, h\rangle_{\mathcal{H}} \in \mathcal{S}^\perp$ in $L^2(\mathbb{P})$. Since $\mathcal{S}$ is dense in $L^2(\mathbb{P})$, we conclude that for every $h \in \mathcal{H}$, $\mathbb{P}$-almost surely :

$$\langle \eta, h\rangle_{\mathcal{H}} = 0.$$

Hence, $\eta \in \mathcal{H}^\perp$ $\mathbb{P}$-almost surely, so $\eta = 0$ in L^q. We have shown that the operator D is closable. □

Recall that in this case, $\mathbb{D}^{1,2}$ is the closure of $\mathcal{S}$ for

$$\forall F \in \mathcal{S}, \|F\|_{1,2}^2 \overset{\text{def.}}{=} \mathbb{E}\left[|F|^2\right] + \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right].$$

We have the following characterization : $F \in \mathbb{D}^{1,2}$ if and only if there exists a sequence $(F_n)_n$ of elements of $\mathbb{D}^{1,2}$ such that $(F_n)_n$ converges in $L^2(\mathbb{P})$ to F and such that $(\mathrm{D}F_n)_n$ is a Cauchy sequence for $\|\cdot\|_{\mathbb{D}^{1,2}}$. It is equivalent to find $F_n \in \mathcal{S}$.

Theorem III.1 : Domain of the Malliavin derivative and chaos expansion

Let $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then, the following are equivalents :

(i) $F \in \mathbb{D}^{1,2}$;

(ii) The series $\sum_n n\mathbb{E}\left[J_n F^2\right]$ converges.

In this case, we have

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = \sum_{n=0}^{+\infty} n\mathbb{E}\left[J_n F^2\right],$$

and for all $n \geqslant 1$, we have $\mathrm{D}(J_n F) = J_{n-1}(\mathrm{D}F)$.

Proof : **[$\Longrightarrow$]** All the game here is to compute $\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right]$ for every $F \in \mathbb{D}^{1,2}$.

• We begin by the simplest case : for $F = H_n(X(h))$. Then, $F \in \mathcal{S}$ so belongs to the domain of the Malliavin derivative. Its derivative is :

$$\mathrm{D}(H_n(X(h)) = H_n'(X(h))h.$$

By the recurrent relation satisfied by $(H_n)_n$, we have :

$$\mathrm{D}(H_n(X(h)) = nH_{n-1}(X(h))h.$$

This also writes

$$\mathrm{D}(J_n(H_n(X(h))) = J_{n-1}\mathrm{D}(H_n(X(h)),$$

which is the relation we want to prove. Finally, if we take norm on both sizes :

$$\mathbb{E}\left[\|\mathrm{D}(H_n(X(h))\|_{\mathcal{H}}^2\right] = n^2\|h\|_{\mathcal{H}}^2 \mathbb{E}\left[H_{n-1}(X(h))^2\right].$$

By the lemma **II.2**, we have an expression of the covariance of Hermite polynomials :

$$\mathbb{E}\left[\|\mathrm{D}(H_n(X(h))\|_{\mathcal{H}}^2\right] = n^2(n-1)!\|h\|_{\mathcal{H}}^2 \|h\|_{\mathcal{H}}^{2(n-1)}.$$

And once again by this lemma, by giving a n from n^2, we get :

$$\mathbb{E}\left[\|\mathrm{D}(H_n(X(h))\|_{\mathcal{H}}^2\right] = n\mathbb{E}\left[H_n(X(h))^2\right].$$

This is the equality we expected for $F = H_n(X(h))$.

• We extend it to all finite linear combination of $H_n(X(h))$ (which is still smooth so still in $\mathbb{D}^{1,2}$). Let $F \in \mathrm{Vect}(H_n(X(h)), h \in \mathcal{H})$, with $\|h\|_{\mathcal{H}} = 1$:

$$F = \sum_{k=1}^{K} \alpha_k H_n(X(h_k)).$$

Then,

$$\mathrm{D}F = n\sum_{k=1}^{K} \alpha_k H_{n-1}(X(h_k))h_k.$$

We expend the expectation of the squared norm of it.

$$\begin{aligned}&\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right]\\ =\;& n^2\sum_{k,l=1}^{K} \alpha_k\alpha_l\mathbb{E}\left[H_{n-1}(X(h_k))H_{n-1}(X(h_l))\right]\langle h_k, h_l\rangle.\end{aligned}$$

But, we know by lemma **II.2** that if (X,Y) is a centered Gaussian couple with $\mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1$, then

$$\mathbb{E}[H_n(X)H_n(Y)] = n!\mathbb{E}[XY]^n.$$

It yields here to, by definition of X being an isonormal Gaussian process :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = n^2(n-1)!\sum_{k,l=1}^{K} \alpha_k\alpha_l\langle h_k, h_l\rangle^n.$$

By giving a "n" from n^2, it writes :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = n\sum_{k,l=1}^{K} n!\alpha_k\alpha_l\langle h_k, h_l\rangle^n.$$

And using once again the equality of the covariance of Hermite polynomials, we get :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = n\sum_{k,l=1}^{K} \alpha_k\alpha_l\mathbb{E}\left[H_n(X(h_k))H_n(X(h_l))\right].$$

By linearity, we have :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = n\mathbb{E}\left[\left(\sum_{k=1}^{K} \alpha_k H_n(X(h_k))\right)^2\right].$$

Meaning :

$$\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = n\mathbb{E}[F^2].$$

The first expression of the statement is true for every $F \in \mathrm{Vect}(H_n(X(h)), h \in \mathcal{H})$. Same for the second equality between D and J_{n-1}, which is just coming from the linearity of the two operators.

• Set $n \in \mathbb{N}$. Let us prove that we can extend it on $\mathfrak{H}_n$. First, notice that we can extend it on $\mathfrak{H}_n \cap \mathbb{D}^{1,2}$ by continuity for the norm

$$\|F\|_{\mathbb{D}^{1,2}}^2 = \mathbb{E}[F^2] + \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right]$$

of D and J_n on $\mathfrak{H}_n \cap \mathbb{D}^{1,2}$. Same for the second equality. Second, let us show that in fact, $\mathfrak{H}_n \subset \mathbb{D}^{1,2}$. Indeed, if $G \in \mathfrak{H}_n$, then there exists a sequence $(\psi_p)_p$ of elements of $\mathrm{Vect}(H_n(X(h))$ such that

$$\psi_p \xrightarrow[p\to+\infty]{L^2} F.$$

Since $\psi_p \in \mathcal{S}$, $\psi_p \in \mathbb{D}^{1,2}$. Moreover, $(\mathrm{D}\psi_p)_p$ is a Cauchy sequence, since by what we proves previously ($\psi_p \in \mathfrak{H}_n$) :

$$\mathbb{E}\left[\|\mathrm{D}\left(\psi_{p+q} - \psi_p\right)\|_{\mathcal{H}}^2\right] = n\mathbb{E}\left[\left(\psi_{p+q} - \psi_p\right)^2\right].$$

So, it means that $G \in \mathbb{D}^{1,2}$, and so we proved that

$$\mathfrak{H}_n \subset \mathbb{D}^{1,2}.$$

Let us notice finally that D sends $\mathfrak{H}_n$ into $\mathfrak{H}_{n-1}(\mathcal{H})$.

• We finally extend it on $\mathbb{D}^{1,2}$. Let us prove that once again, the continuity for the norm $\mathbb{D}^{1,2}$ will allow us to conclude. Let $F \in \mathbb{D}^{1,2}$. Then, by the previous point, since $\mathfrak{H}_n \subset \mathbb{D}^{1,2}$, we have $J_nF \in \mathbb{D}^{1,2}$. For all $N \in \mathbb{N}$, we have by orthogonality of the spaces $\mathfrak{H}_n(\mathcal{H})$ and by the previous points :

$$\mathbb{E}\left[\left\|\mathrm{D}\left(\sum_{n=0}^{N} J_nF\right)\right\|_{\mathcal{H}}^2\right] \overset{(*)}{=} \sum_{n=0}^{N} n\mathbb{E}\left[J_nF^2\right].$$

But, by decomposition theorem on general Hilbert spaces, we have in $L^2(\Omega \to \mathcal{H})$:

$$\mathrm{D}F = \sum_{n=0}^{+\infty} J_n(\mathrm{D}F).$$

So, by the previous point, and since the derivative of a constant is zero :

$$\mathrm{D}F = \sum_{n=0}^{+\infty} \mathrm{D}(J_{n+1}F) = \sum_{n=0}^{+\infty} \mathrm{D}(J_nF).$$

Hence, it means that in $L^2(\Omega \to \mathcal{H})$, we have :

$$\sum_{n=0}^{N} \mathrm{D}(J_nF) \xrightarrow[N\to+\infty]{L^2(\Omega\to\mathcal{H})} \mathrm{D}F.$$

Consequently, the right hand side of $(*)$ converges, so the left hand side too. This proves **(ii)**. Moreover, we can take the limit as $[N \to +\infty]$ to conclude to the equality for every $F \in \mathbb{D}^{1,2}$. Once again, the second equality is true by continuity.

[$\Longleftarrow$] Let $F \in L^2(\mathbb{P})$. We suppose that the series $\sum_n n\mathbb{E}\left[J_nF^2\right]$ is convergent. We will prove that $F \in \mathbb{D}^{1,2}$ by approaching it by its truncated sums in Wiener decomposition. Here's the plan.

- First, since we know that $\sum_{n=0}^{N} J_nF$ converges to F in L^2, let us prove that $J_nF \in \mathbb{D}^{1,2}$ by approaching it by Hermite polynomials.
- Then, we have to show that the convergence is not only in L^2 but in $\mathbb{D}^{1,2}$.

Let us prove it.

• We set $n \in \mathbb{N}$. By definition of belonging to $\mathfrak{H}_n$, there exists a sequence $(\psi_p)_p$ of elements of $\mathrm{Vect}(H_n(X(h))$ such that

$$\psi_p \xrightarrow[p\to+\infty]{L^2} J_nF.$$

We have $\psi_p \in \mathcal{S}$, since H_n is a polynomial, hence $\psi_p \in \mathbb{D}^{1,2}$. To show that $J_nF \in \mathbb{D}^{1,2}$, it enough to show that $(\mathrm{D}\psi_p)_p$ a Cauchy sequence in $L^2(\Omega \to \mathcal{H})$. But, by the computations made in **(i)**, we have in fact, for all $q \in \mathbb{N}$:

$$\mathbb{E}\left[\|\mathrm{D}\left(\psi_{p+q} - \psi_p\right)\|_{\mathcal{H}}^2\right] = n\mathbb{E}\left[|\psi_{p+q} - \psi_p|^2\right].$$

By L^2-convergence of (ψ_p), this sequence is also a Cauchy sequence, the right hand side is as small as we want, and so $(\mathrm{D}\psi_p)_p$ is Cauchy. This concludes that $J_nF \in \mathbb{D}^{1,2}$.

• Now, n is not set, and we set :

$$\phi_n \overset{\text{def.}}{=} \sum_{k=0}^{n} J_kF.$$

By the previous point, $\phi_n \in \mathbb{D}^{1,2}$. By the Wiener-Itô decomposition, $(\phi_n)_n$ converges in L^2 to F. We just have to show that, one more time, $(\mathrm{D}\phi_n)_n$ is a Cauchy sequence on $L^2(\Omega \to \mathcal{H})$. Using once again **(i)** for ϕ_n, we have, for all $p \in \mathbb{N}$:

$$\mathbb{E}\left[\|\mathrm{D}\left(\phi_{n+p} - \phi_n\right)\|_{\mathcal{H}}^2\right] = \sum_{k=0}^{+\infty} k\mathbb{E}\left[J_k\left(\phi_{n+p} - \phi_n\right)^2\right].$$

But, we have :

$$\phi_{n+p} - \phi_n = \sum_{k=n+1}^{n+p} J_k F,$$

so we can easily compute the projections of it :

$$\mathbb{E}\left[\|\mathrm{D}\left(\phi_{n+p} - \phi_n\right)\|_{\mathcal{H}}^2\right] = \sum_{k=n+1}^{n+p} k\mathbb{E}\left[J_k F^2\right].$$

We can bound it :

$$\mathbb{E}\left[\|\mathrm{D}\left(\phi_{n+p} - \phi_n\right)\|_{\mathcal{H}}^2\right] \leqslant \sum_{k=n+1}^{+\infty} k\mathbb{E}\left[J_k F^2\right].$$

Since the series converges by hypothesis, we conclude that $(\mathrm{D}\phi_p)_p$ is indeed a Cauchy sequence, and so that $F \in \mathbb{D}^{1,2}$. We proved the characterization. □

Note : For all $F \in L^2$, $J_n F \in \mathbb{D}^{1,2}$. Moreover, the operator D is an isometry on $\mathfrak{H}_n$.

Corollary III.1 : When the Malliavin derivative is zero

Let $F \in \mathbb{D}^{1,2}$ such that $\mathrm{D}F = 0$ $\mathbb{P}$-almost surely. Then $\mathbb{P}$-almost surely :

$$F = \mathbb{E}[F].$$

III.2 Chain rule and consequences

Before proving it, we use a lemma which helps to know if a limit of elements of $\mathbb{D}^{1,2}$ still belongs to $\mathbb{D}^{1,2}$.

Lemma III.4 : A sufficient condition to be in the domain of the Malliavin derivative

Let $(F_n)_n$ a sequence of $\mathbb{D}^{1,2}$ and let $F \in L^2(\mathbb{P})$. If we suppose that :

(i) The sequence $(F_n)_n$ converges to F in $L^2(\mathbb{P})$;

(ii) The sequence of its derivatives is bounded :

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[\|\mathrm{D}F_n\|_{\mathcal{H}}^2\right] < +\infty.$$

Then $F \in \mathbb{D}^{1,2}$, and for all $U \in L^2(\Omega \to \mathcal{H})$:

$$\mathbb{E}\left[\langle \mathrm{D}F_n, U\rangle_{\mathcal{H}}\right] \xrightarrow[n\to+\infty]{} \mathbb{E}\left[\langle \mathrm{D}F, U\rangle_{\mathcal{H}}\right]$$

We will admit it.

Proposition III.2 : Chain rules

1. Let $\varphi \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R})$ (φ and φ are bounded), and $F \in \mathbb{D}^{1,2}$. Then $\varphi \in \mathbb{D}^{1,2}$ with

$$\mathrm{D}\varphi(F) = \varphi'(F)\ \mathrm{D}F.$$

2. Let $\varphi \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^m, \mathbb{R})$, $p \in \mathbb{N}^*$ and $F = (F^1, \cdots, F^m)$ a real random vector such that every $F^k \in \mathbb{D}^{1,2}$. Then $\varphi(F) \in \mathbb{D}^{1,2}$ and :

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

Proof : The demonstration is the same for the two points, it is just technical details form differential calculus that makes the second point different from the first. We just prove the first point.

• If $\varphi \in \mathcal{C}_{\mathrm{b}}^\infty(\mathbb{R}) \subset \mathcal{C}_{\mathrm{pol}}^\infty(\mathbb{R})$ and $F \in \mathcal{S}$, then $\varphi(F) \in \mathcal{S}$ too and the formula is true :

$$\mathrm{D}\varphi(F) = \varphi'(F)\mathrm{D}F.$$

It remains true for $F \in \mathbb{D}^{1,2}$, since if $(F_n)_n \in \mathcal{S}^{\mathbb{N}}$ converges in $\mathbb{D}^{1,2}$ to F then we have

$$\left|\varphi'(F_n)\right| \leqslant \left\|\varphi'\right\|_\infty,$$

so by dominated convergence, we have

$$\varphi'(F_n) \xrightarrow[n\to+\infty]{L^2} \varphi'(F),$$

so

$$\mathrm{D}\varphi(F_n) \xrightarrow[n\to+\infty]{L^2(\Omega\to\mathcal{H})} \varphi'(F)\mathrm{D}F.$$

Also by dominated convergence, $(\varphi(F_n))_n$ converges to $\varphi(F)$ in L^2. By the lemma, it means that $\varphi(F) \in \mathbb{D}^{1,2}$, and that we have the weak convergence of $(\mathrm{D}\varphi(F_n))_n$. Since this sequence actually converges in $L^2(\Omega \to \mathcal{H})$, it means that the formula remains true for every $F \in \mathbb{D}^{1,2}$.

- To conclude, we approach $\varphi \in \mathcal{C}^1_b(\mathbb{R})$ with the help of an approximation of unity $(\rho_\varepsilon)_{\varepsilon>0}$, that is $\rho_\varepsilon \in \mathcal{C}^\infty$ satisfies that $\rho_\varepsilon \geqslant 0$, $\int_{\mathbb{R}} \rho_\varepsilon = 1$ and $\mathrm{Supp}(\rho_\varepsilon) \subset [-\varepsilon, \varepsilon]$. We define

$$\varphi_\varepsilon \overset{\text{def.}}{=} \varphi * \rho_\varepsilon.$$

Then φ_ε converges uniformly to φ on $\mathbb{R}$. By the precedent point, for all $\varepsilon > 0$, we have, for $F \in \mathbb{D}^{1,2}$:

$$\mathrm{D}\varphi_\varepsilon(F) = \varphi_\varepsilon'(F)\mathrm{D}F.$$

But, by dominated convergence, we have

$$\varphi_\varepsilon'(F)\mathrm{D}F \xrightarrow[\varepsilon\to 0]{L^2(\Omega\to\mathcal{H})} \varphi'(F)\mathrm{D}F.$$

Meaning that $(\mathrm{D}\varphi_\varepsilon(F))_{\varepsilon>0}$ converges on $L^2(\Omega \to \mathcal{H})$. By the lemma, it means that $\varphi(F) \in \mathbb{D}^{1,2}$, and the formula, as the same argument as precedent point. □

We can extend this result by the same argument on Lipschitz functions.

Proposition III.3 : Chain rule for Lipschitz functions

Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ which is K-Lipschitz. Let $F \in \mathbb{D}^{1,2}$. Then $\varphi(F) \in \mathbb{D}^{1,2}$. Moreover, there exists $G \in L^2$ such that $|G| \leqslant K$ $\mathbb{P}$-almost surely and

$$\mathrm{D}\varphi(F) = G\ \mathrm{D}F.$$

Proof : Same argument as previously. We define once again $\varphi_\varepsilon = \varphi * \rho_\varepsilon$, with ρ_ε being an approximation of unity. Then by precedent proposition :

$$\mathrm{D}\varphi_\varepsilon = \varphi_\varepsilon'(F)\ \mathrm{D}F.$$

By dominated convergence,

$$\varphi_\varepsilon(F) \xrightarrow[\varepsilon\to 0]{L^2} \varphi_\varepsilon(F),$$

and we have by the fact that φ is K-Lipschitz :

$$\sup_{\varepsilon>0} \mathbb{E}\left[\|\mathrm{D}\varphi_\varepsilon(F)\|_{\mathcal{H}}\right] \leqslant K\mathbb{E}\left[\|\mathrm{D}F\|\right] < +\infty.$$

Hence, by the lemma $\varphi(F) \in \mathbb{D}^{1,2}$, and $(\mathrm{D}\varphi_\varepsilon(F))_{\varepsilon>0}$ weakly converges to $\mathrm{D}\varphi(F)$. Finally, since $(\varphi_\varepsilon'(F))_{\varepsilon>0}$ is bounded by K in L^2, there exists a subsequence of it that weakly converges to a certain $G \in L^2$ which is bounded by K. By dominated convergence, $(\mathrm{D}\varphi_\varepsilon(F))_{\varepsilon>0}$ weakly converges, up to the extraction, to $G\mathrm{D}F$. By unicity of the weak limit, it means that

$$\mathrm{D}\varphi(F) = G\mathrm{D}F.$$

Which is the expected formula. □

We can imagine a similar result for $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R}$ which is Lipschitz. Those chain rules remains true in $\mathbb{D}^{1,p}$, for every $p \geqslant 1$.

Corollary III.2 : Indicator belonging or not in the Malliavin domain

Let $A \in \mathcal{F}$. Then the following assertions are equivalent :

(i) The indicator function of A belongs to $\mathbb{D}^{1,2}$: $\mathbf{1}_A \in \mathbb{D}^{1,2}$;

(ii) We have $\mathbb{P}(A) \in \{0,1\}$.

In particular, for all $h \in \mathcal{H}$, $\mathbf{1}_{\{X(h)>0\}}$ does not belong to $\mathbb{D}^{1,2}$.

Proof : **[$\Longrightarrow$]** If $F \overset{\text{def.}}{=} \mathbf{1}_A \in \mathbb{D}^{1,2}$, then since $F^2 = F$, we apply the chain rule lemma for $\varphi \in \mathcal{C}^\infty_0(\mathbb{R})$ such that $\varphi(x) = x^2$ on $[0,1]$:

$$\mathrm{D}F = 2F\mathrm{D}F$$

If we suppose that $\mathrm{D}F$ is not equal to zero, then, we would have $F = \frac{1}{2}$, which is impossible. Then, we have $\mathrm{D}F = 0$, meaning by corollary **III.1** that $F = \mathbb{E}[F]$ almost surely, meaning here that

$$\mathbb{P}(A) = \mathbf{1}_A \in \{0,1\}.$$

[$\Longleftarrow$] Conversely, if $\mathbb{P}(A) \in \{0,1\}$, then $\mathbf{1}_A$ is almost surely constant, and so belongs to $\mathbb{D}^{1,2}$.

- Since

$$\mathbb{P}(X(h) > 0) = \frac{1}{2},$$

$\mathbf{1}_{\{X(h)>0\}}$ does not belong to $\mathbb{D}^{1,2}$. □

III.3 White noise

For the rest of this section, $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$, like in previous section. We set an isonormal process X, and a white noise W associated.

Definition III.3

Let $F \in \mathbb{D}^{1,2}$. We note $D_t F \in \mathbb{R}$ the *Malliavin derivative* taken in t of F :

$$\mathrm{D} : \left(\begin{array}{ccc} \mathbb{D}^{1,2} & \longrightarrow & L^2(T) \\ F & \longmapsto & \left(\begin{array}{ccc} T & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \mathrm{D}_t F \end{array} \right) \end{array} \right).$$

That is, if $F = f(X(h_1), \cdots, X(h_n)) \in \mathcal{S}$:

$$\forall t \in T, \mathrm{D}_t F \overset{\text{def.}}{=} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))h_i(t)$$

We can define the multi derivative of a random variable easily here. The general case use tensor products.

Definition III.4

Let $k \in \mathbb{N}^*$ and $s, t \in T$. We define for $F = f(X(h_1), \cdots, X(h_n)) \in \mathcal{S}$:

$$\mathrm{D}^2_{s,t} F \overset{\text{def.}}{=} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X(h_1), \cdots, X(h_n))h_i(t)h_j(s).$$

Remark : In the general case, $\mathrm{D}^2 : \mathcal{S} \longrightarrow \mathcal{H} \otimes \mathcal{H}$ is given by

$$\mathrm{D}^2 F = \sum_{i,j} \partial^2_{i,j} f(X(h_1), \cdots, X(h_n))h_i \otimes h_j.$$

Remark : We can define $\mathrm{D}^k_{t_1,\cdots,t_k}$ by this way. We can show by the exact same way than the general case that this operator is closed. We note $\mathbb{D}^{k,2}$ the domain of its closure, which is the closure of $\mathcal{S}$ for the norm

$$\|\cdot\|_{k,2} \overset{\text{def.}}{=} \left(\mathbb{E}\left[|\cdot|^2 \right] + \sum_{j=1}^{k} \int_{T^j} \mathbb{E}\left[\left| \mathrm{D}^j_{t_1,\cdots,t_k} \cdot \right|^2 \right] \, \mathrm{d}\mu(t_1) \cdots \mathrm{d}\mu(t_k) \right)^{\frac{1}{2}}.$$

In the case of white nose, it is easy to compute the derivative of a L^2 random variable.

Theorem III.2 : Derivative in white noise case

Let $F \in L^2(\mathbb{P})$ having its Wiener chaos decomposition given by $F = \sum_{n=0}^{+\infty} I_n(f_n)$, with $f_n \in L^2(T^n)$ symmetric. Then, we have

$$\mathbb{D}^{1,2} = \left\{ F \in L^2(\mathbb{P}), \sum_{n=1}^{+\infty} n^2 \cdot (n-1)! \, \|f_n\|^2_{L^2(T^n)} < +\infty \right\}.$$

In this case, if $F \in \mathbb{D}^{1,2}$, we have for almost every $t \in T$, in L^2 :

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

More generally, for every $k \in \mathbb{N}^*$, $F \in \mathbb{D}^{k,2}$ if and only if the series $\sum_{n=k}^{+\infty} n^k \cdot (n-k)! \, \|f_n\|^2_{L^2(T^n)}$ converges, and we have in this case, for almost every $t_1, \cdots, t_k \in T$:

$$\mathrm{D}^k_{t_1,\cdots,t_k} F = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} I_{n-k}\left(f_n(\cdot, t_1, \cdots, t_k)\right).$$

Proof : The equality is a consequence of the theorem **III.1** where we express $\mathbb{D}^{1,2}$ in terms of a series. And since for $F \in \mathbb{D}^{1,2}$, we have

$$\mathbb{E}\left[J_n F^2\right] = \mathbb{E}\left[I_n(f_n)^2\right] = n!\|f_n\|^2_{L^2(T^n)},$$

we can conclude to the equality of the two sets.

• Let us prove first the equality for every $F \in \mathcal{S}$. By linearity, it is enough to show this for $F = I_n(f_n)$, where $f_n \in L^2(T^n)$ is symmetric and elementary :

$$f_n = \sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \mathbf{1}_{A_{i_1}} \cdots \mathbf{1}_{A_{i_n}}.$$

By definition, its integral is given by :

$$I_n(f_n) = \sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \prod_{k=1}^n W(A_{i_k}).$$

Since $W(A) = X(\mathbf{1}_A)$, this means that we can compute its derivative :

$$\mathrm{D}_t F = \sum_{j=1}^n \sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \prod_{\substack{k=1\\k\neq j}}^n W(A_{i_k})\ \mathbf{1}_{A_j}(t).$$

But, if we compute the right hand side of the expected equality :

$$I_{n-1}(f_n(\cdot,t)) = \sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \prod_{\substack{k=1\\k\neq n}}^n W(A_{i_k})\ \mathbf{1}_{A_n}(t).$$

By symmetry, we have in fact for every $j \in [\![1,n]\!]$:

$$I_{n-1}(f_n(\cdot,t)) = \sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \prod_{\substack{k=1\\k\neq j}}^n W(A_{i_k})\ \mathbf{1}_{A_j}(t).$$

Hence, we have for the derivative of F :

$$\mathrm{D}_t F = n I_{n-1}(f_n(\cdot,t)).$$

By continuity of D on $\mathbb{D}^{1,2}$, and by the one of I_n in $L^2(T^n)$ (this is an isometry), we conclude that this expression is true for every $F \in \mathfrak{H}_n \cap \mathbb{D}^{1,2} = \mathfrak{H}_n$. (We proved the equality of the sets in theorem **III.1**).

• The case of higher order derivative can be done by induction, by the same arguments as previously. □

Corollary III.3 : An expression for f_n in terms of Malliavin derivative

Let $F \in \mathbb{D}^{\infty,2} \stackrel{\text{def.}}{=} \bigcap_{p\in\mathbb{N}^*} \mathbb{D}^{p,2}$ having its Wiener decomposition given by $F = \sum_{n=0}^{+\infty} I_n(f_n)$, with f_n symmetric. Then, for every $n \in \mathbb{N}^*$, almost everywhere in T^n :

$$f_n = \frac{1}{n!}\mathbb{E}\left[\mathrm{D}^n_\bullet F\right].$$

Proof : We do it for elementary functions and we conclude by density. Let $k \in \mathbb{N}^*$ and $t_1,\cdots,t_k \in T$. Then, for all $n \geqslant k$:

$$\begin{aligned}&\mathbb{E}\left[I_{n-k}(f_n(\cdot,t_1,\cdots,t_k))\right]\\ = &\sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \prod_{j=n-k+1}^n \mathbf{1}_{A_{i_j}}\\ &\cdot\ \mathbb{E}\left[\prod_{j=1}^{n-k} W(A_{i_j})\right].\end{aligned}$$

This expectation is zero except if $n = k$, where we have :

$$\mathbb{E}\left[I_0(f_n(t_1,\cdots,t_n))\right] = \sum_{i_1,\cdots,i_n} a_{i_1,\cdots,i_n} \prod_{j=1}^n \mathbf{1}_{A_{i_j}}.$$

That is f_n. We get by consequently, by the expansion proved in the previous proposition :

$$\mathbb{E}\left[\mathrm{D}^n_{t_1,\cdots,t_n} F\right] = n! f_n.$$

We get the expected relation. □

Let us prove the following property, showing a way to see the relevance of Malliavin derivative. If $A \in \mathcal{B}$, we note

$$\mathcal{F}_A \stackrel{\text{def.}}{=} \sigma\left(W(B), B \in \mathcal{B}, B \subset A, \mu(B) < +\infty\right).$$

Proposition III.4 : Measurability on a subset and derivative

Let $F \in \mathbb{D}^{1,2}$, $A \in \mathcal{B}$. We suppose that F is $\mathcal{F}_A$-measurable. Then, μ-almost everywhere on A^c and $\mathbb{P}$-almost surely,

$$\mathrm{D}_t F = 0.$$

Lemma III.5 : Conditional expectation and Malliavin derivative

Let $F \in L^2$ having its Wiener chaos expansion given by $F = \sum_{n=0}^{+\infty} I_n(f_n)$. Let $A \in \mathcal{B}$.
(i) We have the following expansion in Wiener chaos for $\mathbb{E}[F|\mathcal{F}_A]$:

$$\mathbb{E}\left[F|\mathcal{F}_A\right] = \sum_{n=0}^{+\infty} I_n\left(f_n \mathbf{1}_A^{\otimes n}\right).$$

(ii) If $F \in \mathbb{D}^{1,2}$, we also have $\mathbb{E}[F|\mathcal{F}_A] \in \mathbb{D}^{1,2}$. We can compute the derivative of $\mathbb{E}[F|\mathcal{F}_A]$:

$$\mathrm{D}_t(\mathbb{E}[F|\mathcal{F}_A]) = \mathbb{E}[\mathrm{D}_t F|\mathcal{F}_A]\ \mathbf{1}_A(t),$$

for almost every $t \in T$, and $\mathbb{P}$-almost surely.

Proof of the lemma : (i) By linearity and density, we just have to check this formula for

$$f_n = \mathbf{1}_{B_1 \times \cdots \times B_n},$$

with $B_i \in \mathcal{B}$ two by two disjoints, and $\mu(B_i) < +\infty$. In this case, we just have

$$I_n(f_n) = \prod_{k=1}^{n} W(B_k),$$

and we can compute the conditional expectation in this case.

$$\mathbb{E}\left[I_n(f_n)|\mathcal{F}_A\right] = \mathbb{E}\left[\prod_{k=1}^{n} W(B_k)\ \middle|\ \mathcal{F}_A\right].$$

We make appear A, we get :

$$\mathbb{E}\left[I_n(f_n)|\mathcal{F}_A\right] = \mathbb{E}\left[\prod_{k=1}^{n} W(A\cap B_k) + W\left(A^c \cap B_k\right)\ \middle|\ \mathcal{F}_A\right].$$

We brutally expend the product :

$$\begin{aligned}&\mathbb{E}\left[I_n(f_n)|\mathcal{F}_A\right]\\ &= \sum_{k=0}^{n} \sum_{K\in\mathcal{P}_k(\llbracket 1,n\rrbracket)} \mathbb{E}\Big[\prod_{j\in K} W(A\cap B_k)\\ &\qquad \cdot \prod_{j\notin K} W\left(A^c\cap B_k\right)\ \Big|\ \mathcal{F}_A\Big].\end{aligned}$$

The terms "$W(A\cap B)$" are $\mathcal{F}_A$-measurable, so :

$$\begin{aligned}&\mathbb{E}\left[I_n(f_n)|\mathcal{F}_A\right]\\ &= \sum_{k=0}^{n} \sum_{K\in\mathcal{P}_k(\llbracket 1,n\rrbracket)} \prod_{j\in K} W(A\cap B_k)\\ &\qquad \cdot \mathbb{E}\left[\prod_{j\notin K} W\left(A^c\cap B_k\right)\ \middle|\ \mathcal{F}_A\right].\end{aligned}$$

The terms "$W(A^c\cap B)$" are independent of $\mathcal{F}_A$, so :

$$\begin{aligned}&\mathbb{E}\left[I_n(f_n)|\mathcal{F}_A\right]\\ &= \sum_{k=0}^{n} \sum_{K\in\mathcal{P}_k(\llbracket 1,n\rrbracket)} \prod_{j\in K} W(A\cap B_k)\\ &\qquad \cdot \mathbb{E}\left[\prod_{j\notin K} W\left(A^c\cap B_k\right)\right].\end{aligned}$$

The expectation is zero unless the product is over the empty set, meaning that $K = \llbracket 1,n\rrbracket$: the only non zero term is for $k = n$, and it yields to :

$$\mathbb{E}\left[I_n(f_n)|\mathcal{F}_A\right] = \prod_{j=1}^{n} W(B_j\cap A).$$

We get the definition of I_n of a tensor product :

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

Which is what we want.

(ii) We use **(i)** and twice the proposition **III.4**. We have by **(i)** the expansion in Wiener chaos of $\mathbb{E}[F|\mathcal{F}_A]$, so we can express its derivatives thanks to the previous proposition :

$$\mathrm{D}_t\mathbb{E}[F|\mathcal{F}_A] = \sum_{n=1}^{+\infty} nI_{n-1}\left(f_n(\cdot,t)\mathbf{1}_A^{\otimes(n-1)}\right)\mathbf{1}_A(t).$$

But, by the previous point for the variable $\mathrm{D}_t F$, whose expansion on Wiener chaos is known by the previous proposition, we have :

$$\mathbb{E}\left[\mathrm{D}_t F|\mathcal{F}_A\right] = \sum_{n=1}^{+\infty} nI_{n-1}\left(f_n(\cdot,t)\mathbf{1}_A^{\otimes(n-1)}\right).$$

Finally, we obtain what we want :

$$\mathrm{D}_t\mathbb{E}[F|\mathcal{F}_A] = \mathbf{1}_A(t)\ \mathbb{E}\left[\mathrm{D}_t F|\mathcal{F}_A\right].$$

□

We have everything to prove the proposition.

Proof : If F is $\mathcal{F}_A$-measurable, then the point **(ii)** of the lemma writes

$$\mathrm{D}_t F = \mathbb{E}[\mathrm{D}_t F|\mathcal{F}_A]\ \mathbf{1}_A(t)$$

which is zero when $t \notin A$. □

IV The divergence operator

Recall the definition of an adjoint in unbounded operator theory.

Definition IV.1

Let $(A, D(A))$ an unbounded operator on a Hilbert space E, valued in an Hilbert space F. If $D(A)$ is dense in E, we define the *adjoint* of A the unbounded operator $(A^*, D(A^*))$ on F, where

$$D(A^*) \overset{\text{def.}}{=} \{v \in F, \exists C > 0, \forall u \in D(A), |\langle Au, v\rangle_F| \leqslant C\|u\|_E\},$$

and A^* is uniquely defined by the duality formula :

$$\forall u \in D(A), \forall v \in D(A^*), \langle Au, v\rangle_F = \langle u, A^*v\rangle_E.$$

In our case, δ would be the adjoint of the Malliavin derivative on L^2, whose domain $\mathbb{D}^{1,2}$ contains the set of smooth random variables which is dense in $L^2(\mathbb{P})$.

Definition IV.2

We call *divergence operator* the unbounded operator $(\delta, D(\delta))$ on $L^2(\Omega \to \mathcal{H})$ with values on $L^2(\mathbb{P})$ with :

$$D(\delta) \overset{\text{def.}}{=} \left\{u \in L^2(\Omega \to \mathcal{H}), \exists C \geqslant 0, \forall F \in \mathbb{D}^{1,2}, |\mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right]|^2 \leqslant C\mathbb{E}\left[|F|^2\right]\right\}.$$

and the operator δ is uniquely defined by the duality formula :

$$\forall F \in \mathbb{D}^{1,2}, \forall u \in D(\delta), \mathbb{E}\left[F\delta(u)\right] = \mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right].$$

Remark : it is equivalent to substitute all the "$\forall F \in \mathbb{D}^{1,2}$" by "$\forall F \in \mathcal{S}$".

To make things easier, and to avoid the using of tensor products, we immediately does the things in $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$. The objects introduced in this section could be defined on general real separable Hilbert spaces, we will explain it in a few remarks.

IV.1 Computation on simple elements

This part remains in the case where $\mathcal{H}$ is a general real separable Hilbert space.

A first result is a generalization of our integration by parts result. We can now integrates by parts not only with $h \in \mathcal{H}$ but with every $u \in D(\delta)$.

Lemma IV.1 : Integration by parts

Let $F, G \in \mathbb{D}^{1,2}$ and $u \in D(\delta)$. Then

$$\mathbb{E}[G\langle \mathrm{D}F, u\rangle_{\mathcal{H}}] = \mathbb{E}[FG\ \delta(u)] - \mathbb{E}[F\langle \mathrm{D}G, u\rangle_{\mathcal{H}}].$$

Proof of the lemma : We do the same computations as the case where u is deterministic. We have

$$\mathrm{D}(FG) = G\mathrm{D}F + F\mathrm{D}G.$$

Hence,

$$\mathbb{E}\left[\langle \mathrm{D}(FG), u\rangle\right] = \mathbb{E}\left[F\langle \mathrm{D}G, u\rangle\right] + \mathbb{E}\left[G\langle \mathrm{D}F, u\rangle\right].$$

Finally, the first term gives by duality :

$$\mathbb{E}\left[\langle \mathrm{D}(FG), u\rangle\right] = \mathbb{E}\left[FG\ \delta(u)\right].$$

This completes the proof. □

Let us compute this operator for a class of random variables which would play the role of simple elements.

Lemma IV.2 : Simple elements of $L^2(\Omega \to \mathcal{H})$

Let us note $\mathcal{S}_{\mathcal{H}} \subset L^2(\Omega \to \mathcal{H})$ the set of elements of the type

$$u = \sum_{j=1}^{n} Z_j h_j,$$

with $Z_j \in \mathcal{S}$ smooth and $h_j \in \mathcal{H}$. Then $\mathcal{S}_{\mathcal{H}} \subset D(\delta)$ and for all $u \in \mathcal{S}_{\mathcal{H}}$:

$$\delta(u) = \sum_{j=1}^{n} Z_j X(h_j) - \sum_{j=1}^{n} \langle \mathrm{D} Z_j, h_j \rangle_{\mathcal{H}}.$$

Proof of the lemma : Let $u \in \mathcal{S}_{\mathcal{H}}$ of the proposed form in the lemma. Let $F \in \mathbb{D}^{1,2}$. Then, by linearity :

$$\mathbb{E}\left[\langle \mathrm{D}F, u\rangle\right] = \sum_{j=1}^{n} \mathbb{E}\left[Z_j \langle \mathrm{D}F, h_j\rangle_{\mathcal{H}}\right].$$

By integration by parts formula :

$$\mathbb{E}\left[\langle \mathrm{D}F, u\rangle\right] = \sum_{j=1}^{n} \mathbb{E}\left[Z_j F\ X(h_j)\right] - \mathbb{E}\left[F \langle \mathrm{D}Z_j, h_j\rangle\right].$$

Thanks to this, we can use Cauchy-Schwarz inequality to conclude that

$$\begin{aligned}&\left|\mathbb{E}\left[\langle \mathrm{D}F, u\rangle\right]\right| \\ \leqslant\ &\left[\sum_{j=1}^{n} \mathbb{E}\left[|Z_j\ X(h_j)|^2\right]^{\frac{1}{2}} + \mathbb{E}\left[|\langle \mathrm{D}Z_j, h_j\rangle|^2\right]^{\frac{1}{2}}\right] \mathbb{E}[|F|^2]^{\frac{1}{2}}.\end{aligned}$$

So, $\mathcal{S}_{\mathcal{H}} \subset D(\delta)$. Moreover, we have by unicity of the duality formula the expression of $\delta(u)$.

□

Proposition IV.1 : Homogeneity up to random constant

Let $u \in D(\delta)$. Then for all $F \in \mathbb{D}^{1,2}$ such that the both following expectations are finite :

$$\mathbb{E}\left[F^2 \delta(u)^2\right], \mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}^2\right] < +\infty,$$

then $Fu \in D(\delta)$ and we have :

$$\delta(Fu) = F\delta(u) - \langle \mathrm{D}F, u\rangle_{\mathcal{H}}.$$

Proof : Let $G \in \mathcal{S}$. Then,

$$\mathbb{E}[\langle \mathrm{D}G, Fu\rangle_{\mathcal{H}}] = \mathbb{E}[F\langle \mathrm{D}G, u\rangle_{\mathcal{H}}].$$

By integration by parts formula :

$$\mathbb{E}[\langle \mathrm{D}G, Fu\rangle_{\mathcal{H}}] = \mathbb{E}[FG\ \delta(u)] - \mathbb{E}[G\langle \mathrm{D}F, u\rangle_{\mathcal{H}}].$$

We get by Cauchy-Schwarz :

$$\begin{aligned}&\left|\mathbb{E}\left[\langle \mathrm{D}G, Fu\rangle_{\mathcal{H}}\right]\right| \\ \leqslant\ &\mathbb{E}\left[G^2\right]^{\frac{1}{2}} \mathbb{E}\left[F^2\delta(u)^2\right]^{\frac{1}{2}} + \mathbb{E}\left[G^2\right]^{\frac{1}{2}} \mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}^2\right]^{\frac{1}{2}}.\end{aligned}$$

This proves that $Fu \in D(\delta)$, and it gives us the expression of $\delta(Fu)$. □

We finish by a proposition allowing to determine if a limit of a sequence of elements of $D(\delta)$ is still in $D(\delta)$.

Proposition IV.2 : Limit of elements of the domain of the divergence operator

Let $(u_n)_n \in D(\delta)^{\mathbb{N}}$, $u \in L^2(\Omega \to \mathcal{H})$ and $G \in L^2(\mathbb{P})$. We suppose that :

* The sequence $(u_n)_n$ converges to u in $L^2(\Omega \to \mathcal{H})$;
* The sequence $(\delta(u_n))_n$ converges to G in L^2.

Then, $u \in D(\delta)$ and $G = \delta(u)$.

Proof : To show that $u \in D(\delta)$, we will show that for all $F \in \mathbb{D}^{1,2}$:

$$\mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right] = \mathbb{E}[FG].$$

Indeed, since by Cauchy Schwarz, the map

$$u \longmapsto \mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right]$$

is continuous on $L^2(\Omega \to \mathcal{H})$, we have :

$$\mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right] = \lim_{n\to+\infty} \mathbb{E}\left[\langle \mathrm{D}F, u_n\rangle_{\mathcal{H}}\right].$$

By integration by parts formula on $D(\delta)$, we have :

$$\mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right] = \lim_{n\to+\infty} \mathbb{E}\left[F\delta(u_n)\right].$$

Since, by Cauchy-Schwarz once again, the map

$$G \longmapsto \mathbb{E}[FG]$$

is continuous on L^2, we have :

$$\mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[FG\right].$$

We proved our equality. It is now time to conclude. We have for all $F \in \mathbb{D}^{1,2}$:

$$\left|\mathbb{E}\left[\langle \mathrm{D}F, u\rangle_{\mathcal{H}}\right]\right| \leqslant \mathbb{E}\left[G^2\right]^{\frac{1}{2}} \mathbb{E}\left[F^2\right]^{\frac{1}{2}}.$$

It means that $u \in D(\delta)$. Moreover, by unicity of the adjoint, it directly gives that $G = \delta(u)$. □

IV.2 Differentiation on higher spaces and Heisenberg relation

We defined previously the multiple derivative of an element of $L^2(\mathbb{P})$. We present now the notion of derivative for elements of $L^2(\Omega \to \mathcal{H})$, which could be seen also as $L^2(T \to \Omega)$ for $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$.

Definition IV.3

We consider $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$. Let $u \in \mathcal{S}_{\mathcal{H}}$, given by the lemma **IV.2**. Then, its *Malliavin derivative* will be the element $\mathrm{D}u \in L^2(T^2 \times \Omega)$ defined by :

$$\mathrm{D}u \overset{\text{def.}}{=} \sum_{j=1}^{n} \mathrm{D}Z_j \otimes h_j.$$

In other words, for any $s, t \in T$:

$$\mathrm{D}_t u_s \overset{\text{def.}}{=} \sum_{j=1}^{n} \mathrm{D}_t Z_j \; h_j(s).$$

For general $\mathcal{H}$, the first equality is the one we use, $\mathrm{D}u$ is with values in $\mathcal{H} \otimes \mathcal{H}$. By the same argument as we used for the operator D, this operator is closable. We introduce the following norm on $L^2(\Omega \to \mathcal{H})$:

$$\forall u \in \mathcal{S}_{\mathcal{H}}, \|u\|_{1,2}^2 = \mathbb{E}\left[\|u\|_{\mathcal{H}}^2\right] + \mathbb{E}\left[\int_{T^2} |\mathrm{D}_t u_s|^2 \,\mathrm{d}\mu(t)\,\mathrm{d}\mu(s)\right].$$

Then we define $\mathbb{D}^{1,2}(\mathcal{H})$ as the closure of $\mathcal{S}_{\mathcal{H}}$ for this norm.

Lemma IV.3 : Heisenberg relation

Let $u \in \mathcal{S}_{\mathcal{H}}$. Then, for all $t \in T$, $\mathrm{D}_t u = (\mathrm{D}_t u_s)_{s \in T} \in D(\delta)$, $\delta u \in \mathbb{D}^{1,2}$ and, for all $t \in T$:

$$\mathrm{D}_t(\delta u) - \delta\left(\mathrm{D}_t u\right) = u_t.$$

Proof of the lemma : We compute both terms in this relation, and see where it yields to. We have first the two following expression for $\mathrm{D}_t u$ and $\delta(u)$:

$$\mathrm{D}_t u_s = \sum_{j=1}^{n} \mathrm{D}_t Z_j h_j(s)$$

and

$$\delta(u) = \sum_{j=1}^{n} Z_j X(h_j) - \sum_{j=1}^{n} \int_T \mathrm{D}_t Z_j h_j(t)\,\mathrm{d}\mu(t).$$

It proves the very first part of the lemma.

- First, let us compute $\delta(\mathrm{D}_t u)$. By linearity, we have :

$$\delta(\mathrm{D}_t u) = \sum_{j=1}^{n} \delta(\mathrm{D}_t Z_j \; h_j).$$

By using the expression of δ for simple elements, since $\mathrm{D}_t Z_j \in \mathcal{S}$, it gives :

$$\delta(\mathrm{D}_t u) = \sum_{j=1}^{n} D_t Z_j \delta(h_j) - \sum_{j=1}^{n} \int_T \mathrm{D}^2_{s,t} Z_j h_j(s)\,\mathrm{d}\mu(s).$$

- Second, we compute $\mathrm{D}_t(\delta(u))$. By linearity :

$$\mathrm{D}_t(\delta(u)) = \sum_{j=1}^{n} \mathrm{D}_t(Z_j X(h_j)) - \sum_{j=1}^{n} \mathrm{D}_t\left[\int_T \mathrm{D}_s Z_j h_j(s)\,\mathrm{d}\mu(s)\right].$$

We have for the first term :

$$\mathrm{D}_t(Z_j X(h_j)) = \mathrm{D}_t Z_j X(h_j) + Z_j h_j.$$

Hence,

$$\sum_{j=1}^{n} \mathrm{D}_t(Z_j X(h_j)) = u_t + \sum_{j=1}^{n} \mathrm{D}_t Z_j X(h_j),$$

which is the first term of $\delta_t(\mathrm{D}u)$. Let us compute :

$$\text{(A)} = \sum_{j=1}^{n} \mathrm{D}_t\left[\int_T \mathrm{D}_s Z_j h_j(s)\,\mathrm{d}\mu(s)\right].$$

To do this, we note

$$Z_j = z_j(X(g_1), \cdots, X(g_m)).$$

Then, if we note

$$\text{(A')} = \int_T \mathrm{D}_s Z_j h_j(s)\,\mathrm{d}\mu(s),$$

we have

$$\text{(A')} = \sum_{i=1}^{m} \left(\int_T g_i(s) h_j(s)\,\mathrm{d}\mu(s)\right) \frac{\partial z_j}{\partial x_i}(X(g_1), \cdots, X(g_m)).$$

And we can compute its derivative :

$$\begin{aligned}\mathrm{D}_t\text{(A')} &= \sum_{l=1}^{m}\sum_{i=1}^{m} \left(\int_T g_i(s) h_j(s)\,\mathrm{d}\mu(s)\right) \\ &\cdot \frac{\partial^2 z_j}{\partial x_l \partial x_i}(X(g_1), \cdots, X(g_m)) g_l(t).\end{aligned}$$

We insert everything into the integral, and we find that we get $\mathrm{D}^2_{s,t} Z_j$:

$$\mathrm{D}_t(\mathrm{A'}) = \int_T \mathrm{D}^2_{s,t} Z_j h_j(s) \,\mathrm{d}\mu(s).$$

Finally, (A) is equal to the second term of $\delta_t(\mathrm{D}u)$. This concludes the proof. □

Proposition IV.3 : Inner product

We consider $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$. The space $\mathbb{D}^{1,2}(\mathcal{H})$ is included to the domain $D(\delta)$. Moreover, for all $u, v \in \mathbb{D}^{1,2}(\mathcal{H})$:

$$\mathbb{E}\left[\delta(u)\delta(v)\right] = \mathbb{E}\left[\int_T u_t v_t \,\mathrm{d}\mu(t)\right] + \mathbb{E}\left[\int_{T^2} \mathrm{D}_s u_t \mathrm{D}_t v_s \,\mathrm{d}\mu(t)\,\mathrm{d}\mu(s)\right].$$

Proof : We make a straightforward calculus. By definition of δ as adjoint :

$$\mathbb{E}\left[\delta(u)\delta(v)\right] = \mathbb{E}\left[\int_T v_t \mathrm{D}_t \delta(u) \,\mathrm{d}\mu(t)\right].$$

By Heisenberg relation :

$$\mathbb{E}\left[\delta(u)\delta(v)\right] = \mathbb{E}\left[\int_T v_t u_t \,\mathrm{d}\mu(t)\right] + \mathbb{E}\left[\int_T v_t \delta \mathrm{D}_t u \,\mathrm{d}\mu(t)\right].$$

Since v and $\delta \mathrm{D}_\bullet u$ are in $L^2(T)$, we have :

$$\mathbb{E}\left[\delta(u)\delta(v)\right] = \mathbb{E}\left[\int_T v_t u_t \,\mathrm{d}\mu(t)\right] + \int_T \mathbb{E}\left[v_t \delta \mathrm{D}_t u\right] \,\mathrm{d}\mu(t).$$

And we use again the duality formula for δ for the variables, with a fixed t :

$$\begin{aligned}&\mathbb{E}\left[\delta(u)\delta(v)\right] \\ = \;&\mathbb{E}\left[\int_T v_t u_t \,\mathrm{d}\mu(t)\right] + \int_T \mathbb{E}\left[\int_T \mathrm{D}_t u_s \mathrm{D}_s v_t \,\mathrm{d}\mu(s)\right] \,\mathrm{d}\mu(t).\end{aligned}$$

Which is what we expected. □

We interpret this result like this : if $u \in L^2_{\mathrm{a}}$ is adapted and square integrable (we will define it later) then by Itô's isometry :

$$\mathbb{E}\left[\left|\int_T u_t \,\mathrm{d}B_t\right|^2\right] = \mathbb{E}\left[\int_T u_t^2 \,\mathrm{d}t\right].$$

Since, like we will see later, δ is an extension of the Itô integral, it means here that we lost the isometry property for $u \in D(\delta)$, because of the second term. It is zero whenever u or v is adapted, since in this case, the derivative will be zero on the set where u or v is not measurable.

IV.3 Multiple integrals and divergence

Lemma IV.4 : Expectation of the L^2 norm of an element of $L^2(\Omega \to \mathcal{H})$

Let $u \in L^2(\Omega \to \mathcal{H})$ having its Wiener decomposition given by $u_t = \sum_{n=0}^{+\infty} I_n(f_n(\cdot, t))$, with $f_n \in L^2(T^n \times T)$ symmetric with respect to its n first variables. Then,

$$\mathbb{E}\left[\int_T u(t)^2 \,\mathrm{d}\mu(t)\right] = \sum_{n=0}^{+\infty} n!\, \|f_n\|^2_{L^2(T^{n+1})}.$$

Proof of the lemma : Recall by an ancient proposition that if $f \in L^2(T^m)$ and $g \in L^2(T^q)$ then

$$\mathbb{E}\left[I_m(f) I_q(g)\right] = m! \langle \tilde{f}, \tilde{g} \rangle_{L^2(T^m)} \delta_{q,m},$$

where $\tilde{f}$ is the symmetrization of f. By Fubini-Tonelli, we have :

$$\mathbb{E}\left[\int_T u(t)^2 \,\mathrm{d}\mu(t)\right] = \int_T \mathbb{E}\left[u(t)^2\right] \,\mathrm{d}\mu(t).$$

By the decomposition in Wiener chaos :

$$\mathbb{E}\left[\int_T u(t)^2 \,\mathrm{d}\mu(t)\right] = \int_T \sum_{n=0}^{+\infty} \mathbb{E}\left[I_n(f_n(\cdot,t))^2\right] \,\mathrm{d}\mu(t).$$

Since f_n has its first n components which are symmetric, we have :

$$\mathbb{E}\left[\int_T u(t)^2 \,\mathrm{d}\mu(t)\right] = \int_T \sum_{n=0}^{+\infty} n! \|f_n(\cdot,t)\|^2_{L^2(T^n)} \,\mathrm{d}\mu(t).$$

By a new switching, we finally have :

$$\mathbb{E}\left[\int_T u(t)^2 \,\mathrm{d}\mu(t)\right] = \sum_{n=0}^{+\infty} n! \|f_n\|^2_{L^2(T^{n+1})}.$$

□

Lemma IV.5 : Projection of divergence on Wiener chaos

Let $u \in L^2(\Omega \to \mathcal{H})$ having its Wiener decomposition given by $u_t = \sum_{n=0}^{+\infty} I_n(f_n(\cdot,t))$, with $f_n \in L^2(T^n \times T)$ symmetric with respect to its n first variables. Let $G = I_n(g) \in \mathfrak{H}_n$, with $g \in L^2(T^n)$ symmetric. Then,

$$\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right] = \mathbb{E}\left[G I_n(f_{n-1})\right].$$

Proof of the lemma : The computation uses the same arguments as the previous lemma. Notice first that

$$\mathrm{D}_t G = n I_{n-1}(g(\cdot,t)),$$

by computation of derivatives with a Wiener chaos expansion. As a consequence, after using the Fubini theorem :

$$\begin{aligned}&\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right]\\ =\ & n \sum_{m=0}^{+\infty} \int_T \mathbb{E}\left[I_m(f_m(\cdot,t)) I_{n-1}(g(\cdot,t))\right] \, \mathrm{d}\mu(t).\end{aligned}$$

Since we are in presence of symmetric functions, we have :

$$\begin{aligned}&\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right]\\ =\ & n(n-1)! \int_T \langle f_{n-1}(\cdot,t), g(\cdot,t)\rangle_{L^2(T^{n-1})} \, \mathrm{d}\mu(t).\end{aligned}$$

Which gives

$$\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right] = n! \langle f_{n-1}, g\rangle_{L^2(T^n)}.$$

Since g is symmetric, we have :

$$\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right] = n! \langle \tilde{f}_{n-1}, g\rangle_{L^2(T^n)}.$$

And so :

$$\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right] = \mathbb{E}\left[I_n(f_{n-1}) I_n(g)\right].$$

By definition of G, it gives :

$$\mathbb{E}\left[\int_T u_t \, \mathrm{D}_t G \, \mathrm{d}\mu(t)\right] = \mathbb{E}\left[I_n(f_{n-1}) G\right].$$

□

Remark : Let $f_n(\cdot,t)$ like previously. Then, we have

$$\tilde{f}_n = \frac{1}{n+1} \sum_{k=1}^{n+1} f \circ \tau_k$$

where τ_k is the permutation between the k-th component and the last, τ_{n+1} is the identity.

Theorem IV.1 : Domain of δ and Wiener chaos

Let $u \in L^2(\Omega \to \mathcal{H})$ having its Wiener decomposition given by $u_t = \sum_{n=0}^{+\infty} I_n(f_n(\cdot,t))$, with $f_n \in L^2(T^n \times T)$ symmetric with respect to its n first variables. Then, $u \in D(\delta)$ if and only if the series $\sum_{n\geqslant 0} I_{n+1}(f_n)$ converges in $L^2(\mathbb{P})$. In this case, we have in L^2 :

$$\delta(u) = \sum_{n=0}^{+\infty} I_{n+1}(f_n).$$

Proof : $[\Longrightarrow]$ Let $u \in D(\delta)$. Then, by the previous lemma,

$$\forall G \in \mathfrak{H}_n, \mathbb{E}[G\delta(u)] = \mathbb{E}[G I_n(f_{n-1})].$$

But, $\delta(u)$ expends in Wiener chaos by $\delta(u) = \sum_{n=0}^{+\infty} J_n(\delta(u))$. $J_n(\delta(u))$ is the unique element of $\mathfrak{H}_n$ such that

$$\forall G \in \mathfrak{H}_n, \mathbb{E}[G\delta(u)] = \mathbb{E}[G J_n(\delta(u))].$$

Hence, for all $n \geqslant 1$, $J_n(\delta(u)) = I_n(f_{n-1})$. Since $\sum_n J_n(\delta(u))$ converges in L^2, the series $\sum_n I_n(f_{n-1})$ converges in L^2 and $I_n(f_{n-1})$ is the projection of $\delta(u)$ on $\mathfrak{H}_n$.

$[\Longleftarrow]$ Suppose that $\sum_{n\geqslant 1} I_n(f_{n-1})$ converges in L^2, and note V its limit. Let $G = \sum_{n=0}^{N} I_n(g_n)$. Then, by the previous lemma :

$$\mathbb{E}\left[\int_T u_t \mathrm{D}_t G \, \mathrm{d}\mu(t)\right] = \sum_{n=0}^{N} \mathbb{E}\left[I_n(g_n) I_n(f_{n-1})\right].$$

Hence, by Cauchy-Schwarz and orthogonality of Wiener chaos :

$$\begin{aligned}&\left|\mathbb{E}\left[\int_T u_t \mathrm{D}_t G \, \mathrm{d}\mu(t)\right]\right|\\ \leqslant\ & \mathbb{E}\left[\sum_{n=0}^{N} I_n(f_{n-1})^2\right]^{\frac{1}{2}} \mathbb{E}\left[\sum_{n=0}^{N} I_n(g_n)^2\right]^{\frac{1}{2}}.\end{aligned}$$

We have by Bessel inequality :

$$\left|\mathbb{E}\left[\int_T u_t \mathrm{D}_t G \, \mathrm{d}\mu(t)\right]\right| \leqslant \mathbb{E}\left[V^2\right]^{\frac{1}{2}} \mathbb{E}\left[G^2\right]^{\frac{1}{2}}.$$

Since this equality is true for every random variable having a finite Wiener chaos decomposition, and since all the members are continuous with respect to G for the norm $\|\cdot\|_{1,2}$ in L^2, we conclude that this inequality is true for every $G \in \mathbb{D}^{1,2}$, and then it follows that $u \in D(\delta)$.

□

Corollary IV.1 : Domain of the divergence

We have

$$D(\delta) = \left\{ u \in L^2(\Omega \to \mathcal{H}), \sum_{n=0}^{+\infty} (n+1)! \left\| \tilde{f}_n \right\|^2_{L^2(T^{n+1})} < +\infty \right\}$$

where we decomposed u like in the previous proposition. For all $u \in D(\delta)$, the sum of this series is $\mathbb{E}[\delta(u)^2]$:

$$\mathbb{E}\left[\delta(u)^2\right] = \sum_{n=0}^{+\infty} (n+1)! \left\| \tilde{f}_n \right\|^2_{L^2(T^{n+1})}.$$

Proof : [$\subset$] If $u \in D(\delta)$, then by the previous theorem, we have

$$\mathbb{E}\left[\delta(u)^2\right] = \mathbb{E}\left[\sum_{n=0}^{+\infty} I_{n+1}(f_n)^2\right].$$

By orthogonality in Wiener chaos, and by the expression of the L^2 norm of I_n, we have :

$$\mathbb{E}\left[\delta(u)^2\right] = \sum_{n=0}^{+\infty} (n+1)! \left\| \tilde{f}_n \right\|^2_{L^2(T^{n+1})},$$

which is finite, since $u \in D(\delta)$.

[$\supset$] If $\sum_{n \geqslant 0} (n+1)! \left\| \tilde{f}_n \right\|^2_{L^2(T^{n+1})}$ converges, then

$$\sum_{n=0}^{+\infty} (n+1)! \left\| \tilde{f}_n \right\|^2_{L^2(T^{n+1})} = \sum_{n=0}^{+\infty} \mathbb{E}\left[I_n(f_{n+1})^2\right].$$

By Fubini :

$$\sum_{n=0}^{+\infty} (n+1)! \left\| \tilde{f}_n \right\|^2_{L^2(T^{n+1})} = \mathbb{E}\left[\sum_{n=0}^{+\infty} I_n(f_{n+1})^2\right].$$

Since this sum converges, the expectation is finite, and then $\sum_{n=0}^{+\infty} I_n(f_{n+1})^2$ is finite in L^2. But, if $p, N \in \mathbb{N}$:

$$\mathbb{E}\left[\left(\sum_{n=N}^{p+N} I_{n+1}(f_n)\right)^2\right] \leqslant \mathbb{E}\left[\sum_{n=N}^{+\infty} I_{n+1}(f_n)^2\right]$$

That is $\sum_n I_n(f_{n+1})$ is a Cauchy sequence in L^2, so converges, and so $F \in D(\delta)$ by the previous theorem. □

IV.4 Skorohod integral

IV.4.1 Recalls on Itô's stochastic integration

The divergence operator allows us to extend the Itô's stochastic integral. We give here some properties of this integration.

Let $(B_t)_t$ a Brownian motion, defined for instance on the canonical space : that is $\Omega = \mathcal{C}^0(\mathbb{R}_+)$, $\mathcal{F} = \mathcal{B}(\Omega)$ and $\mathbb{P}$ defined as, for all cylinder, with $A_1, \cdots, A_n \in \mathcal{B}(\mathbb{R})$:

$$C = \{\omega \in \Omega, \omega_{t_1} \in A_1, \cdots, \omega_{t_n} \in A_n\}$$

we have :

$$\mathbb{P}(C) \stackrel{\text{def.}}{=} \int_{A_1 \times \cdots \times A_n} p_{t_1}(x_1) \prod_{k=2}^{n} p_{t_k - t_{k-1}}(x_k - x_{k-1}) \, \mathrm{d}x_n \cdots \, \mathrm{d}x_1,$$

where

$$p_t(x) \stackrel{\text{def.}}{=} \frac{1}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}}.$$

Then, $B_t : \omega \longrightarrow \omega_t$ is a Brownian motion : it is a centered Gaussian process such that $\mathbb{E}[B_t B_s] = s \wedge t$. We want to define an integration theory with (B_t) : $\int_T u_t \, \mathrm{d}B_t$. We can't do this thanks to Stieljes integration because $(B_t)_t$ has an infinite total variation. Indeed, we can prove the following lemma about quadratic variation of (B_t).

Lemma IV.6 : Quadratic variation of Brownian motion

Let for all $t \geqslant 0$ and $n \in \mathbb{N}^*$:

$$S_n(t) \stackrel{\text{def.}}{=} \sum_{k=1}^{2^n} \left(B_{\frac{tk}{2^n}} - B_{\frac{t(k+1)}{2^n}}\right)^2.$$

Then,

$$S_n(t) \xrightarrow[n\to+\infty]{\mathbb{P}\text{-a.s}, L^2} t.$$

Proof of the lemma : • Let us prove first the convergence in L^2. We just need to use that $\mathbb{E}[N^4] = 3$ when $N \sim \mathcal{N}(0,1)$. We just expend

$$\Delta_n \stackrel{\text{def.}}{=} \mathbb{E}\left[|S_n(t) - t|^2\right].$$

We use the independence of the increments and the fact that $B_t \sim \mathcal{N}(0,t)$ to get :

$$\Delta_n = \frac{t}{2^{n-1}},$$

which goes to zero when $[n \to +\infty]$, so the L^2-convergence is checked.

• For the $\mathbb{P}$-almost sure convergence, we use the Borel-Cantelli lemma. Indeed, for all $\varepsilon > 0$, by Bienaymé-Tchebychev inequality,

$$\mathbb{P}(|S_n(t) - t| > \varepsilon) \leqslant \frac{1}{\varepsilon^2 2^{n-1}}.$$

Hence, the series $\sum_n \mathbb{P}(|S_n(t) - t| > \varepsilon)$ converges, so $S_n(t)$ converges $\mathbb{P}$-almost surely to t. □

We will then define this integral by density. We want a property on simple elements then we take the closure of those simple elements, exactly like we did for Wiener integral. In the following, we set T an interval of $\mathbb{R}_+$.

Definition IV.4

We say that a process $U = (U_t)_{t\in T}$ is *elementary* if there exists $t_1 < t_2 < \cdots < t_{n+1}$ elements of T, and $(F_i)_i$ which are $(\mathcal{F}_{t_i})_i$-random variables such that for all $t \in T$:

$$U_t = \sum_{i=1}^{n} F_i \mathbf{1}_{]t_i,t_{i+1}]}.$$

We note $\mathcal{E}_0$ the set of elementary process.

Notice that we opened at the left and closed at the right. In this case, we define our integral like this.

Definition IV.5

Let $U \in \mathcal{E}_0$ defines like in the previous definition. Then, we define its *Itô's integral* as :

$$\int_T U_t \; \mathrm{d}B_t \stackrel{\text{def.}}{=} \sum_{i=1}^{n} F_i \left(B_{t_{i+1}} - B_{t_i}\right).$$

Notice that this a particular case of Wiener integral. Indeed, we are with $\mathcal{H} = L^2(\mathbb{R}_+)$ with $W([0,t]) = B_t$. This why we will write $\int_T U_t \; \mathrm{d}B_t$ or $\int_T U_t \; \mathrm{d}W_t$ in the following for designing the same object.

Proposition IV.4 : Itô's isometry

Let $U \in \mathcal{E}_0$. Then we have :

(i) The Itô integral is centered :

$$\mathbb{E}\left[\int_T U_t \; \mathrm{d}B_t\right] = 0;$$

(ii) We have the following isometry equality between $L^2(\mathbb{P})$ and $L^2(T \times \Omega)$:

$$\mathbb{E}\left[\left|\int_T U_t \; \mathrm{d}B_t\right|^2\right] = \mathbb{E}\left[\int_T U_t^2 \; \mathrm{d}t\right].$$

Proof : **(i)** Suppose that U is given by the definition **IV.4**. Then, since F_i is $\mathcal{F}_{t_i}$-measurable, and since $B_{t_{i+1}} - B_{t_i}$ is independent of $\mathcal{F}_{t_i}$, we have :

$$\mathbb{E}\left[\int_T U_t \; \mathrm{d}B_t\right] = \sum_{i=1}^{n} \mathbb{E}[F_i]\mathbb{E}\left[B_{t_{i+1}} - B_{t_i}\right] = 0.$$

(ii) We expend the square of the sum and use the previous independent property.

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_T U_t \, \mathrm{d}B_t\right)^2\right] \\
= \quad & \sum_{i=1}^{n} \mathbb{E}\left[F_i^2\right] \mathbb{E}\left[\left(B_{t_{i+1}} - B_{t_i}\right)^2\right] \\
&+2\sum_{i<j} \mathbb{E}\left[F_i F_j \left(B_{t_{i+1}} - B_{t_i}\right)\right] \mathbb{E}\left[B_{t_{j+1}} - B_{t_j}\right].
\end{aligned}$$

The second sum is zero. The first is simply :

$$\sum_{i=1}^{n} \mathbb{E}\left[F_i^2\right](t_{i+1} - t_i) = \mathbb{E}\left[\int_T U_t^2 \, \mathrm{d}t\right].$$

This concludes this proof. □

We will extend thanks to this the Itô's integral to any *adapted* process.

Definition IV.6

Let $U = (U_t)_t$ a stochastic process. We say that U is *adapted* with respect to $(\mathcal{F}_t)_t$ if for all $t \in T$, U_t is $\mathcal{F}_t$-measurable.
We will note $L^2_{\mathrm{a}}(T \times \Omega)$ all the adapted processes on T such that :

$$\int_T \mathbb{E}[U_t^2] \, \mathrm{d}t < +\infty.$$

We have chosen $\mathcal{F}_t = \sigma(B_s, s \leqslant t)$.

Lemma IV.7 : Density of elementary processes

We suppose that $T = [a, b]$ is a **segment**. The space $\mathcal{E}_0$ is dense into $L^2_{\mathrm{a}}(T)$ for the norm

$$\|U\|^2_{L^2_{\mathrm{a}}} = \int_T \mathbb{E}[U_t^2] \, \mathrm{d}t.$$

More precisely, every process $U \in L^2_{\mathrm{a}}(T \times \Omega)$ can be approximated by

$$P_n U_t \overset{\text{def.}}{=} 2^n \sum_{i=1}^{2^n} \left(\int_{a+(i-1)\frac{b-a}{2^n}}^{a+i\frac{b-a}{2^n}} U_s \, \mathrm{d}s\right) \mathbf{1}_{]a+i\frac{b-a}{2^n}, a+(i+1)\frac{b-a}{2^n}]}(t).$$

The idea of the proof is wrong in Nualart, it is coming from Karatzas-Shreve.

Proof of the lemma : It is enough to show it for $T = [0, 1]$. Here's the plan :

(i) For all continuous process U, we have $(P_n U)_n$ converging to U for the norm L^2_{a} ;
(ii) We approach $U \in L^2_{\mathrm{a}}(T \times \Omega)$ bounded by a sequence of continuous processes V^n, which can be approximated by $(P_m V^n)_m$ by the point **(i)**. We let $[n \to +\infty]$.
(iii) We approach every process $U \in L^2_{\mathrm{a}}$ by processes $U_M \in L^2_{\mathrm{a}}$ bounded by $M > 0$, and conclude by letting $[M \to +\infty]$.
(iv) We show that for every n, P_n is a continuous linear operator, so that $P_n U$ approaches U like announced in the statement.

Here we go.
(i) Let us note

$$\Delta_n \overset{\text{def.}}{=} \mathbb{E}\left[\int_T |P_n U_t - U_t|^2 \, \mathrm{d}t\right].$$

Then, we write :

$$U_t = 2^n \sum_{i=1}^{2^n} \left(\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} U_t \, \mathrm{d}s\right) \mathbf{1}_{]\frac{i}{2^n}, \frac{i+1}{2^n}]}(t).$$

By definition and by Chasles :

$$\Delta_n = 4^n \sum_{i=1}^{2^n} \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} \mathbb{E}\left[\left|\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} (U_s - U_t) \, \mathrm{d}s\right|^2\right] \mathrm{d}t.$$

Then, by Cauchy-Schwarz inequality :

$$\Delta_n \leqslant 4^n \sum_{i=1}^{2^n} \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} \frac{1}{4^n} \mathbb{E}\left[\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} (U_s - U_t)^2 \, \mathrm{d}s\right] \mathrm{d}t.$$

Finally, since we consider U a continuous process, then almost surely, $(U_t)_t$ is uniformly continuous on $[0, 1]$. If $\varepsilon > 0$, then for n big enough, we have :

$$\Delta_n \leqslant \sum_{i=1}^{2^n} \int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} \mathbb{E}\left[\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} \varepsilon \, \mathrm{d}s\right] \mathrm{d}t = \varepsilon.$$

That proves **(i)**, for every continuous process U.
(ii) Let $U \in L^2_{\mathrm{a}}$, bounded by $M > 0$. We define for $t > 0$:

$$V_t^n \overset{\text{def.}}{=} n \int_{t-\frac{1}{n}}^{t} U_s \, \mathrm{d}s.$$

By Cauchy-Schwarz,

$$\mathbb{E}\left[V_t^{n\,2}\right] \leqslant \int_T \mathbb{E}[U_s^2] \, \mathrm{d}s < +\infty,$$

so V_t^n is $\mathbb{P}$-almost surely finite. Moreover, V^n is a continuous adapted process. $(V_t^n)_n$ converges $\mathbb{P}$-almost surely to U_t. Finally, V^n is also bounded by M. By dominated convergence, it implies that

$$\int_T \mathbb{E}\left[(V_t^n - U_t)^2\right] \, \mathrm{d}t \xrightarrow[n \to +\infty]{} 0.$$

Hence, if we set $\varepsilon > 0$, and n big enough such that

$$\int_T \mathbb{E}\left[(V_t^n - U_t)^2\right] \, \mathrm{d}t \leqslant \frac{\varepsilon}{2},$$

Then, for m big enough, by **(i)** :

$$\int_T \mathbb{E}\left[(P_m V_t^n - V_t^n)^2\right] \, \mathrm{d}t \leqslant \frac{\varepsilon}{2}.$$

Hence, for those m, we have :

$$\int_T \mathbb{E}\left[(P_m V_t^n - U_t)^2\right] \, \mathrm{d}t \leqslant \varepsilon,$$

and we succeed to approach U by elementary processes.

(iii) Finally, for a general $U \in L^2_{\mathsf{a}}(T \times \Omega)$, we set $U_M(t) \stackrel{\text{def.}}{=} U_t \mathbf{1}_{\{|U_t| \leqslant M\}}$. U_M is bounded by M, and belongs to L^2_{a}. By **(ii)**, we can approach U_M by elements of $\mathcal{E}_0$. Moreover,

$$|U_M(t)| \leqslant |U_t|,$$

and $U \in L^2_{\mathsf{a}}$. By dominated convergence theorem,

$$\int_T \mathbb{E}\left[|U_M(t) - U_t|^2\right] \, \mathrm{d}t \xrightarrow[M \to +\infty]{} 0.$$

(iv) Let $U \in L^2_{\mathsf{a}}$. Then, by Cauchy-Schwarz inequality,

$$\int_T \mathbb{E}\left[|P_n u_t|^2\right] \, \mathrm{d}t \leqslant \sum_{i=1}^{2^n} \mathbb{E}\left[\int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} U_s^2 \, \mathrm{d}s\right].$$

Meaning by Chasles relation that

$$\int_T \mathbb{E}\left[|P_n u_t|^2\right] \, \mathrm{d}t \leqslant \int_T \mathbb{E}\left[U_s^2\right] \, \mathrm{d}s.$$

Hence, P_n is continuous. This proves that $P_n U$ approaches U. Indeed, if a sequence $(U^m)_m$ converges to U in L^2_{a}, and such that $(P_n U^m)_n$ converges to U^m in L^2_{a}, like in points **(ii)** and **(iii)**, then for m big enough

$$\int_T \mathbb{E}\left[(U^m(t) - U(t))^2\right] \, \mathrm{d}t \leqslant \frac{\varepsilon}{9}.$$

We set such m. We then have (the factors 3 coming from the convexity of the square) :

$$\begin{aligned} & \int_T \mathbb{E}\left[(P_n U(t) - U(t))^2\right] \, \mathrm{d}t \\ \leqslant \quad & 3 \int_T \mathbb{E}\left[(P_n U(t) - P_n U^m(t))^2\right] \, \mathrm{d}t \\ + \quad & 3 \int_T \mathbb{E}\left[(P_n U^m(t) - U^m(t))^2\right] \, \mathrm{d}t \\ + \quad & 3 \int_T \mathbb{E}\left[(U^m(t) - U(t))^2\right] \, \mathrm{d}t \end{aligned}$$

There exists $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$, the two first members are lower or equal to $3\frac{\varepsilon}{9}$, the first by continuity of P_n, the second by convergence of $P_n U^m$ to U^m. We achieved the proof. □

For the sequel of this subsection, we set T a compact interval of $\mathbb{R}_+$.

Theorem IV.2 : Itô's integral

The map

$$\begin{pmatrix} \mathcal{E}_0 & \longrightarrow & L^2(\mathbb{P}) \\ U & \longmapsto & \int_T U_t \, \mathrm{d}B_t \end{pmatrix}$$

can be extended on $L^2_{\mathsf{a}}(T \times \Omega)$. Moreover, we still have the isometry property for every $u \in L^2_{\mathsf{a}}(T \times \Omega)$.

Proof : This is a consequence of the extension of linear continuous maps theorem, using the fact that L^2_{a} is exactly the closure of $\mathcal{E}_0$ by the previous lemma. □

Then, we know how to define the stochastic Itô integral of adapted process. Let us see an application of this, using the approximation P_n.

Proposition IV.5 : Local property

Let $U \in L^2_{\mathsf{a}}(T \times \Omega)$. If we set

$$G \stackrel{\text{def.}}{=} \left\{\int_T u_t^2 \, \mathrm{d}t = 0\right\} \in \mathcal{F},$$

then

$$\mathbb{P}\left(G \cap \left\{\int_T u_t \, \mathrm{d}B_t = 0\right\}^{\mathsf{c}}\right) = 0.$$

In other words, in G, we have $\int_T u_t \, \mathrm{d}B_t = 0$.

Proof : Once again, we set $T = [0, 1]$. Here's the plan.

(i) We prove it for every $P_n U$ introduced in the lemma, by showing that in G, $P_n U = 0$ almost everywhere in T.

(ii) We conclude by using Itô's isometry and letting $[n \to +\infty]$.

Here we go.

(i) We set on G. The idea is to prove that

$$\int_T (P_n u_t)^2 \, \mathrm{d}t = 0$$

We have :

$$\int_T (P_n u_t)^2 \, \mathrm{d}t = 4^n \sum_{i=1}^{2^n} \left(\int_T \mathbf{1}_{\left]\frac{i}{2^n}, \frac{i+1}{2^n}\right]}(t) \, \mathrm{d}t \right) \left(\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} u_s \, \mathrm{d}s \right)^2.$$

One 2^n simplifies, and we get :

$$\int_T (P_n u_t)^2 \, \mathrm{d}t = 2^n \sum_{i=1}^{2^n} \left(\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} u_s \, \mathrm{d}s \right)^2.$$

By Cauchy-Schwarz inequality, the other 2^n simplifies and :

$$\int_T (P_n u_t)^2 \, \mathrm{d}t \leqslant \sum_{i=1}^{2^n} \int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} u_s^2 \, \mathrm{d}s.$$

By Chasles relation :

$$\int_T (P_n u_t)^2 \, \mathrm{d}t \leqslant \int_T u_s^2 \, \mathrm{d}s = 0.$$

Hence, $P_n u = 0$ almost everywhere in T, and so

$$\int_T P_n u_t \, \mathrm{d}B_t = 0.$$

(ii) By linearity,

$$\begin{aligned} & \mathbb{E}\left[\left| \int_T P_n u_t \, \mathrm{d}B_t - \int_T u_t \, \mathrm{d}B_t \right|^2 \right] \\ = \quad & \mathbb{E}\left[\left| \int_T (P_n u_t - u_t) \, \mathrm{d}B_t \right|^2 \right]. \end{aligned}$$

By Itô's isometry :

$$\begin{aligned} & \mathbb{E}\left[\left| \int_T P_n u_t \, \mathrm{d}B_t - \int_T u_t \, \mathrm{d}B_t \right|^2 \right] \\ = \quad & \mathbb{E}\left[\int_T (P_n u_t - u_t)^2 \, \mathrm{d}t \right]. \end{aligned}$$

By the lemma, the second term goes to zero when $[n \to +\infty]$. Hence,

$$\int_T P_n u_t \, \mathrm{d}B_t \xrightarrow[n \to +\infty]{L^2} \int_T u_t \, \mathrm{d}B_t,$$

so converges up to extraction $\mathbb{P}$-almost surely. Since in G, we have $\mathbb{P}$-almost surely, for all $n \in \mathbb{N}$:

$$\int_T P_n u_t \, \mathrm{d}B_t = 0,$$

it means by almost sure convergence that on G, we have almost surely :

$$\int_T u_t \, \mathrm{d}B_t = 0,$$

which proves our local property. □

Let us recall the Itô formula associated.

Theorem IV.3 : Itô formula

Let $F \in \mathcal{C}^2(\mathbb{R})$, and M a continuous semi-martingale. Then, $F(M)$ is a semi-martingale whose decomposition is given by

$$F(M_t) = F(M_0) + \int_0^t F'(M_s) \, \mathrm{d}M_s + \frac{1}{2} \int_0^t F''(M_s) \, \mathrm{d}\langle M, M \rangle_s.$$

Notice that the term with F'' is only due to stochastic calculus, and doesn't appear in deterministic calculus.

Theorem IV.4 : Itô's integral representation theorem

Let $F \in L^2$. Then, there exists an unique (up to indistinguishability) process $u \in L^2_{\mathsf{a}}(T \times \Omega)$ such that

$$F = \mathbb{E}[F] + \int_T u_t \, \mathrm{d}B_t.$$

To prove it, we introduce the stochastic exponential. This stochastic exponential is defined in a more general way that the one we present here.

Lemma IV.8 : Stochastic exponential

Let $u \in L^2_{\mathsf{a}}(T \times \Omega)$. There exists an unique stochastic process (up to indistinguishability) $\mathcal{E}(u)$ such that $\mathcal{E}(u)_0 = 1$ and

$$\mathrm{d}\mathcal{E}(u)_t = u_t \mathcal{E}(u)_t \, \mathrm{d}B_t.$$

We call it the stochastic exponential of u. It is given by :

$$\mathcal{E}(u) = \exp\left(\int_0^t u_s \, \mathrm{d}B_s - \frac{1}{2} \int_0^t u_s^2 \, \mathrm{d}s \right).$$

Proof of the lemma : If we suppose that X satisfies the EDS, and that X is almost surely strictly positive, then if we apply the Itô formula for $F = \ln$ on $\mathbb{R}_+^*$:

$$\ln X_t = \int_0^t \frac{\mathrm{d}X_s}{X_s} - \frac{1}{2}\int_0^t \frac{\mathrm{d}\langle X, X\rangle_s}{X_s^2}.$$

Since X_s satisfies the EDS, we get the expression of $\mathrm{d}X_s$ and we have

$$\mathrm{d}\langle X, X\rangle_s = u_s^2 X_s^2 \ \mathrm{d}s.$$

Finally, we have

$$\ln X_t = \int_0^t u_s \ \mathrm{d}B_s - \frac{1}{2}\int_0^t u_s^2 \ \mathrm{d}s.$$

□

We admit the following lemma.

Lemma IV.9 : Stochastic exponential and martingales

Let $h \in L^2(T)$. Then $\mathcal{E}(h)$ is a square integrable martingale.

Finally, we introduce a third lemma we could use way much earlier in the introduction of the Wiener chaos. We express it in a general way.

Lemma IV.10 : Density of exponentials

Let $\mathcal{H}$ a real separable Hilbert space, and X an isonormal process on $\mathcal{H}$. Then, the set $\left\{e^{X(h)}, h \in \mathcal{H}\right\}$ is dense in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is the σ-algebra generated by X.

Proof of the lemma : Let $G \in L^2$ such that

$$\forall h \in \mathcal{H}, \mathbb{E}\left[G e^{X(h)}\right] = 0.$$

Then by linearity of X, it means that for every integer $n \in \mathbb{N}^*$, $t_1, \cdots, t_n \in \mathbb{R}$ and $h_1, \cdots, h_n \in \mathcal{H}$:

$$\mathbb{E}\left[G \exp\left(\sum_{k=1}^n t_k X(h_k)\right)\right] = 0.$$

If we define the signed measure on $\mathcal{B}(\mathbb{R}^n)$:

$$\nu(B) \overset{\text{def.}}{=} \mathbb{E}\left[G \mathbf{1}_B(X(h_1), \cdots, X(h_n))\right],$$

then we can compute its Laplace transform. Indeed, if F is a step function, we have :

$$\int_{\mathbb{R}^n} F \ \mathrm{d}\nu = \mathbb{E}\left[G F(X(h_1), \cdots, X(h_n))\right].$$

Since the exponential function can be uniformly approached by step functions on $\mathbb{R}^d$, and since ν is finite, we can conclude that we have :

$$\mathcal{L}\nu(t_1, \cdots, t_n) = \mathbb{E}\left[G \exp\left(\sum_{k=1}^n t_k X(h_k)\right)\right] = 0.$$

Meaning that ν is the null measure, by injectivity of the Laplace transform for finite measure. Hence, G is orthogonal to $\mathcal{F}$, and so $G = 0$, since G is $\mathcal{F}$-measurable. □

We can consequently prove the Itô's representation theorem.

Proof : We set $T = [0, 1]$. Let $G \in L_0^2$ (L^2 and centered) such that

$$\forall u \in L_{\mathsf{a}}^2, \mathbb{E}\left[G \int_T u_t \ \mathrm{d}B_t\right] = 0.$$

- If we prove that $G = 0$, then, since L_0^2 is closed in L^2, and since the set

$$\mathcal{U} \overset{\text{def.}}{=} \left\{\int_T u_t \ \mathrm{d}B_t, u \in L_{\mathsf{a}}^2(\Omega \times T)\right\}$$

in included in L_0^2, this would implies that the reciprocal inclusion is also true : $L_0^2 = \mathcal{U}$.
- Since $\mathcal{E}(u)u \in L_{\mathsf{a}}^2$, we have in fact

$$\forall u \in L_{\mathsf{a}}^2, \mathbb{E}\left[G \int_T u_t \mathcal{E}(u)_t \ \mathrm{d}B_t\right] = 0.$$

But, by the previous lemma about stochastic exponential, this implies that

$$\forall u \in L_{\mathsf{a}}^2, \mathbb{E}\left[G \mathcal{E}(u)_1\right] = 0.$$

By the expression of the stochastic exponential, we get for all $h \in L^2(T)$ (every deterministic process is adapted) :

$$\forall u \in L_{\mathsf{a}}^2, \mathbb{E}\left[G \exp\left(\int_T h_t \ \mathrm{d}B_t\right)\right] \exp\left(\frac{-1}{2}\int_T h_t^2 \ \mathrm{d}t\right) = 0.$$

Finally, since $X(h) = \int_T h_t \ \mathrm{d}B_t$, we have :

$$G \in \left\{e^{X(h)}, h \in L^2(T)\right\}^{\perp}.$$

This set is dense in L^2, by the previous lemma, so we proved that $G = 0$. □

Corollary IV.2 : Conditional expectation of the Itô representation

Let $F \in L^2$ and $u \in L^2(T \times \Omega)$ like previously. Then, for all $t \in T$,

$$\mathbb{E}[F|\mathcal{F}_t] = \mathbb{E}[F] + \int_0^t u_s \ \mathrm{d}B_s.$$

IV.4.2 Skorohod integral as extension of Ito's integration

Let $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$ like previously. We make the identification $L^2(\Omega \to \mathcal{H}) \simeq L^2(\Omega \times T)$. We will simply call the *Skorohod integral of* u the divergence of u : $\delta(u)$. We want to use the following notation :

$$\delta(u) = \int_T u_t \, \mathrm{d}W_t.$$

Until now, we did it in two cases : one with the Wiener integral for elements of $L^2(T)$, and we did it with Itô with elements of $L^2_{\mathrm{a}}(\Omega \times T)$, where here $T \subset \mathbb{R}_+$ is an interval. We want to do this for elements of $D(\delta) \subset L^2(\Omega \times T)$, with a general T.

Proposition IV.6 : Extension of Itô integral

Let $T \subset \mathbb{R}_+$ a compact interval. Then, we have

$$L^2_{\mathrm{a}}(T \times \Omega) \subset D(\delta),$$

and we have for every $u \in L^2_{\mathrm{a}}$:

$$\delta(u) = \int_T u_t \, \mathrm{d}B_t.$$

Proof : • Let us show first that $L^2_{\mathrm{a}} \subset D(\delta)$. Since $D(\delta)$ is closed for the norm L^2_{a}, it is enough to show that $\mathcal{E}_0 \subset D(\delta)$. By linearity, it is enough to prove that if $A \in \mathcal{F}$, with $\mu(A) < +\infty$ (μ is the Lebesgue measure here), and F is A-measurable, then $F\mathbf{1}_{A^c} \in D(\delta)$. First, if $F \in \mathbb{D}^{1,2}$, then we have already shown that $F\mathbf{1}_A \in D(\delta)$, by computation with homogeneity up to a random constant, and we have (the two expectations we have to check that they are finite are indeed finite since $u = \mathbf{1}_A$ is deterministic) :

$$\delta\left(F\mathbf{1}_A\right) = F\delta(\mathbf{1}_A) + \langle \mathrm{D}F, \mathbf{1}_{A^c}\rangle_{\mathcal{H}}.$$

Since F is $\mathcal{F}_A$-measurable, $\mathrm{D}F$ is almost everywhere equal to zero on A, so the bracket is null. This proves the result on $\mathbb{D}^{1,2}$, so on $\mathcal{S}$.

For $F \in L^2(\mathbb{P})$, we have by density of $\mathcal{S}$, the existence of a sequence $(F_n)_n \in \mathcal{S}^{\mathbb{N}}$ such that

$$\mathbb{E}\left[|F_n - F|^2\right] \xrightarrow[n\to+\infty]{} 0.$$

Moreover, we have by the computation on elements of $\mathbb{D}^{1,2}$:

$$\delta\left(F_n\mathbf{1}_{A^c}\right) \xrightarrow[n\to+\infty]{L^2(\mathbb{P})} F\delta(\mathbf{1}_A) + \langle \mathrm{D}F, \mathbf{1}_{A^c}\rangle_{\mathcal{H}}.$$

By the proposition about a limit of elements of $D(\delta)$ belonging or not in $D(\delta)$, we can conclude that $F\mathbf{1}_{A^c} \in D(\delta)$, with :

$$\delta\left(F\mathbf{1}_{A^c}\right) = F\delta(\mathbf{1}_A) + \langle \mathrm{D}F, \mathbf{1}_{A^c}\rangle_{\mathcal{H}}.$$

This proves that $\mathcal{E}_0 \subset D(\delta)$, and since $D(\delta)$ is closed for $\mathbb{E}\left[\int_T \cdot^2 \, \mathrm{d}t\right]$, it proves that $L^2_{\mathrm{a}} \subset D(\delta)$.

• If we compute $\delta(u)$ for $u \in \mathcal{E}_0$:

$$u = \sum_{j=1}^n Z_j \mathbf{1}_{]t_j, t_{j+1}]},$$

with Z_j which is $\mathcal{F}_{t_j}$-measurable. We have $u \in D(\delta)$. By homogeneity up to a random constant :

$$\delta\left(F_j \mathbf{1}_{]t_j,t_{j+1}]}\right) = F_j\delta\left(\mathbf{1}_{]t_j,t_{j+1}]}\right) - \int_T \mathrm{D}_t F_j \mathbf{1}_{]t_j,t_{j+1}]}(t) \, \mathrm{d}t.$$

But, since F_j is $\mathcal{F}_{t_j}$-measurable, $\mathrm{D}F = 0$ almost everywhere on $]t_j, t_{j+1}]$. And by computation on simple elements :

$$\delta\left(\mathbf{1}_{]t_j,t_{j+1}]}\right) = X\left(\mathbf{1}_{]t_j,t_{j+1}]}\right) = B_{t_{j+1}} - B_{t_j}.$$

By linearity, we get the result for every process $u \in \mathcal{E}_0$.

• Let $u \in L^2_{\mathrm{a}}$. Then, using the operator P_n defined previously, and by Itô's isometry :

$$\mathbb{E}\left[\left|\int_T (u_t - P_n u_t) \, \mathrm{d}B_t\right|^2\right] = \|u - P_n u\|_{L^2_{\mathrm{a}}},$$

which goes to zero when $[n \to +\infty]$. Moreover, this also means by **(i)** that

$$\delta(P_n u) \xrightarrow[n\to+\infty]{L^2} \int_T u_t \, \mathrm{d}B_t.$$

Since δ is closed, this finally means that :

$$\delta(u) = \int_T u_t \, \mathrm{d}B_t.$$

δ coincides with the Itô integral. □

Notice that we note "$\mathrm{d}B_t$" to show that this integral is in the Itô's way. Since the Wiener integral and the Itô integral coincides, and since

$$\forall h \in \mathcal{H}, \delta(h) = X(h) = \int_T h_t \, \mathrm{d}W_t,$$

we will note $\delta(u) = \int_T u_t \, \mathrm{d}W_t$, and not $\mathrm{d}B_t$ for general elements of $D(\delta)$.

Definition IV.7

Let $u \in L^2(\Omega \times T)$. We say that u is *Skorohod integrable* if $u \in D(\delta)$. In this case, we note

$$\int_T u_t \, \mathrm{d}W_t \overset{\text{def.}}{=} \delta(u).$$

Proposition IV.7 : Clark-Ocone representation formula

Let $F \in \mathbb{D}^{1,2}$, T an interval of $\mathbb{R}_+$, an B a $(\mathcal{F}_t)_t$-Brownian motion. Then,

$$F = \mathbb{E}[F] + \int_T \mathbb{E}\left[\mathrm{D}_t F | \mathcal{F}_t\right] \, \mathrm{d}B_t.$$

Proof : Two proves are possible, let us make both.

1. We know by Itô's representation theorem that there is an unique (in L^2) $u \in L^2_{\mathrm{a}}$ such that

$$F = \mathbb{E}[F] + \int_T u_t \, \mathrm{d}W_t.$$

We just need to show that $u_t = \mathbb{E}[\mathrm{D}_t F | \mathcal{F}_t]$. To do this, we show that the difference belongs to $L^2_{\mathrm{a}} \cap L^{2\perp}_{\mathrm{a}}$. Let $v \in L^2_{\mathrm{a}}$. Then, $v \in D(\delta)$. By definition of δ by duality, we have

$$\mathbb{E}[F\delta(v)] = \int_T \mathbb{E}[v_t \mathrm{D}_t F] \, \mathrm{d}t.$$

But, since $v \in L^2_{\mathrm{a}}$, since $\mathbb{E}[\delta(v)]$, we also have :

$$\mathbb{E}[F\delta(v)] = \mathbb{E}\left[\left(\int_T v_t \, \mathrm{d}B_t\right)\left(\int_T u_t \, \mathrm{d}B_t\right)\right].$$

By Itô's isometry, we have finally :

$$\int_T \mathbb{E}[v_t \mathrm{D}_t F] \, \mathrm{d}t = \int_T \mathbb{E}[v_t u_t] \, \mathrm{d}B_t.$$

But, since v_t is $\mathcal{F}_t$-measurable, we can write it as :

$$\int_T \mathbb{E}[v_t \mathbb{E}[\mathrm{D}_t F | \mathcal{F}_t]] \, \mathrm{d}t = \int_T \mathbb{E}[v_t u_t] \, \mathrm{d}B_t.$$

And since $\mathbb{E}[\mathrm{D}_\bullet F | \mathcal{F}_\bullet]] \in L^2_{\mathrm{a}}(T \times \Omega)$, this means that we proved that the difference with u is in $L^{2\perp}_{\mathrm{a}}$, so is zero. This concludes the proof.

2. We expend F in Wiener chaos, and compute the expansion of every member of the expected equality :

$$F = \sum_{n=0}^{+\infty} I_n(f_n),$$

with $I_0(f_0) = \mathbb{E}[F]$, and $(f_n)_n$ symmetric. Then, we have

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

And by the lemma **III.4** computing the Wiener chaos expansion of a conditional expectation, we get :

$$\mathbb{E}[\mathrm{D}_t F | \mathcal{F}_t] = \sum_{n=1}^{+\infty} n I_{n-1}\left(f_n(\cdot, t) \mathbf{1}_{[0,t]}^{\otimes(n-1)}\right).$$

Since $(\mathbb{E}[\mathrm{D}_t F | \mathcal{F}_t])_{t \geqslant 0} \in L^2_{\mathrm{a}}$, we can compute the divergence of it in terms of Itô's integral :

$$\delta\left((\mathbb{E}\left[\mathrm{D}_t F | \mathcal{F}_t\right])_{t \geqslant 0}\right) = \int_T \mathbb{E}\left[\mathrm{D}_t F | \mathcal{F}_t\right] \, \mathrm{d}B_t.$$

And since we have the Wiener chaos expansion of $\mathbb{E}[\mathrm{D}_t F | \mathcal{F}_t]$, we can compute the one of its divergence. Let us call

$$g_n(t_1, \cdots, t_n, t) \overset{\text{def.}}{=} (n+1) f_{n+1}(t_1, \dots, t_n, t) \mathbf{1}_{[0,t]}^{\otimes n}(t_1, \cdots, t_n).$$

Then, if we compute its symmetrization, we have :

$$\begin{aligned} &\tilde{g}_n(t_1, \cdots, t_n, t) = \\ = \; &f_{n+1}(t_1, \dots, t_n, t) \\ &\cdot \left(\mathbf{1}_{[0,t]}^{\otimes n}(t_1, \cdots, t_n) + \sum_{k=1}^{n} \mathbf{1}_{[0,t_k]}^{\otimes n}(t_1, \cdots, t, \cdots t_n)\right). \end{aligned}$$

But, the sum in parenthesis is equal almost surely in T^{n+1} to 1. Indeed, if we note $t_{n+1} = t$, then there exists k_0 such that t_{k_0} is the biggest of the t_k. Then, the term in $k = k_0$ is equal to one, and the others are equal to zero unless t_{k_0} is not unique, but the set of $\mathbb{R}^{n+1}$ such that at least two coordinates are equal is negligible. Hence, we have almost every where $\tilde{g}_n = f_{n+1}$ and we get exactly :

$$\delta\left((\mathbb{E}\left[\mathrm{D}_t F | \mathcal{F}_t\right])_{t \geqslant 0}\right) = \sum_{n=1}^{+\infty} I_n(f_n),$$

which is $F - \mathbb{E}[F]$. □

V Ornstein-Ulhenbeck operator

Definition V.1

We define $(P_t)_t$ as the following : for all $F \in L^2(\mathbb{P})$ and $t \geqslant 0$:

$$P_t F \overset{\text{def.}}{=} \sum_{p=0}^{+\infty} e^{-pt} J_p F.$$

V.1 Melher's formula

Hence, $P_t : L^2(\mathbb{P}) \longrightarrow L^2(\mathbb{P})$. We could define a closer definition of this operator from the one-dimensional case.

Proposition V.1 : Melher's formula

Let X' an another Gaussian isonormal process with values in $\mathcal{H}$, with values in $L^2(\Omega', \mathcal{F}', \mathbb{P}')$. Then, for all $F \in L^2(\mathbb{P})$, there exists a measurable map $\psi_F : \mathbb{R}^{\mathcal{H}} \longrightarrow \mathbb{R}$ such that $\psi_F(X) = F$ and

$$P_t F = \mathbb{E}'\left[\psi_F\left(e^{-t}X + \sqrt{1-e^{-2t}}X'\right)\right].$$

Proposition V.2 : Contraction

The operator $P_t : L^q \longrightarrow L^q$ is a contraction :

$$\forall F \in L^q, \mathbb{E}\left[|P_t F|^q\right] \leqslant \mathbb{E}\left[|F|^q\right].$$

Proof : We use Jensen inequality, like in section I.

$$\mathbb{E}[|P_t F|^q] \leqslant \mathbb{E}\left[\mathbb{E}'\left[\left|\psi_F\left(e^{-t}X + \sqrt{1-e^{-2t}}X'\right)\right|^q\right]\right].$$

Since $e^{-t}X + \sqrt{1-e^{-2t}}X'$ as the same law (as processes, so if we look at finite dimensional laws), we have

$$\mathbb{E}[|P_t F|^q] \leqslant \mathbb{E}\left[\mathbb{E}'\left[|\psi_F(X)|^q\right]\right].$$

That is, since $\psi_F(X) = F$:

$$\mathbb{E}[|P_t F|^q] \leqslant \mathbb{E}\left[|F|^q\right].$$

□

V.2 Infinitesimal generator

Definition V.2

We define the unbounded operator $(L, D(L))$ as follows :

$$D(L) \overset{\text{def.}}{=} \left\{F \in L^2(\mathbb{P}), \sum_{n=0}^{+\infty} n^2 \mathbb{E}[(J_n F)^2] < +\infty\right\},$$

and for all $F \in D(L)$, we define on L^2 :

$$LF \overset{\text{def.}}{=} -\sum_{n=0}^{+\infty} n J_n F.$$

Proposition V.3 : Generator of the Ornstein-Uhlebeck semi-group

The operator $(L, D(L))$ is the infinitesimal generator of $(P_t)_t$, that is :

$$D(L) = \left\{F \in L^2, \exists G \in L^2, \mathbb{E}\left[\left|\frac{P_t F - F}{t} - G\right|^2\right] \xrightarrow[t \to 0^+]{} 0\right\},$$

and for all $F \in D(L)$, we have in L^2 :

$$LF = \lim_{t \to 0} \frac{P_t F - F}{t}.$$

Proof : [$\Longrightarrow$] Let $F \in L^2$ such that $\sum_n n^2 \mathbb{E}[J_n F^2]$ converges. Let us show that

$$\mathbb{E}\left[\left|\frac{P_t F - F}{t} - LF\right|^2\right] \xrightarrow[t \to 0]{} 0.$$

By the definitions with the series, and by expanding F in Wiener chaos, we have :

$$\mathbb{E}\left[\left|\frac{P_t F - F}{t} - LF\right|^2\right] = \sum_{p=0}^{+\infty} \left|\frac{e^{-pt} - 1}{t} + p\right|^2 \mathbb{E}[J_p F^2].$$

But, we have the following domination, for all $p \in \mathbb{N}$ and $t \geqslant 0$:

$$\left|\frac{e^{-pt}-1}{t}+p\right|^2 \mathbb{E}[J_pF^2] \leqslant 4p^2\mathbb{E}[J_pF^2].$$

By dominated convergence, we have the expected convergence, and so we proved the inclusion between the definition of L, and the generator of P_t.

[$\Longleftarrow$] Conversely, we suppose that $F \in L^2$ satisfies that there exists $G \in L^2$ such that

$$\mathbb{E}\left[\left|\frac{P_tF-F}{t}-G\right|^2\right] \xrightarrow[t\to 0^+]{} 0.$$

We show that $\sum_n n^2\mathbb{E}[J_nF^2]$ converges and that $G = LF$, in the sense of the definition. First, by continuity of J_n on L^2, and linearity, we have in L^2 :

$$J_nG = \lim_{t\to 0^+} \frac{J_nP_tF - J_nF}{t}.$$

By definition of P_t and unicity of Wiener chaos expansion, we have

$$J_nP_tF = e^{-nt}J_nF (= P_tJ_nF).$$

Consequently, we have :

$$J_nG = \lim_{t\to 0^+} \frac{e^{-nt}-1}{t} J_nF.$$

So :

$$J_nG = -nJ_nF.$$

We deduce that

$$\sum_{n=0}^{+\infty} n^2\mathbb{E}[J_nF^2] = \sum_{n=0}^{+\infty} \mathbb{E}[J_nG^2] = \mathbb{E}[G^2] < +\infty.$$

Moreover, in L^2 :

$$LF = -\sum_{p=0}^{+\infty} pJ_pF = \sum_{p=0}^{+\infty} J_pG = G.$$

We have proven the other inclusion. □

Corollary V.1 : Derivative of P_t

Let $F \in D(L)$. Then for all $t \geqslant 0$:

$$\frac{\mathrm{d}P_tF}{\mathrm{d}t} = LP_tF = P_tLF.$$

We will show an important relation between the last operators we introduced. We have already seen it in the one-dimensional case, and saw some applications.

Theorem V.1 : Relation between L, δ and D

We set $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$. Let $F \in L^2$. Then, $F \in D(L)$ if and only if $F \in \mathbb{D}^{1,2}$ and $\mathrm{D}F \in D(\delta)$. In this case,

$$\delta \mathrm{D}F = -LF.$$

The theorem is still true for real separable Hilbert space $\mathcal{H}$, the choice of $L^2(\mu)$ is to complete the proof.

Proof : [$\Longrightarrow$] Let us suppose that $F \in D(L)$, meaning that $\sum_n n^2\mathbb{E}[J_nF^2]$ converges. Then, the series $\sum_n n\mathbb{E}[J_nF^2]$ converges too, and so $F \in \mathbb{D}^{1,2}$, by the series characterization. Let us show that $\mathrm{D}F \in D(\delta)$, by using this time the definition. Let $G \in \mathbb{D}^{1,2}$. Then by using the expansion of $\mathrm{D}F$ and $\mathrm{D}G$ in Wiener chaos,

$$\mathbb{E}\left[\langle \mathrm{D}G, \mathrm{D}F\rangle\right] = \sum_{n=1}^{+\infty} n^2 \int_T \langle \tilde{g}_n(\cdot,t), \tilde{f}_n(\cdot,t)\rangle_{L^2(T^{n-1})}\, \mathrm{d}\mu(t).$$

Hence,

$$\mathbb{E}\left[\langle \mathrm{D}G, \mathrm{D}F\rangle\right] = \sum_{n=1}^{+\infty} n^2 \langle \tilde{g}_n, \tilde{f}_n\rangle_{L^2(T^n)}.$$

It yields to

$$\mathbb{E}\left[\langle \mathrm{D}G, \mathrm{D}F\rangle\right] = \sum_{n=1}^{+\infty} n\mathbb{E}\left[J_nFJ_nG\right].$$

We conclude by Cauchy-Schwarz inequality that :

$$\left|\mathbb{E}\left[\langle \mathrm{D}G, \mathrm{D}F\rangle\right]\right| \leqslant \left(\sum_{n=1}^{+\infty} n^2\mathbb{E}\left[J_nF^2\right]\right)^{\frac{1}{2}} \mathbb{E}\left[G^2\right]^{\frac{1}{2}}.$$

And so, $\mathrm{D}F \in D(\delta)$.

[$\Longleftarrow$] We suppose that $F \in \mathbb{D}^{1,2}$ and $\mathrm{D}F \in D(\delta)$. Let us show that $J_n(\delta \mathrm{D}F) = nJ_nF$ for every $n \in \mathbb{N}$. To do that, we consider $G \in \mathfrak{H}_n$. Then,

$$\mathbb{E}[G\ \delta \mathrm{D}F] = \mathbb{E}[\langle \mathrm{D}G, \mathrm{D}F\rangle].$$

By the lemma **IV.5** about projection on Wiener chaos of the divergence, since the Wiener chaos expansion of D_tF is

$$\mathrm{D}_tF = \sum_{n=1}^{+\infty} nI_{n-1}(f_n(\cdot,t)),$$

we have :

$$\mathbb{E}[G\ \delta \mathrm{D}F] = \mathbb{E}[GI_n(f_n)].$$

And so, we have $J_n(\delta \mathrm{D}F) = nJ_nF$. It means that :

$$\sum_{n=0}^{+\infty} n^2\mathbb{E}[J_nF^2] = \mathbb{E}[(\delta \mathrm{D}F)^2] < +\infty,$$

so, that $F \in D(L)$ and moreover that

$$\delta \mathrm{D}F = \sum_{n=0}^{+\infty} nJ_nF = -LF.$$

□

Proposition V.4 : An expression of L for smooth random variables

Let $F \in \mathcal{S}$, given by $F = f(X(h_1), \cdots, f(X(h_n))$. Then, $F \in D(L)$ and :

$$LF = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X(h_1), \cdots, X(h_n))\langle h_i, h_j\rangle - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))X(h_i).$$

Proof : We use the expression we just proved. For $F \in \mathcal{S}$, the derivative is given by :

$$\mathrm{D}F = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))h_i.$$

We just need to compute the divergence of it. We use the computation of $\delta(Fu)$ done in a previous proposition. It gives here

$$\begin{aligned} &\delta\left(\frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))h_i\right) \\ =\ &\frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n))\delta(h_i) \\ &-\left\langle \mathrm{D}\frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n)), h_i\right\rangle_{\mathcal{H}}. \end{aligned}$$

Since $\delta(h_i) = X(h_i)$ (computation on simple elements), and since

$$\mathrm{D}\frac{\partial f}{\partial x_i}(X(h_1), \cdots, X(h_n)) = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(X(h_1), \cdots, X(h_n))h_j,$$

we conclude on the equality by linearity. □

V.3 Pseudo inverse and integration by parts formula

Definition V.3

Let $F \in L^2(\mathbb{P})$. We define the *pseudo inverse* of the operator L by :

$$L^{-1}F \stackrel{\text{def.}}{=} -\sum_{p=1}^{+\infty} \frac{1}{p} J_p F.$$

Proposition V.5 : PSEUDO inverse

Let $F \in L^2(\mathbb{P})$. Then $L^{-1}F \in D(L)$ and

$$LL^{-1}F = F - \mathbb{E}[F].$$

Proof : By definition,

$$J_p L^{-1}F = \frac{-1}{p} J_p F.$$

Hence,

$$p^2 \mathbb{E}\left[J_p L^{-1} F^2\right] = \mathbb{E}[J_p F^2]$$

defines a general term of a convergent series, so $L^{-1}F \in D(L)$. Moreover,

$$L(L^{-1}F) = -\sum_{p=0}^{+\infty} pJ_p(L^{-1}F) = \sum_{p=1}^{+\infty} J_p F.$$

So $LL^{-1}F = F - \mathbb{E}[F]$. □

We derive here an another integration by parts formula, which is not the same as the previous with the divergence.

Proposition V.6 : Integration by parts formula with the pseudo inverse

Let $F, G \in \mathbb{D}^{1,2}$ and $g \in \mathcal{C}_b^1(\mathbb{R})$. Then

$$\mathbb{E}[Fg(G)] = \mathbb{E}[F]\mathbb{E}[g(G)] - \mathbb{E}\left[g'(G)\langle \mathrm{D}G, \mathrm{D}L^{-1}F\rangle_{\mathcal{H}}\right].$$

Proof : We use the previous property. We have

$$\mathbb{E}\left[(F - \mathbb{E}[F])g(G)\right] = \mathbb{E}[LL^{-1}F \cdot g(G)].$$

We use $L = -\delta\mathrm{D}$:

$$\mathbb{E}\left[(F - \mathbb{E}[F])g(G)\right] = -\mathbb{E}[\delta\mathrm{D}(L^{-1}F) \cdot g(G)].$$

By duality :

$$\mathbb{E}\left[(F - \mathbb{E}[F])g(G)\right] = -\mathbb{E}\left[\langle \mathrm{D}(L^{-1}F), \mathrm{D}g(G)\rangle_{\mathcal{H}}\right].$$

Finally, by chain rule on C_b^1 functions :

$$\mathbb{E}\left[(F-\mathbb{E}[F])g(G)\right] = -\mathbb{E}\left[g'(G)\langle \mathrm{D}(L^{-1}F), \mathrm{D}G\rangle_{\mathcal{H}}\right].$$

This concludes the proof. □

As an application, we will prove the following theorem about the law of a random variable living in a set Wiener chaos. We prove it for $\mathcal{H} = L^2(T)$.

Theorem V.2 : Absolute continuity with respect to Lebesgue measure in Wiener chaos

Let $q \in \mathbb{N}^*$, $f \in L^2_{\mathrm{S}}(T^q)$ and $F = I_q(f)$. If $\|f\|_{L^2(T^q)} > 0$, then F admits a law which is absolutely continuous with respect to the Lebesgue measure.

Proof : We proceed by induction on q. For $q = 1$, $I_1(f) \sim \mathcal{N}\left(0, \|f\|_{L^2}\right)$, so the initialisation is true. Let $q \in \mathbb{N}^*$. We suppose that for every $g \in L^2_{\mathrm{S}}(T^q)$ non equal to zero almost surely, the random variable $I_q(g)$ admits a density with respect to the Lebesgue measure. We want to prove it for the rank $q+1$. Let $F = I_{q+1}(f)$, with $f \in L^2_{\mathrm{S}}(T^{q+1})$ non equal to zero. The idea is to use $\mathrm{D}F$ which belongs to the previous chaos to use the induction hypothesis. Here's the plan.

1. Let $t_1, \cdots, t_q \in T$ and

$$g = g_{t_1,\cdots,t_q} \overset{\text{def.}}{=} f(t_1, \cdot, t_q, \cdot) \in L^2_{\mathrm{S}}(T).$$

Then there exists $h \in \mathcal{H}$ such that $\langle g, h\rangle_{\mathcal{H}} \neq 0$ and so

$$\mathbb{P}\left(\|\mathrm{D}F\|_{\mathcal{H}} = 0\right) = 0.$$

2. Let $B \in \mathcal{B}(\mathbb{R})$. Then for all $n \in \mathbb{N}^*$:

$$\frac{1}{q+1}\mathbb{E}\left[\mathbf{1}_{B\cap[-n,n]}(F)\|\mathrm{D}F\|^2_{\mathcal{H}}\right] = \mathbb{E}\left[F\int_{-\infty}^F \mathbf{1}_{B\cap[-n,n]}(y)\ \mathrm{d}y\right].$$

3. We moreover suppose that $\lambda(B) = 0$. Then $\mathbb{P}(F \in B) = 0$, meaning that the law of F is absolutely continuous with respect to the Lebesgue measure.

Let us begin.

1. We set $t_1, \ldots, t_q \in T$ and simply denote for now $g = g_{t_1,\cdots,t_q}$. We suppose that for every $h \in \mathcal{H}$, $\langle g, h\rangle_{\mathcal{H}} = 0$. Let $(e_i)_i$ an orthonormal basis of $\mathcal{H}$. We decompose f, and so g :

$$f = \sum_{j_1,\cdots,j_q=1}^{+\infty}\sum_{j_{q+1}=1}^{+\infty} a_{j_1,\cdots,j_{q+1}} e_{j_1}\otimes\cdots\otimes e_{j_{q+1}}.$$

Hence,

$$g = \sum_{j_{q+1}=1}^{+\infty}\left[\sum_{j_1,\cdots,j_q=1}^{+\infty} a^{+\infty}_{j_1,\cdots,j_q=1} e_{j_1}(t_1)\cdots e_{j_q}(t_q)\right] e_{j_{q+1}}.$$

And we have the following expression for $\langle g, h\rangle$:

$$\langle g, h\rangle_{\mathcal{H}} = \sum_{j_{q+1}=1}^{+\infty}\left[\sum_{j_1,\cdots,j_q=1}^{+\infty} a_{j_1,\cdots,j_q} e_{j_1}(t_1)\cdots e_{j_q}(t_q)\right]\langle e_{j_{q+1}}, h\rangle_{\mathcal{H}}.$$

Now, let us note

$$\varphi : (t_1, \cdots, t_q) \longmapsto \langle g_{t_1,\cdots,t_q}, h\rangle_{\mathcal{H}}.$$

By hypothesis, φ is the zero map, so belongs to $L^2(T^q)$. Moreover, we have the decomposition of φ in the orthogonal basis $(e_i)_i$:

$$\varphi = \sum_{j_1,\cdots,j_q=1}^{+\infty}\left[\sum_{j_{q+1}=1}^{+\infty} a_{j_1,\cdots,j_q}\langle e_{j_{q+1}}, h\rangle_{\mathcal{H}}\right] e_{j_1}\otimes\cdots\otimes e_{j_q}.$$

So we can take the norm of it, and use the fact that it is equal to zero :

$$0 = \sum_{j_1,\cdots,j_q=1}^{+\infty}\left[\sum_{j_{q+1}=1}^{+\infty} a_{j_1,\cdots,j_q}\langle e_{j_{q+1}}, h\rangle_{\mathcal{H}}\right]^2.$$

If we take $h = e_k$, then all the sum is zero except for the term in $j_{q+1} = k$:

$$\forall k \in \mathbb{N}^*, a_{j_1,\cdots,j_q,k} = 0.$$

By symmetry of the function f, it means that all the coefficients $a_{j_1,\cdots,j_{q+1}}$ are zero, and so $f = 0$, which is excluded by hypothesis. Hence, we concluded that there exists $h \in \mathcal{H}$ such that for all $t_1, \cdots, t_q \in T$:

$$\langle g_{t_1,\cdots,t_q}, h\rangle_{\mathcal{H}} \neq 0.$$

It means, if we note φ like before that $\|\varphi\|_{L^2(T^q)} > 0$. By induction hypothesis, it means that

$$\mathcal{L}\left(I_q(\varphi)\right) \ll \lambda.$$

But, since $F = I_{q+1}(f)$, we have for all $t \in T$:

$$\mathrm{D}_t F = (q+1)I_q(f(t,\cdot)).$$

Hence, since $f(t,\cdot) = \left\{(t_1,\cdots,t_q) \longmapsto g_{t_1,\cdots,t_q}(t)\right\}$, we have almost surely :

$$\langle \mathrm{D}F, h\rangle = \varphi.$$

Hence,

$$\mathbb{P}\left(\langle \mathrm{D}F, h\rangle_{\mathcal{H}} = 0\right) = \mathbb{P}(\varphi = 0) = 0,$$

since $\mathbb{P}_\varphi \ll \lambda$. Since

$$\left\{\|\mathrm{D}F\|_{\mathcal{H}} = 0\right\} \subset \left\{\langle \mathrm{D}F, h\rangle_{\mathcal{H}} = 0\right\},$$

we finally proved that

$$\mathbb{P}\left(\|\mathrm{D}F\|_{\mathcal{H}} = 0\right) = 0.$$

2. Let $B \in \mathcal{B}(\mathbb{R})$. We set $n \in \mathbb{N}^*$. We want to prove that

$$\frac{1}{q+1}\mathbb{E}\left[\mathbf{1}_{B\cap[-n,n]}(F)\|\mathrm{D}F\|^2_{\mathcal{H}}\right] \overset{(*)}{=} \mathbb{E}\left[\int_{-\infty}^F \mathbf{1}_{B\cap[-n,n]}(z)\ \mathrm{d}z\right].$$

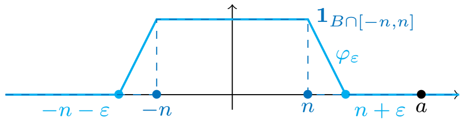
We use a Dynkin argument. Indeed, the set

$$\{B \in \mathcal{B}(\mathbb{R}), (*) \text{ is true}\}$$

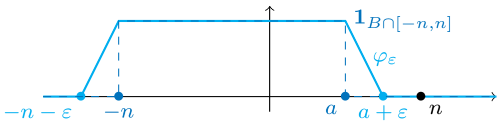
is Dynkin system ($\varnothing$ belongs to this set, it is stable by difference, and increasing limit by property of indicator functions). We just need to prove it for $B =]-\infty, a]$, with $a \in \mathbb{R}$. The idea is to use the integration by part formula we just derive for the map, for all $n \in \mathbb{N}^*$:

$$x \longmapsto \int_{-\infty}^x \mathbf{1}_{B\cap[-n,n]}(y)\ \mathrm{d}y,$$

which is bounded by $2n$, but not C^1 (only Lipschitz). To avoid this problem, we approach $\mathbf{1}_{]-\infty,a]\cap[-n,n]}$ by linear interpolation. For $n \leqslant |a|$, we define φ_ε as :



And we define for $|a| \leqslant n$:



We drew it for $a \geqslant 0$, but we define φ_ε symmetrically for $a \leqslant 0$. Then, φ_ε converges pointwise to $\mathbf{1}_{B\cap[-n,n]}$, is bounded by 1 and has its support included in $[-n-\varepsilon, n+\varepsilon]$. Since φ_ε is continuous, the integration by parts formula for $\int_{-\infty}^{\bullet} \varphi_\varepsilon$ gives

$$\mathbb{E}\left[\varphi_\varepsilon(F)\left\langle \mathrm{D}F, -\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[F\int_{\infty}^{F} \varphi_\varepsilon(z)\ \mathrm{d}z\right].$$

But, since $F = I_{q+1}(f)$, we have $L^{-1}F = \frac{-F}{q+1}$, so we have :

$$\frac{1}{q+1}\mathbb{E}\left[\varphi_\varepsilon(F)\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = \mathbb{E}\left[F\int_{-\infty}^{F} \varphi_\varepsilon(z)\ \mathrm{d}z\right].$$

So, by letting $[\varepsilon \to 0^+]$ (we can do it by dominated convergence, φ_ε has its support included in $[-2n, 2n]$ for ε small enough), we get :

$$\frac{1}{q+1}\mathbb{E}\left[\mathbf{1}_{[-n,n]\cap B}(F)\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = \mathbb{E}\left[F\int_{-\infty}^{F} \mathbf{1}_{[-n,n]\cap B}(z)\ \mathrm{d}z\right].$$

We proved $(*)$ for every $B =]-\infty, a]$, with $a \in \mathbb{R}$. Since the σ-algebra generated by those intervals is exactly $\mathcal{B}(\mathbb{R})$, we conclude by Dynkin lemma that $(*)$ is true for every $B \in \mathcal{B}(\mathbb{R})$.

3. We suppose than $\lambda(B) = 0$. We want to prove that $\mathbb{P}_F(B) = \mathbb{P}(F \in B) = 0$. By hypothesis, for every interval $I \subset \mathbb{R}$, $\int_I \mathbf{1}_B\ \mathrm{d}\lambda = 0$. Since $\lambda(B) = 0$, it means that

$$\frac{1}{q+1}\mathbb{E}\left[\mathbf{1}_{[-n,n]\cap B}(F)\|\mathrm{D}F\|_{\mathcal{H}}^2\right] = 0.$$

Since the left hand side is the expectation of a positive quantity, it means that $\mathbb{P}$-almost surely, we have

$$\mathbf{1}_{[-n,n]\cap B}(F)\|\mathrm{D}F\|_{\mathcal{H}}^2 = 0.$$

By the previous point, $\mathbb{P}\left(\|\mathrm{D}F\|_{\mathcal{H}} > 0\right) = 1$, meaning that $\mathbb{P}$-almost surely, we have $\mathbf{1}_{[-n,n]\cap B}(F) = 0$, so by taking the expectation :

$$\forall n \in \mathbb{N}, \mathbb{P}(F \in B \cap [-n, n]) = 0.$$

By increasing property of the probability, we conclude that $\mathbb{P}(F \in B) = 0$. This shows that $\mathbb{P}_F$ is indeed absolutely continuous with respect to the Lebesgue measure. □

V.4 Nelson's hypercontractivity

V.4.1 Statement

Theorem V.3 : Hypercontractivity

Let $F \in L^p$, for $p > 1$ and let $q(t) = 1 + e^{2t}(p-1)$. Then,

$$\mathbb{E}\left[|P_tF|^q\right]^{\frac{1}{q}} \leqslant \mathbb{E}\left[|F|^p\right]^{\frac{1}{p}}.$$

V.4.2 Equivalence of the norms in Wiener chaos

A very important corollary is the following.

Corollary V.2 : Equivalence of the norms

Let $F \in \mathfrak{H}_p$ and $1 < q < r$. Then,

$$\mathbb{E}\left[|F|^q\right]^{\frac{1}{q}} \leqslant \mathbb{E}\left[|F|^r\right]^{\frac{1}{r}} \leqslant \left(\frac{r-1}{q-1}\right)^{\frac{p}{2}} \mathbb{E}\left[|F|^q\right]^{\frac{1}{q}}.$$

All the L^q-norms are equivalents in every Wiener chaos.

Proof : The first inequality is well known, it is an application of Jensen inequality. For the second inequality, we use Nelson hypercontractivity (in L^q) : let $t \geqslant 0$ such that

$$r = 1 + e^{2t}(q-1).$$

Then,

$$\mathbb{E}\left[|P_tF|^r\right]^{\frac{1}{r}} \leqslant \mathbb{E}\left[|F|^q\right]^{\frac{1}{q}}.$$

But, by definition of the Ornstein-Ulhenbeck operator, we have, since $F \in \mathfrak{H}_p$:

$$P_tF = e^{-pt}F.$$

We finally have

$$\mathbb{E}\left[|F|^r\right]^{\frac{1}{r}} \leqslant e^{pt}\mathbb{E}\left[|F|^q\right]^{\frac{1}{q}},$$

giving the expected inequality. □

VI Applications

VI.1 Poincaré inequality

We begin by a generalization of we saw in the one dimensional case.

Proposition VI.1 : Poincaré inequality

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

$$\operatorname{Var}(F) \leqslant \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right].$$

Proof : We will prove it in two ways, one copying the case where $\Omega = \mathbb{R}$ we treated in section I, and another using explicitly the Malliavin derivative, and its Wiener decomposition.

1. The idea is to use the operators δ, L and D, by observing that for all $F \in L^2$,

$$P_0 F = F \text{ and } P_\infty F = \mathbb{E}[F].$$

The last equality can be shown using dominated convergence. Let us see the variance of F like an expectation of a product :

$$\operatorname{Var}(F) = \mathbb{E}\left[F\left(F - \mathbb{E}[F]\right)\right].$$

We see then an increment of $(P_t)_t$:

$$\operatorname{Var}(F) = \mathbb{E}\left[F\left(P_0 F - P_\infty F\right)\right].$$

And so we can express it as an integral :

$$\operatorname{Var}(F) = -\mathbb{E}\left[F \int_0^{+\infty} \frac{\mathrm{d}P_t F}{\mathrm{d}t}\ \mathrm{d}t\right].$$

But, since $P_t F = \sum_{n=0}^{+\infty} e^{-nt} J_n F$, we have in fact

$$\frac{\mathrm{d}P_t F}{\mathrm{d}t} = L P_t F.$$

So, we have :

$$\operatorname{Var}(F) = -\mathbb{E}\left[F \int_0^{+\infty} L P_t F\ \mathrm{d}t\right].$$

Since $L = -\delta D$:

$$\operatorname{Var}(F) = \mathbb{E}\left[F \int_0^{+\infty} \delta \mathrm{D} P_t F\ \mathrm{d}t\right].$$

We switch integrals :

$$\operatorname{Var}(F) = \int_0^{+\infty} \mathbb{E}\left[F \delta \mathrm{D} P_t F\right]\ \mathrm{d}t.$$

By duality :

$$\operatorname{Var}(F) = \int_0^{+\infty} \mathbb{E}\left[\langle \mathrm{D}F, \mathrm{D}P_t F\rangle_{\mathcal{H}}\right]\ \mathrm{d}t.$$

Using the Wiener chaos expansion, we have $\mathrm{D}P_t F = e^{-t} P_t \mathrm{D}F$. So, by Cauchy-Schwarz, we have :

$$\operatorname{Var}(F) \leqslant \int_0^{+\infty} e^{-t}\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}} \|P_t \mathrm{D}F\|_{\mathcal{H}}\right]\ \mathrm{d}t.$$

And by Cauchy-Schwarz once again :

$$\operatorname{Var}(F) \leqslant \int_0^{+\infty} e^{-t}\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right]^{\frac{1}{2}} \mathbb{E}\left[\|P_t \mathrm{D}F\|_{\mathcal{H}}^2\right]^{\frac{1}{2}}\ \mathrm{d}t.$$

By contraction property of P_t (here on $\mathcal{H}$) :

$$\operatorname{Var}(F) \leqslant \int_0^{+\infty} e^{-t}\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right]\ \mathrm{d}t = \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right].$$

And here is our inequality.

2. This proof is quicker but is more difficult, since it uses non elementary tools. We expend F in Wiener chaos :

$$F - \mathbb{E}[F] = \sum_{p=1}^{+\infty} J_p F.$$

So,

$$\operatorname{Var}(F) = \sum_{p=1}^{+\infty} \mathbb{E}[J_p F^2].$$

And now, the result is immediate by the theorem about the characterization of $\mathbb{D}^{1,2}$ in terms of convergence of a series :

$$\operatorname{Var}(F) \leqslant \sum_{p=1}^{+\infty} p\mathbb{E}[J_p F^2] = \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\right].$$

The inequality is proved. □

VI.2 Variance expansions

We set here $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$ with μ a non atomic measure.

Proposition VI.2 : Variance expansions for $\mathbb{D}^{\infty,2}$ variables

Let $F \in \mathbb{D}^{\infty,2} = \bigcap_{p \geqslant 1} \mathbb{D}^{p,2}$.

i. We have the following expansion :

$$\operatorname{Var}(F) = \sum_{p=1}^{+\infty} \frac{\|\mathbb{E}\left[\mathrm{D}^p F\right]\|_{L^2(T^p)}^2}{p!}.$$

ii. If we suppose moreover that

$$\frac{\mathbb{E}\left[\|\mathrm{D}^p F\|^2_{L^2(T^p)}\right]}{p!} \xrightarrow[p\to+\infty]{} 0,$$

then we have the following expansion :

$$\mathrm{Var}(F) = \sum_{p=1}^{+\infty} \frac{(-1)^{p+1}\mathbb{E}\left[\|\mathrm{D}^p F\|^2_{L^2(T^p)}\right]}{p!}.$$

Proof : We do the same proof as the case where $\Omega = \mathbb{R}$ and $\mathbb{P} = \gamma$.

i. We use a corollary we proves in section III about Malliavin derivative. More precisely, we proves that if $F \in \mathbb{D}^{\infty,2}$ then

$$F = \mathbb{E}[F] + \sum_{n=1}^{+\infty} \frac{I_n\left(\mathbb{E}\left[\mathrm{D}^n F\right]\right)}{n!}.$$

So, by taking the expectation of it :

$$\mathbb{E}[F^2] = \mathbb{E}[F]^2 + \sum_{n=1}^{+\infty} \frac{n!\|\mathbb{E}\left[\mathrm{D}^n F\right]\|^2_{L^2(T^n)}}{(n!)^2},$$

which exactly what we expected.

ii. We introduce

$$g(t) \overset{\text{def.}}{=} \mathbb{E}\left[\left(P_{\frac{-\ln t}{2}} F\right)^2\right].$$

1. Let us show that for all $p \in \mathbb{N}$:

$$g^{(p)}(t) = \mathbb{E}\left[\left\|P_{\frac{-\ln t}{2}} \mathrm{D}^p F\right\|^2_{L^2(T^p)}\right].$$

To do this, we just need to treat the case $p = 1$, the higher cases would be identical. By differentiating a composition of functions :

$$g'(t) = \frac{-1}{t}\mathbb{E}\left[P_{\frac{-\ln t}{2}} F \; L P_{\frac{-\ln t}{2}} F\right].$$

Since $L = -\delta\mathrm{D}$:

$$g'(t) = \frac{1}{t}\mathbb{E}\left[P_{\frac{-\ln t}{2}} F \; \delta\mathrm{D} P_{\frac{-\ln t}{2}} F\right].$$

By duality :

$$g'(t) = \frac{1}{t}\mathbb{E}\left[\left\|\mathrm{D} P_{\frac{-\ln t}{2}} F\right\|^2_{L^2(T)}\right].$$

Finally, we have $\mathrm{D}P_t = e^{-t} P_t \mathrm{D}$, so :

$$g'(t) = \mathbb{E}\left[\left\|P_{\frac{-\ln t}{2}} \mathrm{D} F\right\|^2_{L^2(T)}\right].$$

2. Notice that we can extend g on $[0, 1]$ with

$$g(0) = \mathbb{E}[F]^2 \text{ and } g(1) = \mathbb{E}[F^2].$$

By Taylor formula with integral reminder around 1, we have for all $N \in \mathbb{N}^*$, since

$$g^{(p)}(1) = \mathbb{E}\left[\|\mathrm{D}^p F\|^2_{L^2(T^p)}\right],$$

we get :

$$\begin{aligned} g(0) - g(1) = & \sum_{p=0}^{N} \frac{(-1)^p \mathbb{E}\left[\|\mathrm{D}^p F\|^2_{L^2(T^p)}\right]}{p!} \\ & + \quad (-1)^N \int_0^1 \frac{t^N}{N!} g^{(N+1)}(t) \; \mathrm{d}t. \end{aligned}$$

Since by contraction property, we have

$$\left|\int_0^1 \frac{t^N}{n!} g^{(p+1)}(t) \; \mathrm{d}t\right| \leqslant \frac{\mathbb{E}\left[\left\|\mathrm{D}^{N+1} F\right\|^2_{L^2(T^{N+1})}\right]}{(N+1)!}.$$

This goes to zero by hypothesis. So, we concludes in our expansion **ii.**.

Note : Same remark as the one-dimensional case. We could have proved our expansion **(i)** by using this method and expanding in 0. Nevertheless, it uses the hypothesis formulated in **ii.** we don't need to suppose for **i.**. □

References

— *Normal approximations with Malliavin calculus : from Stein's method to universality*, Ivan Nourdin & Giovanni Peccati, 2012 ;

— *The Malliavin Calculus and related topics*, David Nualart, 1995 ;

— *An Introduction to Wiener Analysis*, Ali Süleyman Üstünel, 1995.

Part II - Stein's method

Table of contents

I Stein's method for one-dimensional normal approximation

I.1 Stein's lemma and Stein's equation

We want to quantify how far a law is far from the normal distribution. To do this, we use this characterization of the Gaussian normal distribution.

Lemma I.1 : Stein's lemma

Let N a real random variable. The following assertions are equivalents :

(i) N follows the law $\mathcal{N}(0,1)$;

(ii) For all differentiable $f:\mathbb{R}\longrightarrow\mathbb{R}$ such that $f'\in L^1(\gamma)$, we have $\mathbb{E}[|f'(N)|],\mathbb{E}[|Nf(N)|]<+\infty$ and

$$\mathbb{E}[Nf(N)]=\mathbb{E}[f'(N)].$$

Proof of the lemma : **[$\Longrightarrow$]** We suppose $N\sim\mathcal{N}(0,1)$. Let $f:\mathbb{R}\longrightarrow\mathbb{R}$ differentiable such that $f'\in L^1(\gamma)$. Since $\mathbb{E}[N]=0$, we can suppose that $f(0)=0$. Then, immediately, $\mathbb{E}[|f'(N)|]<+\infty$. Moreover,

$$\mathbb{E}\left[|Nf(N)|\right]=\int_{\mathbb{R}}|xf(x)|\,\mathrm{d}\gamma(x).$$

We write f as the integral of its derivative. In order to use triangular inequality, we have to be careful with the bounds of the integrals :

$$\mathbb{E}\left[|Nf(N)|\right]\leqslant\int_0^{+\infty}x\left(\int_0^x\left|f'(t)\right|\,\mathrm{d}t\right)\mathrm{d}\gamma(x)-\int_{-\infty}^0x\left(\int_0^{-x}\left|f'(t)\right|\,\mathrm{d}t\right)\mathrm{d}\gamma(x).$$

By performing $x\leftarrow -x$ in the second integral, we get that the second integral in equal to the first :

$$\mathbb{E}\left[|Nf(N)|\right]\leqslant 2\int_0^{+\infty}x\left(\int_0^x\left|f'(t)\right|\,\mathrm{d}t\right)\,\mathrm{d}\gamma(x).$$

By Fubini's theorem :

$$\mathbb{E}\left[|Nf(N)|\right]\leqslant 2\int_0^{+\infty}|f'(t)|\left(\int_t^{+\infty}x\,\mathrm{d}\gamma(x)\right)\,\mathrm{d}t.$$

Finally, since this integral can be computed, we have :

$$\mathbb{E}\left[|Nf(N)|\right]\leqslant 2\int_0^{+\infty}|f'(t)|\,\mathrm{d}\gamma(t).$$

So,

$$\mathbb{E}\left[|Nf(N)|\right]\leqslant 2\mathbb{E}[|f'(N)|]<+\infty.$$

Finally, we show that $\mathbb{E}[Nf(N)]=\mathbb{E}[f'(N)]$. This is again a Fubini's theorem which allows us to conclude (an integration by parts cannot be done, we don't know how $f(t)e^{\frac{-t^2}{2}}$ behave on $\pm\infty$). We have :

$$\begin{aligned}\int_{\mathbb{R}} xf(x)\ \mathrm{d}\gamma(x) \\ = \int_0^{+\infty}\int_{\mathbb{R}} \mathbf{1}_{0\leqslant t\leqslant x} xf'(t)\ \mathrm{d}t\ \mathrm{d}\gamma(x) \\ -\int_{-\infty}^0\int_{\mathbb{R}} \mathbf{1}_{x\leqslant t\leqslant 0} xf'(t)\ \mathrm{d}t\ \mathrm{d}\gamma(x).\end{aligned}$$

We switch, and use the fact, once again, that $\int_t^{+\infty} x\ \mathrm{d}\gamma(x) = e^{\frac{-t^2}{2}}$ to conclude :

$$\int_{\mathbb{R}} xf(x)\ \mathrm{d}\gamma(x) = \int_{\mathbb{R}} f'(t)\ \mathrm{d}\gamma(t).$$

That is the wanted equality.

[$\Longleftarrow$] Suppose that the real random variable N satisfies the condition **(ii)**. There are two ways to prove the reciprocal sens.

1. For $f(x) = x^n$, we have :

$$\mathbb{E}[N^{n+1}] = n\mathbb{E}[N^{n-1}].$$

Since for $f(x) = 1$, we have $\mathbb{E}[N] = 0$, and $f(x) = x$, $\mathbb{E}[N^2] = 1$, it means that N admits moments at every order, and by the induction relation, N admits the exact same moments as a Gaussian $\mathcal{N}(0,1)$. We just have to show that it means that N follows $\mathcal{N}(0,1)$. To do this, we compute the characteristic function of N. By Fubini,

$$\mathbb{E}\left[e^{\mathrm{i}tN}\right] = \sum_{n=0}^{+\infty} \frac{\mathrm{i}^{2n}t^{2n}}{(2n)!}\frac{(2n)!}{2^n n!}.$$

This leads to

$$\mathbb{E}\left[e^{\mathrm{i}tN}\right] = \sum_{n=0}^{+\infty} \frac{(-1)^n t^{2n}}{2^n n!} = e^{\frac{-t^2}{2}}.$$

By Lévy's theorem, $N \sim \mathcal{N}(0,1)$.

2. Let us show that the characteristic function of N is the one of $\mathcal{N}(0,1)$ by getting a differential equation. Let $t\in\mathbb{R}$. For $f = \cos(t\cdot)$, then $\mathbb{E}[|f'(N)|] < +\infty$ and we have :

$$t\mathbb{E}[\sin tN] = -\mathbb{E}[N\cos tN]$$

Same for $f = \sin(t\cdot)$:

$$t\mathbb{E}[\cos tN] = \mathbb{E}[N\sin tN]$$

Hence,

$$t\mathbb{E}\left[e^{\mathrm{i}tN}\right] = \mathbb{E}\left[N\left(\sin tN - \mathrm{i}\cos(tN)\right)\right] = -\mathbb{E}\left[\mathrm{i}Ne^{\mathrm{i}tN}\right].$$

Since $N\in L^1$ (by taking $f(x) = 1$), we have, with φ_N the characteristic function of N :

$$\varphi_N'(t) = \mathrm{i}\mathbb{E}\left[Ne^{\mathrm{i}tN}\right].$$

Finally, we get the differential equation :

$$\varphi_N'(t) = -t\varphi_N(t)$$

with initial value equal to 1, so we have

$$\varphi_N(t) = e^{\frac{-t^2}{2}},$$

meaning that $N\sim\mathcal{N}(0,1)$. □

Corollary I.1 : Stein's lemma for general Gaussian

Let $N\in L^2$, $\mu\in\mathbb{R}$ and $\sigma^2 > 0$. Then the following assertions are equivalents :

(i) N follows the law $\mathcal{N}(\mu,\sigma^2)$;

(ii) For all differentiable $f:\mathbb{R}\longrightarrow\mathbb{R}$ such that $\mathbb{E}[|f'(N)|]$ is finite, then $\mathbb{E}\left[\left|\frac{N-\mu}{\sigma^2}f(N)\right|\right] < +\infty$ and

$$\mathbb{E}\left[f'(N)\right] = \mathbb{E}\left[\frac{N-\mu}{\sigma^2}f(N)\right].$$

Proof : [$\Longrightarrow$] We suppose that $N\sim\mathcal{N}(\mu,\sigma^2)$. Let $Z = \frac{N-\mu}{\sigma}\sim\mathcal{N}(0,1)$. Then by Setin's lemma, for all $g:\mathbb{R}\longrightarrow\mathbb{R}$ such that $\mathbb{E}[|g'(Z)|] < +\infty$, we have $\mathbb{E}[|Zg(Z)|] < +\infty$ and

$$\mathbb{E}[g'(Z)] = \mathbb{E}[Zg(Z)].$$

Let $f:\mathbb{R}\longrightarrow\mathbb{R}$ derivable such that $\mathbb{E}[|f'(N)|] < +\infty$. We define

$$g(z) \overset{\text{def.}}{=} f(\sigma z + \mu).$$

Then g is derivable, with $g'(z) = \sigma f'(\sigma z+\mu)$ and

$$\mathbb{E}\left[|g'(Z)|\right] = \sigma\mathbb{E}[|f'(N)|] < +\infty.$$

Hence, we have $\mathbb{E}[|Zg(Z)|] < +\infty$, so :

$$\mathbb{E}\left[|Zg(Z)|\right] = \mathbb{E}\left[\left|\frac{N-\mu}{\sigma}f(N)\right|\right] < +\infty,$$

and we also have

$$\mathbb{E}[g'(Z)] = \mathbb{E}[Zg(Z)].$$

It rewrites here as :

$$\sigma\mathbb{E}[f'(N)] = \mathbb{E}\left[\frac{N-\mu}{\sigma}f'(N)\right].$$

This concludes the direct implication.

[$\Longleftarrow$] Same idea as the direct implication. We define Z as the direct implication. Let us show that for every $g:\mathbb{R}\longrightarrow\mathbb{R}$ derivable such that $\mathbb{E}[|g'(Z)|] < +\infty$ then $\mathbb{E}[|Zg(Z)|] < +\infty$ and

$$\mathbb{E}[Zg(Z)] = \mathbb{E}[g'(Z)].$$

This would implies that $Z\sim\mathcal{N}(0,1)$ by Stein's lemma, and so that $N\sim\mathcal{N}(\mu,\sigma^2)$. Let g derivable with $g'(Z)\in L^1$. We define like before

$$f(x) \overset{\text{def.}}{=} g\left(\frac{x-\mu}{\sigma}\right).$$

Then $f(N) = g(Z)$, $f'(x) = \frac{1}{\sigma}g'\left(\frac{x-\mu}{\sigma}\right)$. So, $\mathbb{E}[|f'(N)|] < +\infty$, and by hypothesis, it means that

$$\mathbb{E}\left[\left|\frac{N-\mu}{\sigma^2}f(N)\right|\right] < +\infty.$$

So, we have

$$\mathbb{E}\left[|Zg(Z)|\right] = \sigma\mathbb{E}\left[\left|\frac{N-\mu}{\sigma^2}f(N)\right|\right] < +\infty.$$

Finally, by hypothesis, we have

$$\mathbb{E}\left[f'(N)\right] = \mathbb{E}\left[\frac{N-\mu}{\sigma^2}f(N)\right].$$

In terms of g, it means that

$$\frac{1}{\sigma}\mathbb{E}\left[g'(Z)\right] = \frac{1}{\sigma}\mathbb{E}\left[Zg(Z)\right].$$

So $Z\sim\mathcal{N}(0,1)$, and $N\sim\mathcal{N}(\mu,\sigma^2)$. □

Corollary I.2 : A first step to multi dimensional Stein's lemma

Let $C \in \mathcal{S}_d^+(\mathbb{R})$ and $N = (N_1, \cdots, N_d) \sim \mathcal{N}(0, C)$. Then, for all $f \in \mathcal{C}_b^1(\mathbb{R}^d)$ and for all $i \in [\![1, d]\!]$:

$$\mathbb{E}\left[N_i f(N_1, \cdots, N_d)\right] = \sum_{j=1}^{d} C_{i,j} \mathbb{E}\left[\frac{\partial f}{\partial x_j}(N_1, \cdots, N_d)\right].$$

Proof : **1.** We do first the case where C is diagonal. We note $\lambda_1, \cdots, \lambda_d$ the diagonal entries of the matrix. Then $(N_1, \cdots, N_d)$ are mutually independent. So :

$$\begin{aligned} &\mathbb{E}\left[N_i f(N_1, \cdots, N_d)\right] \\ = &\int_{\mathbb{R}^{d-1}} \mathbb{E}\left[N_i f(x_1, \cdots, N_i, \ldots, x_d)\right] \prod_{j \neq i} \frac{e^{\frac{-x_j^2}{2\lambda_j}}}{\sqrt{\lambda_j}} \frac{\widehat{\mathrm{d}x_i}}{\sqrt{(2\pi)^{d-1}}}. \end{aligned}$$

Where $\widehat{\mathrm{d}x_i}$ means the product of $\mathrm{d}x_j$ without the the term $\mathrm{d}x_i$. By Stein's lemma in the one-dimensional case, we get :

$$\begin{aligned} &\mathbb{E}\left[N_i f(N_1, \cdots, N_d)\right] \\ = &\lambda_i \int_{\mathbb{R}^{d-1}} \mathbb{E}\left[\frac{\partial f}{\partial x_i}(x_1, \cdots, N_i, \ldots, x_d)\right] \\ &\cdot \prod_{j \neq i} \frac{e^{\frac{-x_j^2}{2\lambda_j}}}{\sqrt{\lambda_j}} \frac{\widehat{\mathrm{d}x_i}}{\sqrt{(2\pi)^{d-1}}}. \end{aligned}$$

And so

$$\mathbb{E}\left[N_i f(N_1, \cdots, N_d)\right] = \lambda_i \mathbb{E}\left[\frac{\partial f}{\partial x_i}(N_1, \cdots, N_d)\right].$$

2. In general case, we diagonalize $C = OD^tO$, with $O \in \mathcal{O}_d(\mathbb{R})$. We have

$$\mathbb{E}[N_i f(N)] = \int_{\mathbb{R}^d} x_i f(x) e^{\frac{-1}{2}\langle x, C^{-1}x\rangle} \frac{\mathrm{d}x}{\sqrt{(2\pi)^d \det(C)}}.$$

We have

$$\langle x, C^{-1}x\rangle = \langle O^{-1}x, D^{-1}O^{-1}x\rangle.$$

We set the change of variables $y = O^{-1}x$. We note $\omega_{i,j}$ the entries of the matrix O. We set $g = f \circ O^{-1}$.

$$\mathbb{E}[N_i f(N)] = \sum_{j=1}^{d} \omega_{i,j} \int_{\mathbb{R}^d} y_j g(y) e^{\frac{-1}{2}\langle y, D^{-1}y\rangle} \frac{\mathrm{d}y}{\sqrt{(2\pi)^d \det(D)}}.$$

Then, by the diagonal case, the integral gives :

$$\begin{aligned} &\mathbb{E}[N_i f(N)] \\ = &\sum_{j=1}^{d} \omega_{i,j} \lambda_j \int_{\mathbb{R}^d} \frac{\partial g}{\partial y_j}(y) e^{\frac{-1}{2}\langle y, D^{-1}y\rangle} \frac{\mathrm{d}y}{\sqrt{(2\pi)^d \det(D)}}. \end{aligned}$$

Hence, we are in the situation :

$$\mathbb{E}[N_i f(N)] = \int_{\mathbb{R}^d} \left[OD\nabla g(y)\right]_i e^{\frac{-1}{2}\langle y, D^{-1}y\rangle} \frac{\mathrm{d}y}{\sqrt{(2\pi)^d \det(D)}}.$$

Since $OD = CO$, we have :

$$\mathbb{E}[N_i f(N)] = \int_{\mathbb{R}^d} \left[CO\nabla g(y)\right]_i e^{\frac{-1}{2}\langle y, D^{-1}y\rangle} \frac{\mathrm{d}y}{\sqrt{(2\pi)^d \det(D)}}.$$

We have by differentiation of composed functions :

$$\nabla f(x) = O\nabla g(y)$$

Hence, by using again the change of variable $x = Oy$, we have :

$$\mathbb{E}[N_i f(N)] = \int_{\mathbb{R}^d} \left[C\nabla f(x)\right]_i e^{\frac{-1}{2}\langle x, C^{-1}x\rangle} \frac{\mathrm{d}x}{\sqrt{(2\pi)^d \det(C)}}.$$

And so we proved that

$$\mathbb{E}[N_i f(N)] = \sum_{j=1}^{d} C_{i,j} \mathbb{E}\left[\frac{\partial f}{\partial x_j}(N)\right].$$

This is the desired conclusion. □

We can express a Stein lemma for higher derivative order.

Corollary I.3 : Hermite-Stein lemma

Let N a real random variable, and $m \in \mathbb{N}^*$. The following assertions are equivalents :

(i) N follows the law $\mathcal{N}(0, 1)$;

(ii) For all $k \in [\![1, m]\!]$, for all k times differentiable functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ having all its derivative having at most a polynomial growth, we have $\mathbb{E}[|H_m(N)f(N)|] < +\infty$ and

$$\mathbb{E}[H_m(N)f(N)] = \mathbb{E}[H_{m-k}(N)f^{(k)}(N)].$$

Proof : For this proof, we note S^k the set of functions which are k times differentiable having, for it and for its derivatives, at most a polynomial growth.

[$\Longrightarrow$] We proceed by induction on m. The case $m = 1$ is the Stein's lemma. For $m \geqslant 1$, we suppose that **(ii)** is true. Let us check **(ii)** for $m + 1$.

• We claim that is it enough to prove it for $k = 1$. Indeed, let $k \in [\![2, m+1]\!]$, and $f \in S^k$. Since N is Gaussian,

$$\mathbb{E}\left[|H_{m+1}(N)f(N)|\right] < +\infty.$$

Since :

$$\mathbb{E}\left[H_{m+1-k}(N)f^{(k)}(N)\right] = \mathbb{E}\left[H_{m-(k-1)}(N)f^{(k)}(N)\right],$$

we have by induction hypothesis that :

$$\mathbb{E}\left[H_{m+1-k}(N)f^{(k)}(N)\right] = \mathbb{E}\left[H_m(N)f'(N)\right]$$

If we prove the case $k=1$, we could conclude that

$$\mathbb{E}\left[H_{m+1-k}(N)f^{(k)}(N)\right] = \mathbb{E}\left[H_{m+1}(N)f(N)\right].$$

• Let us prove the $k=1$ case. Let $f \in S^1$. Hence, since $N \sim \mathcal{N}(0,1)$, N admits moments at every order, and so

$$\mathbb{E}\left[|H_m(N)f(N)|\right] < +\infty.$$

Using the fact that f has a growth at most polynomial (no bound terms), and the fact that

$$\frac{\mathrm{d}^m}{\mathrm{d}x^m}\left[e^{\frac{-x^2}{2}}\right] = (-1)^m H_m(x) e^{\frac{-x^2}{2}},$$

we have by integration by parts that

$$\mathbb{E}\left[H_{m+1}(N)f(N)\right] = \mathbb{E}\left[H_m(N)f'(N)\right].$$

So we proved the case $k=1$, and so **(ii)**.

[$\Longleftarrow$] We proceed once again by induction on m. For $m=1$, this is not exactly the Stein's lemma, since we consider a smaller set of function. But, nevermind, the proof is the same, we just use the fact that every polynomial (or the cosine and the sine) belongs to S^∞ (the intersection of the S^k). Hence, the case $m=1$ is proved.

We suppose **(ii)** $\Longrightarrow$ **(i)** for $m \geqslant 1$. We suppose **(ii)** for $m+1$: for every $k \in [\![1, m+1]\!]$, and $g \in S^k$, we have $\mathbb{E}\left[|H_{m+1}(N)g(N)|\right] < +\infty$ and

$$\mathbb{E}\left[H_{m+1-k}(N)g^{(k)}(N)\right] = \mathbb{E}\left[H_{m+1}(N)g(N)\right].$$

We want to prove that $N \sim \mathcal{N}(0,1)$. By induction hypothesis, it is enough to prove that for every $f \in S^k$, we have $\mathbb{E}\left[|H_m(N)f(N)|\right] < +\infty$, and :

$$\mathbb{E}\left[H_{m-k}(N)f^{(k)}(N)\right] = \mathbb{E}\left[H_m(N)f(N)\right].$$

For the finite expectation, we just need to see that the hypothesis for $m+1$ case implies that N admits moment at every order, so

$$\mathbb{E}\left[|H_m(N)f(N)|\right] < +\infty.$$

Let $F(x) = \int_0^x f(t)\,\mathrm{d}t$. Then $F \in S^{k+1}$ and $F' = f$. By applying our hypothesis for $k=1$, and $g = F \in S^{k+1} \subset S^1$, we have

$$\mathbb{E}\left[H_m(N)f(N)\right] = \mathbb{E}\left[H_{m+1}(N)F(N)\right].$$

Since $F \in S^{k+1}$,

$$\mathbb{E}\left[H_m(N)f(N)\right] = \mathbb{E}\left[H_{m+1-(k+1)}(N)F^{(k+1)}(N)\right].$$

Which gives in terms of f :

$$\mathbb{E}\left[H_m(N)f(N)\right] = \mathbb{E}\left[H_{m-k}(N)f^{(k)}(N)\right].$$

This implies by induction hypothesis that $N \sim \mathcal{N}(0,1)$ and concludes our induction. □

The idea is the following : how far a law is from a Gaussian law ? To know it, the lemma will help us : we want to measure how far a variable satisfies the Stein's lemma, that is : if F a is random variable satisfying

$$\mathbb{E}\left[f'(F) - Ff(F)\right] \ll 1,$$

is F close to a $\mathcal{N}(0,1)$ law ? This is why we try to solve the *Stein's equation*.

Definition I.1

Let $N \sim \mathcal{N}(0,1)$, $h : \mathbb{R} \longrightarrow \mathbb{R}$ a measure function such that $\mathbb{E}[|h(N)|] < +\infty$. We call *Stein's equation* the following ordinary equation where f is the unknown and is absolutely continuous :

$$f'(x) - xf(x) = h(x) - \mathbb{E}[h(N)].$$

In our case one-dimensional case, the solution is easily known.

Proposition I.1 : Solution of the Stein's equation

The set of the solution of the Stein's equation is an affine space of dimension 1 given by :

$$f(x) = Ce^{\frac{x^2}{2}} + f_h(x),$$

where $C \in \mathbb{R}$ and f_h is the unique solution satisfying

$$e^{\frac{-x^2}{2}} f_h(x) \xrightarrow[x \to \pm\infty]{} 0.$$

It is given by

$$f_h(x) = e^{\frac{x^2}{2}} \int_{-\infty}^{x} (h(y) - \mathbb{E}[h(N)]) e^{\frac{-y^2}{2}}\,\mathrm{d}y.$$

Proof : • A quick way to prove it is to multiply the equation by $e^{\frac{-x^2}{2}}$ and see appear a derivative :

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[e^{\frac{-x^2}{2}}f(x)\right] = e^{\frac{-x^2}{2}}\left(h(x) - \mathbb{E}[h(N)]\right).$$

Since we suppose that $\mathbb{E}[|h(N)|] < +\infty$, then

$$\int_{\mathbb{R}} |h(y)|e^{\frac{-y^2}{2}}\,\mathrm{d}y < +\infty,$$

and we can integrate it from $-A$ to x, with $A > |x|$:

$$e^{\frac{-x^2}{2}}f(x) = e^{\frac{-A^2}{2}}f(-A) + \int_{-A}^{x} e^{\frac{-y^2}{2}}\left(h(y) - \mathbb{E}[h(N)]\right)\,\mathrm{d}y.$$

Since the integral terms admits a finite limit when $[A \to +\infty]$, then the case of $f(A)e^{\frac{-A^2}{2}}$. Hence, we can define

$$C \overset{\text{def.}}{=} \lim_{A\to-\infty} e^{\frac{-A^2}{2}}f(A) < +\infty.$$

The constant C only depends on f, and we have :

$$f(x) = Ce^{\frac{x^2}{2}} + e^{\frac{x^2}{2}}\int_{-\infty}^{x} e^{\frac{-y^2}{2}}\left(h(y) - \mathbb{E}[h(N)]\right)\,\mathrm{d}y.$$

Conversely, every function of this type solves the Stein's equation.

• The function f_h is solution, with $C = 0$. Moreover, for every $x \in \mathbb{R}$ and $y \in \mathbb{R}$:

$$\left|e^{\frac{-y^2}{2}}\left(h(y) - \mathbb{E}[h(N)]\right)\mathbf{1}_{\{y<x\}}\right| \leqslant e^{\frac{-y^2}{2}}\left|h(y) - \mathbb{E}[h(N)]\right|,$$

which is integrable on $\mathbb{R}$. By dominated convergence, we conclude that :

$$f_h(x) \xrightarrow[x\to-\infty]{} 0.$$

Moreover, in $+\infty$, we have

$$\int_{-\infty}^{+\infty} h(y)e^{\frac{-y^2}{2}}\,\mathrm{d}y = \mathbb{E}[h(N)]\int_{-\infty}^{+\infty} e^{\frac{-y^2}{2}}\,\mathrm{d}y.$$

Hence,

$$f_h(x) \xrightarrow[x\to+\infty]{} 0.$$

• Finally, if we suppose that g is a solution going to zero at $\pm\infty$, then there exists $C \in \mathbb{R}$ such that

$$g(x) - f_h(x) = Ce^{\frac{x^2}{2}}.$$

The left hand side is going to zero at $\pm\infty$. This the case of the right hand side if and only if $C = 0$, so $g = f_h$. There is an unique solution of the Stein's equation going to zero at $\pm\infty$. □

Here's the deal. We want to measure how far a random variable is from the law $\mathcal{N}(0,1)$. To do this, we introduce the following distance.

Definition I.2

Let $\mathcal{H}$ be a subset of $\mathcal{F}(\mathbb{R},\mathbb{R})$. We say that $\mathcal{H}$ is *separative* if for all random variables F, G such that $h(F), h(G) \in L^1$, we have :

$$\forall h \in \mathcal{H}, \mathbb{E}[h(F)] = \mathbb{E}[h(G)] \Longrightarrow F \overset{\text{law}}{=} G.$$

In this case, we define for all random variables F, G such that $h(F), h(G) \in L^1$ for all $h \in \mathcal{H}$:

$$d_{\mathcal{H}}(F,G) \overset{\text{def.}}{=} \sup_{h\in\mathcal{H}} \left|\mathbb{E}[h(F)] - \mathbb{E}[h(G)]\right|.$$

By the previous proposition, we immediately have the following corollary.

Corollary I.4 : Distance and Stein's equation

Let F a random variable and $N \sim \mathcal{N}(0,1)$. Then, if $\mathcal{H} \subset \{h : \mathbb{R} \longrightarrow \mathbb{R}, \mathbb{E}[|h(N)|] < +\infty\}$, we have

$$d_{\mathcal{H}}(F,N) = \sup_{h\in\mathcal{H}} \left|\mathbb{E}\left[f_h'(F) - Ff_h(F)\right]\right|.$$

In other words, the supremum does not contains explicitly N.

I.2 Some estimations of the distance for different class of functions

We will here explain the different spaces $\mathcal{H}$ we could take to estimate the previous distance. In particular, we will prove Stein's bound we used in the precedent part for Poincaré's second type inequality for the one-dimensional case.

I.2.1 Total variation distance

Definition I.3

We set $d \in \mathbb{N}^*$ and

$$\mathcal{H} = \left\{ \mathbf{1}_B, B \in \mathcal{B}\left(\mathbb{R}^d\right) \right\}.$$

The *total variation distance* is the distance in law for this space $\mathcal{H}$: for all random variables F, G,

$$d_{\mathrm{TV}}(F,G) \overset{\text{def.}}{=} \sup_{B \in \mathcal{B}(\mathbb{R}^d)} |\mathbb{P}(F \in B) - \mathbb{P}(G \in B)|.$$

If $d = 1$, the definition is equivalent to

$$d_{\mathrm{TV}}(F,G) = \sup_{x \in \mathbb{R}} |\mathbb{P}(F \leqslant x) - \mathbb{P}(G \leqslant x)|.$$

We can estimate a first *Stein's bound* for this distance.

Proposition I.2 : Stein's bound for total variation distance

i. Let $h : \mathbb{R} \longrightarrow [0,1]$ a measurable function. Then, there exists a version of f_h' such that we have :

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

ii. We note $\mathcal{F}_{\mathrm{TV}}$ the set of all absolutely continuous functions on $\mathbb{R}$ bounded by $\sqrt{\frac{\pi}{2}}$ and such that there exists a version of their derivative bounded by 2. Then, we have the following *Stein's bound*, for $F \in L^1$ and $N \sim \mathcal{N}(0,1)$:

$$d_{\mathrm{TV}}(F,N) \leqslant \sup_{f \in \mathcal{F}_{\mathrm{TV}}} \left|\mathbb{E}\left[f'(F) - Ff(F)\right]\right|.$$

Proof : i. Since

$$\int_{-\infty}^{x} (h(y) - \mathbb{E}[h(N)]) e^{\frac{-y^2}{2}} \, \mathrm{d}y = -\int_{x}^{+\infty} (h(y) - \mathbb{E}[h(N)]) e^{\frac{-y^2}{2}} \, \mathrm{d}y,$$

since the integral over $\mathbb{R}$ is zero, we have, by distinguishing the cases where $x < 0$ and $x > 0$, and since $|h| \leqslant 1$:

$$|f_h(x)| \leqslant e^{\frac{x^2}{2}} \int_{|x|}^{+\infty} e^{\frac{-y^2}{2}} \, \mathrm{d}y.$$

We call $S(x)$ the right hand side. Let us show that S admits its maximum at 0. We have

$$S(0) = \int_0^{+\infty} e^{\frac{-y^2}{2}} \, \mathrm{d}y = \sqrt{\frac{\pi}{2}}.$$

To do this, we compute the derivative for $x > 0$ and $x < 0$. For $x > 0$,

$$S'(x) = x e^{\frac{x^2}{2}} \int_x^{+\infty} e^{\frac{-y^2}{2}} \, \mathrm{d}y - 1.$$

Using the fact that the bound of the integral starts in x :

$$S'(x) \leqslant e^{\frac{x^2}{2}} \int_x^{+\infty} y e^{\frac{-y^2}{2}} \, \mathrm{d}y - 1.$$

We can compute this integral :

$$S'(x) \leqslant e^{\frac{x^2}{2}} e^{\frac{-x^2}{2}} - 1 = 0.$$

Hence, S is decreasing on $\mathbb{R}_+$. We do the same on $\mathbb{R}_-$. S has the expression

$$S(x) = e^{\frac{x^2}{2}} \int_{-\infty}^{x} e^{\frac{-y^2}{2}} \, \mathrm{d}y$$

on $\mathbb{R}_-$, so that S is the product of two positive increasing functions so S is increasing on $\mathbb{R}_-$. We conclude that for every $x \in \mathbb{R}$, $S(x) \leqslant S(0)$, and so that

$$\|f_h\|_\infty \leqslant \sqrt{\frac{\pi}{2}}.$$

For f_h', we use the differential equation satisfied by f_h. Since h takes its values in $[0,1]$, we have :

$$|f_h'(x)| \leqslant |x| e^{\frac{x^2}{2}} \int_{|x|}^{+\infty} e^{\frac{-y^2}{2}} \, \mathrm{d}y + 1.$$

We use the same bound trick as before :

$$|f_h'(x)| \leqslant e^{\frac{x^2}{2}} \int_{|x|}^{+\infty} y e^{\frac{-y^2}{2}} \, \mathrm{d}y + 1.$$

We can compute this integral, and we obtain finally :

$$\left\|f_h'\right\|_\infty \leqslant 2.$$

ii. We already have the estimation by Stein's equation :

$$d_{\mathrm{TV}}(F,N) = \sup_{\substack{h = \mathbf{1}_B \\ B \in \mathcal{B}(\mathbb{R})}} \left|\mathbb{E}\left[f_h'(F) - Ff_h(F)\right]\right|$$

But, if $h = \mathbf{1}_B$, $f_h \in \mathcal{F}_{\mathrm{TV}}$, so the Stein's bound is now immediate :

$$d_{\mathrm{TV}}(F,N) \leqslant \sup_{f \in \mathcal{F}_{\mathrm{TV}}} \left|\mathbb{E}\left[f'(F) - Ff(F)\right]\right|.$$

This concludes the proof. □

I.2.2 Kolmogorov distance

Definition I.4

We set $d \in \mathbb{N}^*$ and

$$\mathcal{H} = \left\{ \bigotimes_{i=1}^{d} \mathbf{1}_{]-\infty, z_i]}, z_i \in \mathbb{R} \right\}.$$

The *Kolmogorov distance* is the distance in law for this set $\mathcal{H}$. In other words, for all random variables F, G :

$$d_{\mathrm{Kol}}(F,G) \overset{\text{def.}}{=} \sup_{z_1,\cdots,z_d \in \mathbb{R}} \left| \mathbb{P}\left(F \in \prod_{i=1}^{d}]-\infty, z_i] \right) - \mathbb{P}\left(G \in \prod_{i=1}^{d}]-\infty, z_i] \right) \right|.$$

Then, we can see that $d_{\mathrm{Kol}} \leqslant d_{\mathrm{TV}}$. We can have a better Stein's bound.

Proposition I.3 : Stein's bound for Kolmogorov distance

i. Let $z \in \mathbb{R}$ and $h = \mathbf{1}_{]-\infty, z]}$. We set $f_z = f_h$. Then, there exists a version of f_z' such that we have :

$$\|f_z\|_\infty \leqslant \sqrt{\frac{\pi}{8}} \quad \text{and} \quad \|f_z'\|_\infty \leqslant 1.$$

ii. We note $\mathcal{F}_{\mathrm{Kol}}$ the set of all absolutely continuous functions on $\mathbb{R}$ bounded by $\sqrt{\frac{\pi}{8}}$ and such that there exists a version of their derivative bounded by 1. Then, we have the following *Stein's bound*, for $F \in L^1$ and $N \sim \mathcal{N}(0,1)$:

$$d_{\mathrm{Kol}}(F,N) \leqslant \sup_{f \in \mathcal{F}_{\mathrm{TV}}} \left| \mathbb{E}\left[f'(F) - F f(F) \right] \right|.$$

A very nice property of this distance is the following : under a condition of continuity of a cumulative distribution function, the Kolmogorov's distance allows to put a metric on the law convergence.

Theorem I.1 : Kolmogorov distance and convergence in distribution

(i) If a sequence of random variables converges in Kolmogorov sense, then it converges in law.
(ii) Let $(X_n)_n$ a sequence of random variables, and X a random variable such that its cumulative distribution function $\Phi : x \longmapsto \mathbb{P}(X \leqslant x)$ is continuous on $\mathbb{R}$. Then $(X_n)_n$ converges in law to X if and only if $(X_n)_n$ converges to X for d_{Kol}.

Proof : **(i)** This is in fact the case for every distance we present here. The fact that the cumulative distribution functions converge to an another cumulative distribution function at every point of continuity is equivalent to the convergence in distribution.
(ii) We know that

$$\mathbb{P}(X_n \leqslant x) - \mathbb{P}(X \leqslant x) \xrightarrow[n \to +\infty]{} 0,$$

for every $x \in \mathbb{R}$, and we want to prove

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(X_n \leqslant x) - \mathbb{P}(X \leqslant x)| \xrightarrow[n \to +\infty]{} 0.$$

That is proving uniform convergence from simple convergence. To do that, we use Dini's theorem. So, we set on a compact. More precisely, if we consider

$$f_n : x \in \left[\frac{-\pi}{2}, \frac{\pi}{2} \right] \longmapsto \begin{cases} 0 & \text{if} \quad x = \frac{-\pi}{2} \\ \mathbb{P}(X_n \leqslant \tan(x)) & \text{if} \quad x \in \left] \frac{-\pi}{2}, \frac{\pi}{2} \right[\\ 1 & \text{if} \quad x = \frac{\pi}{2} \end{cases}$$

and f like f_n but for X, then $(f_n)_n$ is a sequence of functions that converges pointwise to f on the compact $\left[\frac{-\pi}{2}, \frac{\pi}{2} \right]$ and such that for every $n \in \mathbb{N}$, f_n is an increasing function. Finally, f is continuous by composition of continuous functions. By second Dini's theorem, it means that

$$\sup_{x \in \left[\frac{-\pi}{2}, \frac{\pi}{2} \right]} |f_n(x) - f(x)| \xrightarrow[n \to +\infty]{} 0.$$

Since at the bounds, the quantity inside the supremum is zero, we have :

$$\sup_{x \in \left] \frac{-\pi}{2}, \frac{\pi}{2} \right[} |\mathbb{P}(X_n \leqslant \tan x) - \mathbb{P}(X \leqslant \tan x)| \xrightarrow[n \to +\infty]{} 0.$$

Finally, since $\tan$ is a bijection between $\left] \frac{-\pi}{2}, \frac{\pi}{2} \right[$ and $\mathbb{R}$, we finally have what we expected :

$$\sup_{z \in \mathbb{R}} |\mathbb{P}(X_n \leqslant z) - \mathbb{P}(X \leqslant z)| \xrightarrow[n \to +\infty]{} 0.$$

That is

$$d_{\mathrm{Kol}}(X_n, X) \xrightarrow[n \to +\infty]{} 0.$$

So this proves our equivalence. □

We can explicit a counter example if the cumulative distribution function is not continuous. We have

$$\delta_{\frac{1}{n}} \xrightarrow[n\to+\infty]{\text{law}} \delta_0,$$

with δ_0 having a non continuous cumulative distribution function. And we have for all $n \in \mathbb{N}^*$:

$$d_{\mathrm{Kol}}\left(\delta_{\frac{1}{n}}, \delta_0\right) = \sup_{z\in\mathbb{R}} \left|\mathbf{1}_{\left\{z\leqslant\frac{1}{n}\right\}} - \mathbf{1}_{\{z\leqslant 0\}}\right| = 1.$$

So the distance cannot go to zero.

I.2.3 Wasserstein distance

Definition I.5

We set $d \in \mathbb{N}^*$ and

$$\mathcal{H} = \mathsf{Lip}(1),$$

the set of all function $h : \mathbb{R}^d \longrightarrow \mathbb{R}$ which are 1-lipschitzian. We call *Wasserstein distance* the distance in law associated. We note it d_{W}.

Proposition I.4 : Stein's bound for Wasserstein distance

i. Let $h : \mathbb{R} \longrightarrow \mathbb{R}$ a K-Lipschitz function. Then $f_h \in \mathcal{C}^1(\mathbb{R})$, we have, if $N \sim \mathcal{N}(0,1)$:

$$f_h(x) = -\int_0^{+\infty} \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \mathbb{E}\left[Nh\left(e^{-t}x + \sqrt{1-e^{-2t}}N\right)\right] \,\mathrm{d}t.$$

Finally, we have $\|f_h'\|_\infty \leqslant K\sqrt{\frac{2}{\pi}}$.

ii. We note $\mathcal{F}_{\mathrm{W}}$ the set of function C^1 on $\mathbb{R}$ such that their derivative is bounded by $\sqrt{\frac{2}{\pi}}$. For $F \in L^2$ and $N \sim \mathcal{N}(0,1)$, we have

$$d_{\mathrm{W}}(F, N) \leqslant \sup_{F\in\mathcal{F}_W} \left|\mathbb{E}\left[f'(F) - Ff(F)\right]\right|.$$

Proof : We could simply check that the left hand side satisfies the Stein's equation and the limit condition, like do the reference. We will actually derive it, by doing before an analysis step. The idea is to use Malliavin calculus, in dimension 1.

- Since

$$f_h(x) = e^{\frac{x^2}{2}} \int_{-\infty}^x (h(y) - \mathbb{E}[h(N)])\, e^{\frac{-y^2}{2}} \,\mathrm{d}y,$$

and since h is continuous, f_h is indeed C^1.

- With the notation of the chapter **I** of the part about Malliavin calculus, if we consider F_h a antiderivative of f_h, then the Stein's equation writes, if we suppose $F_h \in \mathcal{S}$:

$$LF_h(x) = h(x) - \mathbb{E}[h(N)]$$

But, by the tricks we used in this chapter, we have

$$h(x) - \mathbb{E}[h(N)] = P_0h(x) - P_\infty h(x) = -\int_0^{+\infty} \frac{\mathrm{d}P_t h}{\mathrm{d}t}(x) \,\mathrm{d}t.$$

By relation between derivative of P_t and L :

$$h(x) - \mathbb{E}[h(N)] = -\int_0^{+\infty} LP_t h(x) \,\mathrm{d}t.$$

Hence, a possible definition of F_h would be

$$F_h(x) = -\int_0^{+\infty} P_t h(x) \,\mathrm{d}t.$$

However, the convergence of this integral is not guarantied under our hypothesis. That's why we will introduce $\mathbb{E}[h(N)] = P_\infty h$, and define

$$F_h(x) \overset{\text{def.}}{=} \int_0^{+\infty} \{\mathbb{E}[h(N)] - P_t h(x)\} \,\mathrm{d}t.$$

Let us show that F_h is well-defined. Let

$$\text{(A)} \overset{\text{def.}}{=} \int_0^{+\infty} \mathbb{E}\left[\left|h(N) - h\left(e^{-t}x + \sqrt{1-e^{-2t}}N\right)\right|\right].$$

Then, since h is K-Lipschitz, we have

$$\text{(A)} \leqslant K\int_0^{+\infty} \mathbb{E}\left[\left|\left(1-\sqrt{1-e^{-2t}}\right)N - e^{-t}x\right|\right].$$

We cut in two :

$$\text{(A)} \leqslant K|x| \int_0^{+\infty} e^{-t} \,\mathrm{d}t + K\mathbb{E}[|N|] \int_0^{+\infty} \left(1-\sqrt{1-e^{-2t}}\right) \,\mathrm{d}t.$$

We can up-bound the second integral, since

$$1-\sqrt{1-e^{-2t}} = \frac{e^{-2t}}{1+\sqrt{1-e^{-2t}}} \leqslant e^{-2t}.$$

Finally,

$$\text{(A)} \leqslant K|x| + \frac{K}{2}\sqrt{\frac{\pi}{2}} < +\infty,$$

so F_h is well-defined. Hence, we just need to differentiate to define f_h :

$$f_h(x) \overset{\text{def.}}{=} \int_0^{+\infty} e^{-t}\mathbb{E}\left[h'\left(e^{-t}x+\sqrt{1-e^{-2t}}N\right)\right].$$

We can indeed define h' since h is Lipschitz. Then, by Stein's lemma for the map

$$y \longmapsto h\left(e^{-t}x+\sqrt{1-e^{-2t}}y\right),$$

we have our expected definition :

$$f_h(x) \overset{\text{def.}}{=} \int_0^{+\infty} \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\mathbb{E}\left[Nh\left(e^{-t}x+\sqrt{1-e^{-2t}}N\right)\right].$$

We just need to do the synthesis step to conclude.

• With such a definition, we just need to show that F_h is indeed in the domain of L. To do this, we will show that F_h is in fact in $\mathcal{S}$. To stay consistent, we write $\tilde{f}_h$ the integral function we just defined before, and $\tilde{F}_h$ the anti-derivative. $\tilde{f}_h$ is C^1, by dominated convergence, since :

$$\mathbb{E}\left[\left|\frac{e^{-2t}}{\sqrt{1-e^{-2t}}}h'\left(e^{-t}x+\sqrt{1-e^{-2t}}N\right)\right|\right] \leqslant K\sqrt{\frac{2}{\pi}}\frac{e^{-2t}}{\sqrt{1-e^{-2t}}},$$

which is integrable on $\mathbb{R}_+$. We have as a consequence by differentiating the integral :

$$\left|\tilde{f}'(x)\right| \leqslant K\sqrt{\frac{2}{\pi}}$$

So $\tilde{f}_h'$ is bounded by $K\sqrt{\frac{2}{\pi}}$, and so $\tilde{F}_h \in D(L)$. As a consequence, $\tilde{f}_h$ satisfies the Stein's equation.

• To be sure that $\tilde{f}_h$ is indeed f_h, we have one last thing to check : the fact that

$$e^{\frac{-x^2}{2}}\tilde{f}_h(x) \xrightarrow[x\to\pm\infty]{} 0.$$

But, this property is true by the bound we just give for $\tilde{f}_h'$:

$$|\tilde{f}_h(x)| \leqslant |\tilde{f}_h(0)| + K\sqrt{\frac{2}{\pi}}|x|.$$

So we finally have $\tilde{f}_h = f_h$ like expected. □

I.2.4 An application for centered Gaussian

We can apply those bounds here for Gaussian random variables.

Proposition I.5 : Estimations for Gaussian random variables

Let $\sigma_1, \sigma_2 > 0$, $N_1 \sim \mathcal{N}(0,\sigma_1^2)$ and $N_2 \sim \mathcal{N}(0,\sigma_2^2)$. Then,

(i) For total variation distance :

$$d_{\mathrm{TV}}(N_1,N_2) \leqslant \frac{2}{\max\{\sigma_1^2,\sigma_2^2\}}\left|\sigma_1^2-\sigma_2^2\right|.$$

(ii) For Kolmogorov distance :

$$d_{\mathrm{Kol}}(N_1,N_2) \leqslant \frac{1}{\max\{\sigma_1^2,\sigma_2^2\}}\left|\sigma_1^2-\sigma_2^2\right|.$$

(iii) For Wasserstein distance :

$$d_{\mathrm{W}}(N_1,N_2) \leqslant \sqrt{\frac{2}{\pi}}\frac{1}{\max\{\sigma_1,\sigma_2\}}\left|\sigma_1^2-\sigma_2^2\right|.$$

Proof : Let $N_0 \sim \mathcal{N}(0,1)$, and we suppose that $\sigma_1 \leqslant \sigma_2$. The idea is to use the fact that $N_2 \overset{\text{law}}{=} \sigma_2 N_0$, and use Stein's bound for a Gaussian.

(i) Like we just told, we have

$$d_{\mathrm{TV}}(N_1,N_2) = d_{\mathrm{TV}}(N_1,\sigma_2 N_0).$$

Since the map $B \longmapsto \sigma_2 B$ from $\mathcal{B}(\mathbb{R})$ to itself is a bijection, we have

$$d_{\mathrm{TV}}(N_1,N_2) = d_{\mathrm{TV}}\left(\frac{N_1}{\sigma_2},N_0\right).$$

Now, we just use the Stein's bound for total variation, and compute thanks to Stein's lemma all the quantities that appear. We have

$$d_{\mathrm{TV}}(N_1,N_2) \leqslant \sup_{f\in\mathcal{F}_{\mathrm{TV}}}\left|\mathbb{E}\left[f'\left(\frac{N_1}{\sigma_2}\right)-\frac{N_1}{\sigma_2}f\left(\frac{N_1}{\sigma_2}\right)\right]\right|.$$

Let $f \in \mathcal{F}_{\mathrm{TV}}$. We have $\frac{N_1}{\sigma_2} \sim \mathcal{N}\left(0,\left(\frac{\sigma_1}{\sigma_2}\right)^2\right)$, and

$$\mathbb{E}\left[\left|f'\left(\frac{N_1}{\sigma_2}\right)\right|\right] \leqslant 2 < +\infty.$$

By Stein's lemma :

$$\mathbb{E}\left[f'\left(\frac{N_1}{\sigma_2}\right)\right] = \mathbb{E}\left[\frac{\sigma_2 N_1}{\sigma_1^2}f\left(\frac{N_1}{\sigma_2}\right)\right].$$

Since we know that f' is bounded by 2, we express everything only in terms of f'. Here, we get :

$$d_{\mathrm{TV}}(N_1,N_2) \leqslant \sup_{f\in\mathcal{F}_{\mathrm{TV}}}\left|\mathbb{E}\left[\left(1-\left(\frac{\sigma_1}{\sigma_2}\right)^2\right)f'\left(\frac{N_1}{\sigma_2}\right)\right]\right|.$$

And so,

$$d_{\mathrm{TV}}(N_1,N_2) \leqslant 2\left(1-\left(\frac{\sigma_1}{\sigma_2}\right)^2\right).$$

(ii) This is essentially the same argument. We have again :

$$d_{\mathrm{TV}}(N_1, N_2) \leqslant \sup_{f \in \mathcal{F}_{\mathrm{Kol}}} \left| \mathbb{E}\left[f'\left(\frac{N_1}{\sigma_2}\right) - \frac{N_1}{\sigma_2} f\left(\frac{N_1}{\sigma_2}\right) \right] \right|.$$

This time, f' is bounded by 1, so it gives :

$$d_{\mathrm{TV}}(N_1, N_2) \leqslant \left(1 - \left(\frac{\sigma_1}{\sigma_2}\right)^2\right).$$

(iii) For Wasserstein distance, it is not true that

$$d_{\mathrm{W}}(F, \alpha G) = d_{\mathrm{W}}\left(\frac{1}{\alpha} F, G\right),$$

since the map $h \longmapsto h(\alpha \cdot)$ does not takes values in Lip(1), but in Lip(α). It does not matter, we can do nevertheless the computation in this case. Hence, by our little argumentation, we have

$$d_{\mathrm{W}}(N_1, N_2) \leqslant \sup_{g \in \mathrm{Lip}(\sigma_2)} \left| \mathbb{E}\left[g\left(\frac{N_1}{\sigma_2}\right) - g(N_0) \right] \right|.$$

By the proposition about the Stein's bound for Wasserstein distance, if h is K-Lipschitz, then f_h is C^1 and $K\sqrt{\frac{\pi}{2}}$-Lipschitz. It follows that by Stein's equation :

$$d_{\mathrm{W}}(N_1, N_2) \leqslant \sup_{f \in \mathcal{C}^1 \cap \mathrm{Lip}\left(\sqrt{\frac{2}{\pi}}\sigma_2\right)} \left| \mathbb{E}\left[f'\left(\frac{N_1}{\sigma_2}\right) - \frac{N_1}{\sigma_2} f\left(\frac{N_1}{\sigma_2}\right) \right] \right|.$$

By Stein's lemma,

$$d_{\mathrm{W}}(N_1, N_2) \leqslant \sup_{f \in \mathcal{C}^1 \cap \mathrm{Lip}\left(\sqrt{\frac{2}{\pi}}\sigma_2\right)} \left| \left(1 - \left(\frac{\sigma_1}{\sigma_2}\right)^2\right) \mathbb{E}\left[f'\left(\frac{N_1}{\sigma_2}\right) \right] \right|.$$

And so,

$$d_{\mathrm{W}}(N_1, N_2) \leqslant \sigma_2 \sqrt{\frac{2}{\pi}} \left(1 - \left(\frac{\sigma_1}{\sigma_2}\right)^2\right).$$

This show our estimation. □

In fact, we can explicitly express the Kolmogorov distance for two centered Gaussian variables.

Proposition I.6 : Kolmogorov distance for Gaussian variables

Let $\sigma_1, \sigma_2 > 0$. Then

$$d_{\mathrm{Kol}}(\mathcal{N}(0,\sigma_1^2), \mathcal{N}(0,\sigma_2^2)) = \Phi\left(\frac{\sigma_2}{\sqrt{\sigma_2^2 - \sigma_1^2}} \sqrt{\ln\left(\frac{\sigma_2^2}{\sigma_1^2}\right)}\right) - \Phi\left(\frac{\sigma_1}{\sqrt{\sigma_2^2 - \sigma_1^2}} \sqrt{\ln\left(\frac{\sigma_2^2}{\sigma_1^2}\right)}\right),$$

where $\Phi(x) = \mathbb{P}(N \leqslant x)$ is the cumulative distribution function of $\mathcal{N}(0,1)$.

Proof : We suppose that $\sigma_1 \leqslant \sigma_2$. Then if we define the function

$$f : x \longmapsto \mathbb{P}(N_1 \leqslant x) - \mathbb{P}(N_2 \leqslant x),$$

then f is differentiable on $\mathbb{R}$. Its derivative is given by

$$\sqrt{2\pi} f'(x) = \frac{1}{\sigma_1} e^{\frac{-x^2}{2\sigma_1^2}} - \frac{1}{\sigma_2} e^{\frac{-x^2}{2\sigma_2^2}}.$$

Then, $f'(x)$ is null if and only if (we suppose that $\sigma_1 < \sigma_2$) :

$$x^2 = \frac{2\sigma_1^2\sigma_2^2}{\sigma_2^2 - \sigma_1^2}.$$

Finally, let us notice that

$$d_{\mathrm{Kol}}(N_1, N_2) = \max\left\{ \sup_{x \geqslant 0} f(x), -\sup_{x \leqslant 0} f(x) \right\}.$$

Since f is odd, we have finally that the supremum is taken on $\mathbb{R}_+$, and so the value of x is known. □

I.3 Berry-Esséen theorem and Central Limit Theorem

We want to precise the convergence in law for the CLT. The Stein's method allows us to do it for the Kolmogorov's distance, if we add the hypothesis of being in L^3.

Theorem I.2 : Berry-Esséen theorem

Let $(X_i)_i$ a sequence of real random variables L^3 which are independent and identically distributed. We suppose that $\mathbb{E}[X_i] = 0$ and $\mathbb{E}[X_i^2] = 1$. Finally, we set

$$S_n \overset{\text{def.}}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i.$$

Then,

$$d_{\mathrm{Kol}}(S_n, \mathcal{N}(0,1)) \leqslant \frac{C\mathbb{E}\left[|X_1|^3\right]}{\sqrt{n}}.$$

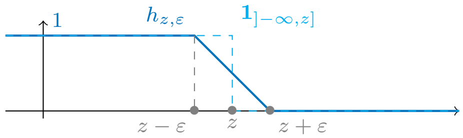
Since the convergence in Kolmogorov sense implies the convergence in law, this proves the Central Limit Theorem.

Proof : The first idea would be to take $h = \mathbf{1}_{]-\infty,z]}$, for $z \in \mathbb{R}$ and directly apply a Stein's method. With a few computations quite close to those we will perform later, we would have to estimate

$$\left|\mathbb{E}\left[h(S_n) - h(S_n^i)\right]\right|,$$

where $S_n^i = S_n - \frac{1}{\sqrt{n}} X_i \overset{\text{law}}{=} S_{n-1}$. Since h is not absolutely continuous, a satisfying estimation is quite hard to derive. Here, "satisfying" means finding an estimation in the form $O\left(\frac{1}{n^\alpha}\right)$, $\alpha > 0$. For our theorem, $\alpha = \frac{1}{2}$ would do the trick. To avoid this problem, we approach h by an absolute continuous function, and compute everything with this approximation. We will show that we can find the result for h after all. Here's the plan.

1. We set $z \in \mathbb{R}$, and $\varepsilon > 0$. We consider $h_{z,\varepsilon}$ like :



Then, we have, for all $n \geqslant 1$ and if $N \sim \mathcal{N}(0,1)$:

$$d_{\mathrm{Kol}}(S_n, N) \leqslant \frac{4\varepsilon}{\sqrt{2\pi}} + \sup_{z \in \mathbb{R}} \left|\mathbb{E}\left[h_{z,\varepsilon}(S_n) - h_{z,\varepsilon}(N)\right]\right|.$$

2. Let

$$C_n \overset{\text{def.}}{=} \sup_{(X_1,\cdots,X_n)\in A_n} \frac{\sqrt{n}\, d_{\mathrm{Kol}}(S_n, N)}{\mathbb{E}\left[|X_1|^3\right]},$$

where

$$A_n = \left\{ (X_1,\cdots,X_n) \in L^3(\Omega)^n \;\middle|\; \begin{array}{rl} X_1,\cdots,X_n & \text{iid} \\ \mathbb{E}[X_1] & = 0 \\ \mathbb{E}[X_1^2] & = 1 \end{array} \right\}$$

Then $C_n \leqslant \sqrt{n}$.

3. By a Stein's method, we have the estimation for **1.** :

$$d_{\mathrm{Kol}}(S_n, N) \leqslant \frac{4\varepsilon}{\sqrt{2\pi}} + \frac{A\mathbb{E}\left[|X_1|^3\right]}{\sqrt{n}} + \frac{BC_{n-1}\mathbb{E}\left[|X_1|^3\right]}{\varepsilon n}.$$

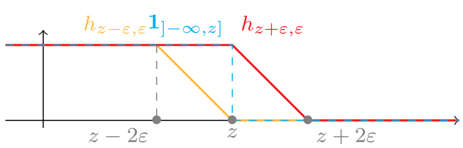
4. We optimize in ε, and we conclude that $C_n \leqslant \alpha\sqrt{C_{n-1}} + \beta$, and so that $(C_n)_n$ is bounded. This will conclude the proof.

Let us begin the proof of the Berry-Esséen theorem.

1. By definition of $h_{z,\varepsilon}$, we have the inequalities :

$$h_{z-\varepsilon,\varepsilon} \leqslant \mathbf{1}_{]-\infty,z]} \leqslant h_{z+\varepsilon,\varepsilon}.$$

Indeed :



So we have the inequalities :

$$\mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(S_n)\right] \leqslant \mathbb{P}(S_n \leqslant z) \leqslant \mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(S_n)\right].$$

• Let us try to show the same inequality for N, but with reversed roles. To do this, we want to estimate

$$\Delta_\varepsilon \overset{\text{def.}}{=} \mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(N)\right] - \mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(N)\right].$$

To do this, we use the expression of $h_{z,\varepsilon}$:

$$h_{z,\varepsilon}(x) = \mathbf{1}_{]-\infty,z-\varepsilon]}(x) + \left(\frac{z+\varepsilon-x}{2\varepsilon}\right)\mathbf{1}_{]z-\varepsilon,z+\varepsilon]}(x).$$

So, we have :

$$\begin{aligned} &\Delta_\varepsilon \\ = \quad & \mathbb{P}(N \leqslant z) - \mathbb{P}(N \leqslant z-\varepsilon) \\ + \quad & \frac{z+2\varepsilon}{2\varepsilon}\mathbb{P}(z \leqslant N \leqslant z+2\varepsilon) - \frac{z}{2\varepsilon}\mathbb{P}(z-2\varepsilon \leqslant N \leqslant z) \\ - \quad & \frac{1}{2\varepsilon}\mathbb{E}\left[N\mathbf{1}_{\{z \leqslant N \leqslant z+2\varepsilon\}}\right] + \frac{1}{2\varepsilon}\mathbb{E}\left[N\mathbf{1}_{\{z-2\varepsilon \leqslant N \leqslant z\}}\right]. \end{aligned}$$

We write as an integral and arrange it :

$$\begin{aligned} &\Delta_\varepsilon \\ = \quad & \int_{z-2\varepsilon}^{z+2\varepsilon} \rho(x)\,\mathrm{d}x \\ + \quad & \frac{1}{2\varepsilon}\left(\int_z^{z+\varepsilon}(z-x)\rho(x)\,\mathrm{d}x - \int_{z-2\varepsilon}^{z}(z-x)\rho(x)\,\mathrm{d}x\right), \end{aligned}$$

where ρ is the density of $\mathcal{N}(0,1)$. The terms in $z-x$ is lower than 2ε in absolute values. So :

$$|\Delta_\varepsilon| \leqslant 2\int_{z-2\varepsilon}^{z+2\varepsilon} \rho(x)\,\mathrm{d}x.$$

Since $|\rho| \leqslant \frac{1}{\sqrt{2\pi}}$, we finally have :

$$|\Delta_\varepsilon| \leqslant \frac{4\varepsilon}{\sqrt{2\pi}}.$$

• So we can conclude on the inequalities on $\mathbb{P}(N \leqslant z)$. We have

$$\mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(N)\right] \leqslant \mathbb{P}(N \leqslant z) \leqslant \mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(N)\right].$$

So, by in terms of Δ_ε :

$$\mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(N)\right] - \Delta_\varepsilon \leqslant \mathbb{P}(N \leqslant z) \leqslant \mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(N)\right] + \Delta_\varepsilon.$$

And by the work done before :

$$\mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(N)\right] - \frac{4\varepsilon}{\sqrt{2\pi}} \leqslant \mathbb{P}(N \leqslant z) \leqslant \mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(N)\right] + \frac{4\varepsilon}{\sqrt{2\pi}}.$$

Finally, we have

$$\begin{aligned} & \mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(S_n)\right] - \mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(S_n)\right] + \frac{4\varepsilon}{\sqrt{2\pi}} \\ \leqslant \quad & \mathbb{P}(S_n \leqslant z) - \mathbb{P}(N \leqslant z) \\ \leqslant \quad & \mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(S_n)\right] - \mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(S_n)\right] - \frac{4\varepsilon}{\sqrt{2\pi}}. \end{aligned}$$

So, we proved that

$$d_{\mathrm{Kol}}(S_n, N) \leqslant \sup_{z \in \mathbb{R}} \left|\mathbb{E}\left[h_{z+\varepsilon,\varepsilon}(S_n)\right] - \mathbb{E}\left[h_{z-\varepsilon,\varepsilon}(S_n)\right]\right| + \frac{4\varepsilon}{\sqrt{2\pi}}.$$

2. If $(X_1,\cdots,X_n) \in A_n$, then we have by Jensen's inequality :

$$\mathbb{E}\left[|X_1|^3\right] \leqslant \mathbb{E}\left[X_1^2\right]^{\frac{3}{2}} = 1.$$

Moreover, we have $d_{\mathrm{Kol}} \leqslant 1$. So, we have $C_n \leqslant \sqrt{n}$. This doesn't proof that $(C_n)_n$ is bounded, but this estimation will be helpful.

3. For more easy notations, we note $h = h_{z,\varepsilon}$ and $f = f_h$ the associated solution of Stein's equation. By the proposition about the estimation for Kolmogorov distance, f is C^1 and satisfies $\|f\|_\infty \leqslant \sqrt{\frac{\pi}{2}}$ and $\|f'\|_\infty \leqslant 2$. Then, by Stein's equation :

$$\mathbb{E}\left[h(S_n) - h(N)\right] = \mathbb{E}\left[f'(S_n) - S_n f(S_n)\right].$$

In terms of $(X_i)_i$:

$$\mathbb{E}\left[h(S_n) - h(N)\right] = \sum_{i=1}^n \mathbb{E}\left[\frac{f'(S_n)}{n} - \frac{X_i}{\sqrt{n}} f(S_n)\right].$$

We make appear with f the derivative f' thanks to the random variable

$$S_n^i \overset{\text{def.}}{=} S_n - \frac{X_i}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{j \neq i} X_j.$$

The idea is to find a recurrent inequality on C_n to prove that this sequence is bounded. We can have access to this with S_n^i which is equal in law to S_{n-1}, up to a factor. The rand variable S_n^i is independent of X_i, so

$$\mathbb{E}\left[X_i f\left(S_n^i\right)\right] = 0.$$

So, we have :

$$\mathbb{E}\left[h(S_n) - h(N)\right] = \sum_{i=1}^n \mathbb{E}\left[\frac{f'(S_n)}{n} - \frac{X_i}{\sqrt{n}}\left(f(S_n) - f\left(S_n^i\right)\right)\right].$$

There exists a random variable θ independent of $(X_i)_i$ such that $\theta \sim \mathcal{U}([0,1])$ and :

$$f(S_n) - f\left(S_n^i\right) = \frac{X_i}{\sqrt{n}} f'\left(S_n^i + \frac{\theta X_i}{\sqrt{n}}\right).$$

So, we have :

$$\mathbb{E}\left[h(S_n) - h(N)\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[f'(S_n) - X_i^2 f'\left(S_n^i + \frac{\theta X_i}{\sqrt{n}}\right)\right].$$

We now use the differential equation satisfied by f'. We note

$$\tilde{f}(x) \overset{\text{def.}}{=} x f(x)$$

Then, we have :

$$\begin{aligned} & \mathbb{E}\left[h(S_n) - h(N)\right] \\ = & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\tilde{f}(S_n) - X_i^2 \tilde{f}\left(S_n^i + \frac{\theta X_i}{\sqrt{n}}\right)\right] \\ + & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[h(S_n) - X_i^2 h\left(S_n^i + \frac{\theta X_i}{\sqrt{n}}\right)\right] \\ - & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[1 - X_i^2\right]\mathbb{E}[h(N)]. \end{aligned}$$

Since $\mathbb{E}[X_i^2] = 1$, the last term is zero. Since X_i is independent of S_n^i, we have

$$\mathbb{E}\left[X_i^2 \tilde{f}\left(S_n^i\right)\right] = \mathbb{E}\left[\tilde{f}\left(S_n^i\right)\right].$$

Same for h. We cut in four the expression :

$$\begin{aligned} & \mathbb{E}\left[h(S_n) - h(N)\right] \\ = & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\tilde{f}(S_n) - \tilde{f}\left(S_n^i\right)\right] \\ + & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[X_i^2\left(\tilde{f}\left(S_n^i\right) - \tilde{f}\left(S_n^i + \frac{\theta X_i}{\sqrt{n}}\right)\right)\right] \\ + & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[h(S_n) - h(S_n^i)\right] \\ + & \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[X_i^2\left(h(S_n^i) - h\left(S_n^i + \frac{\theta X_i}{\sqrt{n}}\right)\right)\right]. \end{aligned}$$

We write

$$\mathbb{E}\left[h(S_n) - h(N)\right] = \frac{1}{n}\sum_{i=1}^n \left\{(\mathsf{A}) + (\mathsf{B}) + (\mathsf{C}) + (\mathsf{D})\right\}.$$

We treat one by one each term.

A. We will the following inequality, using the bounds satisfied by f :

$$|\tilde{f}(x) - \tilde{f}(y)| \leqslant \sqrt{\frac{\pi}{2}}|x-y| + 2|y||x-y|.$$

Here, we get for (A) :

$$|(\mathsf{A})| \leqslant \sqrt{\frac{\pi}{2}}\mathbb{E}\left[|S_n - S_n^i|\right] + 2\mathbb{E}\left[\left|S_n^i\right|\left|S_n^i - S_n\right|\right].$$

By Cauchy-Schwarz on the second term :

$$|(\mathsf{A})| \leqslant \sqrt{\frac{\pi}{2n}}\mathbb{E}\left[|X_i|\right] + \frac{2}{\sqrt{n}}\mathbb{E}\left[\left|S_n^i\right|^2\right]^{\frac{1}{2}}\mathbb{E}\left[X_i^2\right]^{\frac{1}{2}}.$$

We have $\mathbb{E}[X_i^2] = 1$ and by triangular inequality, $\mathbb{E}[\left|S_n^i\right|^2]^{\frac{1}{2}} \leqslant 1$. We finally have the following bound for (A) :

$$|(\mathsf{A})| \leqslant \frac{1}{\sqrt{n}}\left(2 + \sqrt{\frac{\pi}{2}}\right).$$

B. We use the exact same inequality.

$$|(\mathsf{B})| \leqslant \sqrt{\frac{\pi}{2n}}\mathbb{E}\left[\theta\,|X_i|^3\right] + \frac{2}{\sqrt{n}}\mathbb{E}\left[\theta|X_i|^3|S_n^i|\right].$$

By mutual independence of θ, X_i, and S_n^i, we have :

$$|(\mathsf{B})| \leqslant \sqrt{\frac{\pi}{8n}}\mathbb{E}\left[|X_i|^3\right] + \frac{1}{\sqrt{n}}\mathbb{E}\left[|X_i|^3\right]\mathbb{E}[|S_n^i|].$$

And since $\mathbb{E}[|S_n^i|] \leqslant 1$, we have the inequality :

$$|(\mathsf{B})| \leqslant \frac{\mathbb{E}\left[|X_i|^3\right]}{\sqrt{n}}\left(1 + \sqrt{\frac{\pi}{8}}\right).$$

C. This estimation is the point that makes the use of $h_{z,\varepsilon}$ instead of directly use the indicator function, like we said in the beginning of the proof. Indeed, we have the equality :

$$h(y) - h(x) = (y-x)\int_0^1 h'(x + s(y-x))\,\mathrm{d}s.$$

The derivative of h is zero except on $[z-\varepsilon, z+\varepsilon]$ where it is equal to $\frac{-1}{2\varepsilon}$. We write as

$$h(y) - h(x) = \frac{y-x}{-2\varepsilon}\mathbb{E}\left[\mathbf{1}_{[z-\varepsilon,z+\varepsilon]}(x + \tilde{\theta}(y-x))\right],$$

where $\tilde{\theta} \sim \mathcal{U}([0,1])$. So, we have

$$|(\mathsf{C})| \leqslant \frac{\mathbb{E}[|X_i|]}{2\varepsilon\sqrt{n}}\mathbb{P}\left(z - \varepsilon - \frac{\tilde{\theta}X_i}{\sqrt{n}} \leqslant S_n^i \leqslant z + \varepsilon - \frac{\tilde{\theta}X_i}{\sqrt{n}}\right).$$

Since $\mathbb{E}[|X_i|] \leqslant 1$, we have the following inequality :

$$|(\mathsf{C})| \leqslant \frac{1}{2\varepsilon\sqrt{n}}\sup_{\substack{t\in[0,1]\\ y\in\mathbb{R}}}\mathbb{P}\left(z - \varepsilon - \frac{ty}{\sqrt{n}} \leqslant S_n^i \leqslant z + \varepsilon - \frac{ty}{\sqrt{n}}\right).$$

This supremum may look scary, but we can in fact estimate it in terms of the Kolmogorov distance of S_n^i with $\mathcal{N}(0,1)$. Knowing the fact that $S_n^i \overset{\text{law}}{=} \sqrt{\frac{n-1}{n}}S_{n-1}$, we could conclude. Let $N \sim \mathcal{N}(0,1)$. Let $a < b$. Then :

$$\begin{aligned} & \mathbb{P}(a \leqslant S_n^i \leqslant b) \\ = & \mathbb{P}\left(\sqrt{\frac{n}{n-1}}a \leqslant S_{n-1} \leqslant \sqrt{\frac{n}{n-1}}b\right) \\ - & \mathbb{P}\left(\sqrt{\frac{n}{n-1}}a \leqslant N \leqslant \sqrt{\frac{n}{n-1}}b\right) \\ + & \mathbb{P}\left(\sqrt{\frac{n}{n-1}}a \leqslant N \leqslant \sqrt{\frac{n}{n-1}}b\right) \end{aligned}$$

Since $N \sim \mathcal{N}(0,1)$, the third term is easily estimated. The two first can be estimated in terms of Kolmogorov distance :

$$\mathbb{P}(a \leqslant S_n^i \leqslant b) \leqslant d_{\mathrm{Kol}}(S_{n-1}, \mathcal{N}(0,1)) + \frac{b-a}{\sqrt{2\pi}}\sqrt{\frac{n}{n-1}}.$$

Going back to (C), we have for all $t \in [0,1]$ and $y \in \mathbb{R}$, and by definition of C_{n-1} :

$$\begin{aligned} &\mathbb{P}\left(z - \varepsilon - \frac{ty}{\sqrt{n}} \leqslant S_n^i \leqslant z + \varepsilon - \frac{ty}{\sqrt{n}}\right) \\ \leqslant\ & \frac{C_{n-1}\mathbb{E}\left[|X_i|^3\right]}{\sqrt{n-1}} + \frac{2\varepsilon}{\sqrt{2\pi}}\sqrt{\frac{n}{n-1}}. \end{aligned}$$

Hence, we have :

$$|(\mathsf{C})| \leqslant \frac{C_{n-1}\mathbb{E}\left[|X_1|^3\right]}{2\varepsilon\sqrt{n(n-1)}} + \frac{1}{\sqrt{2\pi(n-1)}}.$$

Finally, by using $n-1 \geqslant \frac{n}{2}$ and $\mathbb{E}\left[|X_1|^3\right] \leqslant \mathbb{E}\left[|X_1|^3\right]^2$, we have :

$$|(\mathsf{C})| \leqslant \frac{C_{n-1}\mathbb{E}\left[|X_1|^3\right]^2}{\sqrt{2}\varepsilon n} + \frac{1}{\sqrt{\pi n}}.$$

D. We use the same estimations as **C.**. Indeed, we have by the estimation on the increments of h (in terms of $\tilde{\theta}$) :

$$\begin{aligned} |(\mathsf{D})| \leqslant\ & \frac{1}{2\varepsilon\sqrt{n}}\mathbb{E}\left[\theta|X_i|^3\right] \\ & \cdot \mathbb{P}\left(z - \varepsilon - \frac{\tilde{\theta}\theta X_i}{\sqrt{n}} \leqslant S_n^i \leqslant z + \varepsilon - \frac{\tilde{\theta}\theta X_i}{\sqrt{n}}\right). \end{aligned}$$

So, we have by independence of θ with X_i :

$$\begin{aligned} |(\mathsf{D})| \leqslant\ & \frac{1}{4\varepsilon\sqrt{n}}\mathbb{E}\left[|X_i|^3\right] \\ & \cdot \sup_{\substack{t\in[0,1]\\ y\in\mathbb{R}}} \mathbb{P}\left(z - \varepsilon - \frac{ty}{\sqrt{n}} \leqslant S_n^i \leqslant z + \varepsilon - \frac{ty}{\sqrt{n}}\right). \end{aligned}$$

By the estimation of this supremum made in **C.**, we get :

$$|(\mathsf{D})| \leqslant \frac{\mathbb{E}\left[|X_i|^3\right]}{4\varepsilon\sqrt{n}}\left(\frac{C_{n-1}\mathbb{E}\left[|X_i|^3\right]}{\sqrt{n-1}} + \frac{2\varepsilon}{\sqrt{2\pi}}\sqrt{\frac{n}{n-1}}\right).$$

We write it as :

$$|(\mathsf{D})| \leqslant \frac{C_{n-1}\mathbb{E}\left[|X_i|^3\right]^2}{4\varepsilon\sqrt{n(n-1)}} + \frac{\mathbb{E}\left[|X_i|^3\right]}{\sqrt{8\pi(n-1)}}.$$

And conclude using $n-1 \geqslant \frac{n}{2}$:

$$|(\mathsf{D})| \leqslant \frac{C_{n-1}\mathbb{E}\left[|X_i|^3\right]^2}{2\sqrt{2}\varepsilon n} + \frac{\mathbb{E}\left[|X_i|^3\right]}{\sqrt{4\pi n}}.$$

• We can go back to the estimation of $\mathbb{E}[h(S_n) - h(N)]$. We estimated the four terms that appeared in the estimation we made. We write as increasing powers of $\frac{\mathbb{E}[|X_1|^3]}{\sqrt{n}}$.

$$\begin{aligned} & |\mathbb{E}\left[h(S_n) - h(N)\right]| \\ \leqslant\ & \frac{1}{\sqrt{n}}\left(2 + \sqrt{\frac{\pi}{2}} + \frac{1}{\sqrt{\pi}}\right) \\ +\ & \frac{\mathbb{E}[|X_1|^3]}{\sqrt{n}}\left(1 + \sqrt{\frac{\pi}{8}} + \frac{1}{\sqrt{4\pi}}\right) \\ +\ & \frac{C_{n-1}\mathbb{E}[|X_1|^3]^2}{\varepsilon n}\left(\frac{1}{\sqrt{2}} + \frac{1}{2\sqrt{2}}\right). \end{aligned}$$

We simplify the constants (we need them explicitly) by taking their upper integer part. We also use once again the fact that $1 \leqslant \mathbb{E}\left[|X_1|^3\right]$ to finally have the estimation :

$$|\mathbb{E}\left[h(S_n) - h(N)\right]| \leqslant \frac{6\mathbb{E}\left[|X_1|^3\right]}{\sqrt{n}} + \frac{2C_{n-1}\mathbb{E}\left[|X_1|^3\right]^2}{\varepsilon n}.$$

This estimation is independent of z. So, we have proved that

$$d_{\mathrm{Kol}}(S_n, \mathcal{N}(0,1)) \leqslant \frac{4\varepsilon}{\sqrt{2\pi}} + \frac{2C_{n-1}\mathbb{E}\left[|X_1|^3\right]^2}{n\varepsilon} + \frac{6\mathbb{E}\left[|X_1|^3\right]}{\sqrt{n}}.$$

4. The optimization in ε gives $\varepsilon = \sqrt{\frac{\sqrt{2\pi}C_{n-1}\mathbb{E}[|X_1|^3]^2}{2n}}$ and

$$d_{\mathrm{Kol}}(S_n, \mathcal{N}(0,1)) \leqslant 2\sqrt{\frac{8C_{n-1}\mathbb{E}\left[|X_1|^3\right]^2}{n\sqrt{2\pi}}} + \frac{6\mathbb{E}\left[|X_1|^3\right]}{\sqrt{n}}.$$

So,

$$\frac{\sqrt{n}d_{\mathrm{Kol}(S_n,\mathcal{N}(0,1))}}{\mathbb{E}\left[|X_1|^3\right]} \leqslant \sqrt{\frac{32C_{n-1}}{\sqrt{2\pi}}} + 6 \leqslant 4\sqrt{C_{n-1}} + 6.$$

We consequently have the recurrent inequality :

$$C_n \leqslant 4\sqrt{C_{n-1}} + 6.$$

To conclude, we find the fixed point on $\mathbb{R}_+$ of $x \longmapsto 4\sqrt{x} + 6$. Here, this fixed point is lower than 21. We conclude by induction that $C_n \leqslant 21$, for all $n \in \mathbb{N}$. Indeed, $C_1 \leqslant \sqrt{1}$ by **2.** (that's why we needed the constant explicitly, it was to prove the initialisation part of the induction). So, $(C_n)_n$ is bounded, and we conclude that for all $n \in \mathbb{N}$:

$$d_{\mathrm{Kol}}(S_n, \mathcal{N}(0,1)) \leqslant \frac{21\mathbb{E}\left[|X_1|^3\right]}{\sqrt{n}}.$$

This proves the theorem. □

I.4 A discrete version : Chen-Stein lemma and bounds

We did all our theory around the centered reduced Gaussian random variables. But, we can do the same in a discrete case, for Poisson random variables.

Lemma I.2 : Chen-Stein's lemma

Let X a random variable taking values in $\mathbb{N}$ and $\lambda > 0$. Then the following assertions are equivalent :

(i) X follows $\mathcal{P}(\lambda)$;

(ii) For every bounded function $f : \mathbb{N} \longrightarrow \mathbb{R}$,

$$\mathbb{E}[Xf(X)] = \lambda\mathbb{E}[f(X+1)].$$

To prove it, we will use a new method we could apply before. We will directly use the Stein's equation. Here's the discrete version.

Definition I.6

Let $A \subset \mathbb{N}$, $\lambda > 0$, and $Z \sim \mathcal{P}(\lambda)$. We call *Stein-Chen's equation* the functional equality

$$\lambda f_A(k+1) - k f_A(k) = \mathbf{1}_A(k) - \mathbb{P}(Z \in A),$$

where the unknown is $f_A : \mathbb{N} \longrightarrow \mathbb{R}$ satisfies the initial condition $f_A(0) = 0$.

Like in the Gaussian case, we can express a solution to this equation.

Proposition I.7 : Solution to Stein-Chan's equation

The Stein-Chen's equation admits an unique solution given by :

$$f_A(k+1) = \frac{\mathbb{P}(Z \leqslant k, Z \in A) - \mathbb{P}(Z \leqslant k)\mathbb{P}(Z \in A)}{\lambda \mathbb{P}(Z = k)}.$$

Moreover, the solution is bounded :

$$\sup_{A \subset \mathbb{N}} \sup_{k \in \mathbb{N}} |f_A(k)| \leqslant \frac{e^\lambda}{\lambda}.$$

Proof : We could prove it by induction, but it would mean that we already knew the answer. Let us derive it. To do this, we introduce a sequence $(\psi_n)_{n \geqslant 1}$, with $\psi_1 \neq 0$ such that

$$\forall k \in \mathbb{N}^*, \psi_{k+1} = \frac{\lambda}{k}\psi_k.$$

Such ψ_n is given by

$$\psi_n = \psi_1 \frac{\lambda^{n-1}}{(n-1)!}.$$

So, taking $\psi_n = \mathbb{P}(Z = n-1)$ do the trick. Thanks to this sequence, we have for our functional equation :

$$f_A(k+1) - \frac{\psi_k}{\psi_{k+1}} f_A(k) = \frac{\mathbf{1}_A(k) - \mathbb{P}(Z \in A)}{\lambda}.$$

So, if we introduce $\Delta u_n = u_{n+1} - u_n$, we have :

$$\Delta(f_A(k)\psi_k) = \frac{\psi_{k+1}\left(\mathbf{1}_A(k) - \mathbb{P}(Z \in A)\right)}{\lambda}.$$

So, if we sum, we have :

$$f_A(k+1)\psi_{k+1} - f_A(0)\psi_0 = \frac{1}{\lambda}\sum_{j=0}^{k} \psi_{j+1}\left(\mathbf{1}_A(j) - \mathbb{P}(Z \in A)\right),$$

where ψ_0 is arbitrary. Since $f_A(0) = 0$, we are really close to our result, up to the computation of the sum. We have :

$$\begin{aligned} & \sum_{j=0}^{k} \psi_{j+1}\left(\mathbf{1}_A(j) - \mathbb{P}(Z \in A)\right) \\ = & \sum_{j=0}^{k} \mathbb{P}(Z = j)\mathbf{1}_A(j) \\ - & \sum_{j=0}^{k} \mathbb{P}(Z = j)\mathbb{P}(Z \in A). \end{aligned}$$

Hence,

$$\begin{aligned} & \sum_{j=0}^{k} \psi_{j+1}\left(\mathbf{1}_A(j) - \mathbb{P}(Z \in A)\right) \\ = \; & \mathbb{P}(Z \leqslant k, Z \in A) - \mathbb{P}(Z \leqslant k)\mathbb{P}(Z \in A). \end{aligned}$$

And we have our expression for $f_A(k+1)$. Let us show that f_A is bounded. To do this, we use the fact that

$$\mathbb{P}(Z \leqslant k, Z \in A) \leqslant \mathbb{P}(Z \in A).$$

We have thanks to this :

$$|f_A(k+1)| \leqslant \frac{\mathbb{P}(Z \in A)\mathbb{P}(Z \geqslant k+1)}{\lambda \mathbb{P}(Z = k)}.$$

But, the sequence $\left(\frac{\mathbb{P}(Z \geqslant k+1)}{\mathbb{P}(Z = k)}\right)_k$ is decreasing. We prove it later on the following lemma by expending the probabilities. So, we have

$$|f_A(k+1)| \leqslant \frac{e^\lambda}{\lambda}.$$

This concludes the proof. □

We have everything to prove the Chen-Stein's lemma.

Proof of the lemma : **[$\Longrightarrow$]** If X follows the law of Poisson, and $f : \mathbb{N} \longrightarrow \mathbb{R}$ is bounded then

$$\mathbb{E}[Xf(X)] = \sum_{n=0}^{+\infty} n f(n) e^{-\lambda}\frac{\lambda^n}{n!}.$$

The term $n = 0$ is zero, and for $n \geqslant 1$, the factor "n" simplifies with the factorial. We hence have :

$$\mathbb{E}[Xf(X)] = \sum_{n=0}^{+\infty} f(n+1)e^{-\lambda}\frac{\lambda^{n+1}}{n!}.$$

Meaning that

$$\mathbb{E}[Xf(X)] = \lambda\mathbb{E}[f(X+1)].$$

[$\Longleftarrow$] We set $Z \sim \mathcal{P}(\lambda)$, $A \subset \mathbb{N}$ and f_A the solution of the Chen-Stein's equation. We suppose that for every $f : \mathbb{N} \longrightarrow \mathbb{R}$ bounded :

$$\mathbb{E}[Xf(X)] = \lambda\mathbb{E}[f(X+1)].$$

Then, if we apply it to f_A :

$$\mathbb{E}[\lambda f_A(X+1) - Xf_A(X)] = 0.$$

By Chen-Stein's equation :

$$\mathbb{E}[\mathbf{1}_A(X) - \mathbb{P}(Z \in A)] = 0.$$

In other words, for any $A \subset \mathbb{N}$, we have

$$\mathbb{P}(Z \in A) = \mathbb{P}(X \in A),$$

this means that $X \overset{\text{law}}{=} Z$, so $X \sim \mathcal{P}(\lambda)$. □

Thanks to this, we are able to prove a Chen-Stein's bound. Before that, a lemma which will be helpful for the next proposition.

Lemma I.3 : Difference of Chen-Stein's solution

Let $j \in \mathbb{N}$. Then, for all $k \in \mathbb{N}^*$:

$$f_{\{j\}}(k+1) - f_{\{j\}}(k) > 0 \iff j = k.$$

Proof of the lemma : We treat every possible cases. We have

$$\begin{aligned}\Delta_j(k) &\overset{\text{def.}}{=} f_{\{j\}}(k+1) - f_{\{j\}}(k) \\ &= \frac{\mathbb{P}(Z \leqslant k, Z=j) - \mathbb{P}(Z \leqslant k)\mathbb{P}(Z=j)}{\lambda\mathbb{P}(Z=k)} \\ &\quad - \frac{\mathbb{P}(Z \leqslant k-1, Z=j) - \mathbb{P}(Z \leqslant k-1)\mathbb{P}(Z=j)}{\lambda\mathbb{P}(Z=k-1)}.\end{aligned}$$

- If $j < k$, then

$$\mathbb{P}(Z \leqslant k, Z=j) = \mathbb{P}(Z \leqslant k)$$

and

$$\mathbb{P}(Z \leqslant k-1, Z=j) = \mathbb{P}(Z=j).$$

It yields to

$$\Delta_j(k) = \frac{\mathbb{P}(Z=j)}{\lambda}\left(\frac{\mathbb{P}(Z \geqslant k+1)}{\mathbb{P}(Z=k)} - \frac{\mathbb{P}(Z \geqslant k)}{\mathbb{P}(Z=k-1)}\right).$$

Let us prove that the second factor is negative. We call it (A) :

$$\text{(A)} \overset{\text{def.}}{=} \frac{\mathbb{P}(Z \geqslant k+1)}{\mathbb{P}(Z=k)} - \frac{\mathbb{P}(Z \geqslant k)}{\mathbb{P}(Z=k-1)}.$$

We expend in series the probabilities that appear :

$$\text{(A)} = \frac{k!}{\lambda^k}\sum_{j=k+1}^{+\infty}\frac{\lambda^j}{j!} - \frac{(k-1)!}{\lambda^{k-1}}\sum_{j=k}^{+\infty}\frac{\lambda^j}{j!}.$$

By indexing the sums, we get :

$$\text{(A)} = \sum_{j=1}^{+\infty}\lambda^j\left(\frac{k!}{(k+j)!} - \frac{(k-1)!}{(j+k-1)!}\right).$$

But :

$$\frac{k!}{(k+j)!} = \frac{1}{(k+1)(k+2)\cdots(k+j)}.$$

So :

$$\frac{k!}{(k+j)!} < \frac{1}{k(k+1)\cdots(k+j-1)} = \frac{(k-1)!}{(j+k-1)!}.$$

So (A) < 0, and we proved that

$$f_{\{j\}}(k+1) - f_{\{j\}}(k) < 0.$$

- If $j > k$, then

$$\mathbb{P}(Z \leqslant k, Z=j) = \mathbb{P}(Z \leqslant k-1, Z=j) = 0.$$

This time, we have

$$\Delta_j(k) = \frac{\mathbb{P}(Z=j)}{\lambda}\left(\frac{\mathbb{P}(Z \leqslant k-1)}{\mathbb{P}(Z=k-1)} - \frac{\mathbb{P}(Z \leqslant k)}{\mathbb{P}(Z=k)}\right).$$

We cannot unfortunately use the previous case since we don't know if $\mathbb{P}(Z=k) < \mathbb{P}(Z=k-1)$, it depends on the position of k with respect to λ. So, we will just repeat the same argument to conclude of the negativity of the second factor :

$$\text{(B)} \overset{\text{def.}}{=} \frac{\mathbb{P}(Z \leqslant k-1)}{\mathbb{P}(Z=k-1)} - \frac{\mathbb{P}(Z \leqslant k)}{\mathbb{P}(Z=k)}.$$

We have by expanding the terms :

$$\text{(B)} = \frac{(k-1)!}{\lambda^{k-1}}\sum_{j=0}^{k-1}\frac{\lambda^j}{j!} - \frac{k!}{\lambda^k}\sum_{j=0}^{k}\frac{\lambda^j}{j!}.$$

We put the term in $j=0$ of the second sum aside, and index the second sum to have the familiar sum :

$$\text{(B)} = \sum_{j=1}^{k}\frac{1}{\lambda^{k-j}}\left[\frac{(k-1)!}{(j-1)!} - \frac{k!}{j!}\right] - \frac{k!}{\lambda^k} - \frac{k!}{\lambda^k}.$$

It is familiar because it looks like the opposite of the sum (A), but it is not because if the indexes. (B) is indeed negative since :

$$\text{(B)} = \sum_{j=1}^{k}\frac{k!}{\lambda^{k-j}}\left[\frac{1}{k(j-1)!} - \frac{1}{j!}\right] - \frac{k!}{\lambda^k},$$

and we sum for $j \leqslant k$, so

$$\text{(B)} \leqslant \sum_{j=1}^{k}\frac{k!}{\lambda^{k-j}}\left[\frac{1}{j(j-1)!} - \frac{1}{j!}\right] - \frac{k!}{\lambda^k} = \frac{-k!}{\lambda^k},$$

so (B) < 0, and $\Delta_j(k) < 0$.

- If $j = k$, we prove that $\Delta_j(k) > 0$. We have

$\mathbb{P}(Z \leqslant k, Z=k) = \mathbb{P}(Z=k)$ and $\mathbb{P}(Z \leqslant k-1, Z=k) = 0$.

So,

$$\Delta_k(k) = \frac{1 - \mathbb{P}(Z \leqslant k)}{\lambda} + \frac{\mathbb{P}(Z=k)\mathbb{P}(Z \leqslant k-1)}{\lambda\mathbb{P}(Z=k-1)}.$$

We can simplify using the fact that $\mathbb{P}(Z=k) = \frac{\lambda}{k}\mathbb{P}(Z=k-1)$, so

$$\Delta_k(k) = \frac{1 - \mathbb{P}(Z \leqslant k)}{\lambda} + \frac{\mathbb{P}(Z \leqslant k-1)}{k} > 0.$$

This concludes the proof. □

Proposition I.8 : Chen-Stein's bound for total variation distance

i. Let $\lambda > 0$, and $A \subset \mathbb{N}$. Then $f_A \in \Psi(\lambda)$, where

$$\Psi(\lambda) = \left\{ f : \mathbb{N} \longrightarrow \mathbb{R} \;\middle|\; \begin{array}{rcl} \forall i, j \in \mathbb{N}, |f(i) - f(j)| & \leqslant & \frac{1-e^{-\lambda}}{\lambda}|i-j| \\ \forall k \in \mathbb{N}, |f(k)| & \leqslant & \frac{e^{\lambda}}{\lambda} \end{array} \right\}.$$

ii. Let X a random variable and $Z \sim \mathcal{P}(\lambda)$. Then :

$$\sup_{A \subset \mathbb{N}} |\mathbb{P}(X \in A) - \mathbb{P}(Z \in A)| \leqslant \sup_{f \in \Psi(\lambda)} |\mathbb{E}\left[\lambda f(X+1) - X f(X)\right]|.$$

Proof : **i.** Let us see first that if $A \cap B = \varnothing$ then

$$f_{A \sqcup B} = f_A + f_B.$$

As a consequence, we can write for every finite $A \subset \mathbb{N}$:

$$f_A = \sum_{a \in A} f_{\{a\}}.$$

• By the lemma, we have for $a \neq k$, $f_{\{a\}}(k+1) - f_{\{a\}}(k) \leqslant 0$, so that we can consider the sum $\sum_{\substack{a \in A \\ k \neq a}} f_{\{a\}}(k+1) - f_{\{a\}}(k)$, which could be infinite. Hence, we have by additivity, and positivity of $f_{\{k\}}(k+1) - f_{\{k\}}(k)$, whether $k \in A$ or not :

$$f_A(k+1) - f_A(k) \leqslant \sum_{\substack{a \in A \\ k \neq a}} f_{\{a\}}(k+1) - f_{\{a\}}(k) + f_{\{k\}}(k+1) - f_{\{k\}}(k)$$

By negativity of the first term :

$$f_A(k+1) - f_A(k) \leqslant f_{\{k\}}(k+1) - f_{\{k\}}(k).$$

By the computation of the lemma :

$$f_A(k+1) - f_A(k) \leqslant \frac{\mathbb{P}(Z \geqslant k+1)}{\lambda} + \frac{\mathbb{P}(Z \leqslant k-1)}{k}.$$

And we expand the second sum :

$$f_A(k+1) - f_A(k) \leqslant \frac{\mathbb{P}(Z \geqslant k+1)}{\lambda} + \frac{1}{k} \sum_{j=0}^{k-1} \frac{\lambda^j}{j!}.$$

We index :

$$f_A(k+1) - f_A(k) \leqslant \frac{\mathbb{P}(Z \geqslant k+1)}{\lambda} + \frac{1}{\lambda} \sum_{j=1}^{k} \frac{\lambda^j}{k(j-1)!}.$$

Finally, we use the fact that $k \geqslant j$ in the sum to have :

$$f_A(k+1) - f_A(k) \leqslant \frac{\mathbb{P}(Z \geqslant k+1)}{\lambda} + \frac{1}{\lambda} \sum_{j=1}^{k} \frac{\lambda^j}{j!} = \frac{\mathbb{P}(Z > 0)}{\lambda}.$$

That is

$$f_A(k+1) - f_A(k) \leqslant \frac{\mathbb{P}(Z \geqslant k+1)}{\lambda} + \frac{1}{\lambda} \sum_{j=1}^{k} \frac{\lambda^j}{j!} = \frac{1 - e^{-\lambda}}{\lambda}.$$

We prove the inequality without the absolute values for every $A \subset \mathbb{N}$. We just need to introduce it to conclude.

• By additivity, we observe that for every $A \subset \mathbb{N}$ (and since $f_{\mathbb{N}} = 0$) that :

$$f_A(k+1) + f_{A^c}(k+1) = f_A(k) + f_{A^c}(k).$$

So,

$$f_A(k+1) - f_A(k) = -\left(f_{A^c}(k+1) - f_{A^c}(k)\right).$$

Suppose that $f_A(k+1) - f_A(k) \leqslant 0$, then :

$$|f_A(k+1) - f_A(k)| = f_{A^c}(k+1) - f_{A^c}(k) \leqslant \frac{1 - e^{-\lambda}}{\lambda}.$$

And if $f_A(k+1) - f_A(k) \geqslant 0$, the previous case concludes to the inequality. We proved that for every $k \in \mathbb{N}$ and $A \subset \mathbb{N}$:

$$|f_A(k+1) - f_A(k)| \leqslant \frac{1 - e^{-\lambda}}{\lambda}.$$

So, by triangular inequality, it follows that $f_A \in \Psi(\lambda)$.

ii. By Stein's equation :

$$\sup_{A \subset \mathbb{N}} |\mathbb{P}(X \in A) - \mathbb{P}(Z \in A)| = \sup_{A \subset \mathbb{N}} |\mathbb{E}\left[\lambda f_A(X+1) - X f_A(X)\right]|.$$

And since $f_A \in \Psi(\lambda)$, we have

$$\sup_{A \subset \mathbb{N}} |\mathbb{P}(X \in A) - \mathbb{P}(Z \in A)| \leqslant \sup_{f \in \Psi(\lambda)} |\mathbb{E}\left[\lambda f(X+1) - X f(X)\right]|.$$

We proved our Stein's bound. □

We can prove by this way our classic theorem about Poisson approximation.

Theorem I.3 : Poisson approximation

Let $(Y_n)_n$ a sequence of random variables, $(p_n)_n$ a sequence of elements of $]0,1[$ such that :

(i) $Y_n \sim \mathcal{B}(n, p_n)$;

(ii) $p_n \xrightarrow[n \to +\infty]{} 0$;

(iii) $np_n \xrightarrow[n \to +\infty]{} \lambda > 0$.

Then $(Y_n)_n$ converges in total variation sense (so in law) to $\mathcal{P}(\lambda)$.

Proof : We consider $X_1^n, \cdots, X_n^n$ independent of law $\mathcal{B}(p_n)$ such that

$$Y_n = \sum_{k=1}^{n} X_k^n.$$

And we consider $V_n = Y_N - X_1^n$. Then for all $f : \mathbb{N} \longrightarrow \mathbb{R}$ bounded :

$$\mathbb{E}[Y_n f(Y_n)] = \sum_{k=0}^{n} k f(k) \binom{n}{k} p_n^k (1-p_n)^{n-k}.$$

The term $k = 0$ is zero. Using the relation

$$k\binom{n}{k} = n\binom{n-1}{k-1},$$

we get by indexing the sum :

$$\mathbb{E}[Y_n f(Y_n)] = np_n \sum_{k=0}^{n-1} f(k+1)\binom{n-1}{k} p_n^k (1-p_n)^{n-1-k}.$$

In terms of $V_n \overset{\text{law}}{=} Y_{n-1}$, we have :

$$\mathbb{E}[Y_n f(Y_n)] = np_n \mathbb{E}[f(V_n+1)].$$

Hence, by Chen-Stein's bounds, we get, if $Z \sim \mathcal{P}(\lambda)$:

$$\begin{aligned} & \sup_{A \subset \mathbb{N}} |\mathbb{P}(Y_n \in A) - \mathbb{P}(Z \in A)| \\ \leqslant & \sup_{f \in \Psi(\lambda)} |\lambda \mathbb{E}[f(Y_n+1)] - np_n \mathbb{E}[f(V_n+1)]|. \end{aligned}$$

By adding and substracting by $\lambda \mathbb{E}[f(V_n+1)]$, we get :

$$\begin{aligned} & \sup_{A \subset \mathbb{N}} |\mathbb{P}(Y_n \in A) - \mathbb{P}(Z \in A)| \\ \leqslant & |\lambda| \sup_{f \in \Psi(\lambda)} |\mathbb{E}\left[f(Y_n+1) - f(V_n+1)\right]| \\ + & |np_n - \lambda| \sup_{f \in \Psi(\lambda)} |\mathbb{E}[f(V_n+1)]| \end{aligned}$$

For the first term, we have for all $f \in \Psi(\lambda)$:

$$|\mathbb{E}\left[f(Y_n+1) - f(V_n+1)\right]| \leqslant \frac{1-e^{-\lambda}}{\lambda} |\mathbb{E}[X_1^n]| = \frac{1-e^{-\lambda}}{\lambda} p_n.$$

Hence the first term goes to zero. For the second, we have for all $f \in \Psi(\lambda)$:

$$|np_n - \lambda| \sup_{f \in \Psi(\lambda)} |\mathbb{E}[f(V_n+1)]| \leqslant \frac{e^\lambda}{\lambda} |np_n - \lambda|,$$

so the second term also goes to zero. So, we proved that

$$\begin{aligned} & \sup_{A \subset \mathbb{N}} |\mathbb{P}(Y_n \in A) - \mathbb{P}(Z \in A)| \\ \leqslant & \left(1 - e^{-\lambda}\right) p_n + \frac{e^\lambda}{\lambda} |np_n - \lambda|. \end{aligned}$$

This proves our theorem. □

II Univariate normal approximations

II.1 General approximations in Malliavin calculus

In this part, we will estimate some distance with respect with the term $\langle \mathrm{D}F, -\mathrm{D}L^{-1}F \rangle_{\mathcal{H}}$, coming from integration by parts formula for Ornstein-Ulhenbeck generator. We will see later why it is relevant to do so. We begin by a lemma giving an expression of $\mathrm{D}L^{-1}F$ in terms of $(P_t)_t$.

Lemma II.1 : An expression of the derivative of the pseudo inverse

Let $F \in \mathbb{D}^{1,2}$ with $\mathbb{E}[F] = 0$. Then $L^{-1}F \in \mathbb{D}^{1,2}$ and

$$-\mathrm{D}L^{-1}F = \int_0^{+\infty} e^{-t} P_t \mathrm{D}F \ \mathrm{d}t.$$

Proof of the lemma : First, let us see that the decomposition in Wiener chaos of $L^{-1}F$ is

$$L^{-1}F = -\sum_{p=1}^{+\infty} \frac{J_p F}{p}.$$

So,

$$\sum_{p=0}^{+\infty} p\mathbb{E}\left[\left|J_p L^{-1} F\right|^2\right] = \sum_{p=1}^{+\infty} \mathbb{E}\left[J_p F^2\right].$$

So,

$$\sum_{p=0}^{+\infty} p\mathbb{E}\left[\left|J_p L^{-1} F\right|^2\right] \leqslant \sum_{p=0}^{+\infty} \mathbb{E}\left[J_p F^2\right] < +\infty.$$

We notice that $L^{-1}F$ is always in $\mathbb{D}^{1,2}$, whenever $F \in \mathbb{D}^{1,2}$ or not. Moreover, we have the following decomposition in Wiener chaos (in $\mathcal{H}$) of $\mathrm{D}L^{-1}F$:

$$\mathrm{D}L^{-1}F = \sum_{p=1}^{+\infty} \frac{1}{p} J_{p-1}(\mathrm{D}F).$$

On the other hand, we have

$$P_t \mathrm{D}F = \sum_{p=0}^{+\infty} e^{-pt} J_p(\mathrm{D}F).$$

So, if we integrate (and switch sum and integral in $L^2(\Omega \to \mathcal{H})$) :

$$\int_0^{+\infty} e^{-t} P_t \mathrm{D}F \ \mathrm{d}t = \sum_{p=0}^{+\infty} \left(\int_0^{+\infty} e^{-(p+1)t} \ \mathrm{d}t\right) J_p(\mathrm{D}F).$$

We compute this integral (it is equal to $\frac{1}{p+1}$) and conclude on the equality we expect. □

Proposition II.1 : Start of the game

i. Let $F \in \mathbb{D}^{1,2}$ such that $\mathbb{E}[F]=0$ and $\mathbb{E}[F^2]=1$, and $f:\mathbb{R}\longrightarrow\mathbb{R}$ of class C^1 and K-Lipschitz. Then

$$|\mathbb{E}\left[f'(F)-Ff(F)\right]| \leqslant K\mathbb{E}\left[\left|1-\left\langle \mathrm{D}F,-DL^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

ii. Moreover, if $F\in\mathbb{D}^{1,4}$ then $\left\langle \mathrm{D}F,-DL^{-1}F\right\rangle_{\mathcal{H}}\in L^2$ and

$$\mathbb{E}\left[\left|1-\left\langle \mathrm{D}F,-DL^{-1}F\right\rangle_{\mathcal{H}}\right|\right] \leqslant \sqrt{\mathrm{Var}\left(\langle \mathrm{D}F,-DL^{-1}F\rangle_{\mathcal{H}}\right)}.$$

Proof : **i.** We note :

$$\text{(A)} \overset{\text{def.}}{=} \left|\mathbb{E}\left[f'(F)-Ff(F)\right]\right|.$$

Since f is C^1 with its derivative being bounded, we can apply the integration by parts formula for the pseudo-inverse to compute $\mathbb{E}[F(F)]$:

$$\text{(A)} = \left|\mathbb{E}\left[f'(F)\left(1-\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right)\right]\right|.$$

Hence by triangular inequality, and since f is K-Lipschitz :

$$\text{(A)} \leqslant K\mathbb{E}\left[\left|1-\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

ii. We suppose that $F\in\mathbb{D}^{1,4}$. Let us prove that $\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\in L^2$. We have $L^{-1}F\in\mathbb{D}^{1,4}$ too, since the series $\sum_{n\geqslant 1} n\mathbb{E}[J_nF^4]$ is convergent. Then, we can compute the derivative of $L^{-1}F$. By Cauchy-Schwarz, we have :

$$\mathbb{E}\left[\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}^2\right] \leqslant \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^2\left\|\mathrm{D}L^{-1}F\right\|_{\mathcal{H}}^2\right].$$

By Cauchy Schwarz again, we have :

$$\mathbb{E}\left[\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}^2\right] \leqslant \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^4\right]^{\frac{1}{2}}\mathbb{E}\left[\left\|\mathrm{D}L^{-1}F\right\|_{\mathcal{H}}^4\right]^{\frac{1}{2}}.$$

We can estimate the norm of $\mathrm{D}L^{-1}F$, by the lemma :

$$\mathbb{E}\left[\left\|\mathrm{D}L^{-1}F\right\|_{\mathcal{H}}^4\right] = \mathbb{E}\left[\left\|\int_0^{+\infty} e^{-t}P_t\mathrm{D}F\,\mathrm{d}t\right\|_{\mathcal{H}}^4\right].$$

By Jensen inequality on the measure $e^{-t}\mathbf{1}_{\mathbb{R}_+}\,\mathrm{d}t$:

$$\mathbb{E}\left[\left\|\mathrm{D}L^{-1}F\right\|_{\mathcal{H}}^4\right] \leqslant \mathbb{E}\left[\int_0^{+\infty} e^{-t}\|P_t\mathrm{D}F\|_{\mathcal{H}}^4\mathrm{d}t\right].$$

We switch expectation and integral and use the contraction property of (P_t) :

$$\mathbb{E}\left[\left\|\mathrm{D}L^{-1}F\right\|_{\mathcal{H}}^4\right] \leqslant \int_0^{+\infty} e^{-t}\mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^4\right]\mathrm{d}t.$$

Since $\int_{\mathbb{R}_+} e^{-t}\,\mathrm{d}t=1$, we finally have :

$$\mathbb{E}\left[\left\|\mathrm{D}L^{-1}F\right\|_{\mathcal{H}}^4\right] \leqslant \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^4\right].$$

And so :

$$\mathbb{E}\left[\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}^2\right] \leqslant \mathbb{E}\left[\|\mathrm{D}F\|_{\mathcal{H}}^4\right] < +\infty.$$

We proved that $\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\in L^2$. But, by integration by parts formula, we have this time :

$$\mathbb{E}\left[\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right] = \mathbb{E}[F^2]=1.$$

So, by Cauchy-Schwarz inequality, we have :

$$\mathbb{E}\left[\left|1-\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right] \leqslant \mathrm{Var}\left(\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right)^{\frac{1}{2}}.$$

This concludes the proof. □

It is consequently possible to use our Stein's bounds to have the following proposition.

Proposition II.2 : Estimations for the different distances in terms of Malliavin calculus

Let $F\in\mathbb{D}^{1,2}$ such that $\mathbb{E}[F]=0$ and $\mathbb{E}[F^2]=\sigma^2>0$. Let $N\sim\mathcal{N}(0,\sigma^2)$. Then, we have :

(i) For Wasserstein distance :

$$d_{\mathrm{W}}(F,N) \leqslant \sqrt{\frac{2}{\sigma^2\pi}}\mathbb{E}\left[\left|\sigma^2-\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

(ii) For total variation distance :

$$d_{\mathrm{TV}}(F,N) \leqslant \frac{2}{\sigma^2}\mathbb{E}\left[\left|\sigma^2-\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

(iii) For Kolmogorov distance :

$$d_{\mathrm{Kol}}(F,N) \leqslant \frac{1}{\sigma^2}\mathbb{E}\left[\left|\sigma^2-\left\langle \mathrm{D}F,-\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

Proof : The computations are likely the same as those for estimating the distance between centred Gaussian random variables.

A. First, let us do it for $\sigma = 1$. This is just application of what we just saw. Indeed, by Stein's bound, we have :

$$d_{\mathrm{W}}(F,N) \leqslant \sup_{f\in\mathcal{F}_{\mathrm{W}}} \left|\mathbb{E}\left[f'(F) - Ff(F)\right]\right|,$$

where $\mathcal{F}_{\mathrm{W}}$ is the set of C^1 functions being $\sqrt{\frac{2}{\pi}}$-Lipschitz. Hence, by the previous proposition :

$$d_{\mathrm{W}}(F,N) \leqslant \sqrt{\frac{2}{\pi}}\mathbb{E}\left[\left|1 - \left\langle \mathrm{D}F, -\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

Same argument for d_{TV} (where f is 2-Lipschitz) and for d_{Kol} (where f is 1-Lipschitz).

B. We can prove one by one our inequalities for general $\sigma > 0$.

i. By definition :

$$d_{\mathrm{W}}(F,N) = \sup_{h\in\mathrm{Lip}(1)} |\mathbb{E}[h(F)] - \mathbb{E}[h(N)]|.$$

We would use the Stein's equation and the previous proposition, but we need to have random variables with variance equal to 1. To do this, we just need to use the bijection $h \longmapsto h(\sigma\cdot)$ between $\mathrm{Lip}(1)$ and $\mathrm{Lip}(\sigma)$:

$$d_{\mathrm{W}}(F,N) = \sup_{h\in\mathrm{Lip}(\sigma)} \left|\mathbb{E}\left[h\left(\frac{F}{\sigma}\right)\right] - \mathbb{E}\left[h\left(\frac{N}{\sigma}\right)\right]\right|.$$

By Stein's equation, and by seeing that $f_h \in C^1 \cap \mathrm{Lip}\left(\sigma\sqrt{\frac{2}{\pi}}\right)$ if $h \in \mathrm{Lip}(\sigma)$ (we proved it for $\sigma = 1$, we conclude by scaling), we finally have :

$$d_{\mathrm{W}}(F,N) \leqslant \sup_{f\in\mathcal{C}^1\cap\mathrm{Lip}\left(\sigma\sqrt{\frac{2}{\pi}}\right)} \left|\mathbb{E}\left[f'\left(\frac{F}{\sigma}\right) - \frac{F}{\sigma}f\left(\frac{F}{\sigma}\right)\right]\right|.$$

We are exactly in the case **A.**, so :

$$d_{\mathrm{W}}(F,N) \leqslant \sigma\sqrt{\frac{2}{\pi}}\mathbb{E}\left[\left|1 - \frac{1}{\sigma^2}\left\langle \mathrm{D}F, -\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}}\right|\right].$$

ii. The computations are the same, and even are easier since $B \longmapsto \sigma B$ is a bijection between $\mathcal{B}(\mathbb{R})$ and itself. So, we have directly

$$d_{\mathrm{TV}}(F,N) = d_{\mathrm{TV}}\left(\frac{F}{\sigma}, \frac{N}{\sigma}\right),$$

and we conclude by using **A.**. Same for $\mathrm{d}_{\mathrm{Kol}}$. □

II.2 A first estimation on Wiener chaos

Before starting, remind the Prohorov theorem, stated as a lemma here.

Lemma II.2 : Prohorov theorem

Let $(\mathbb{P}_n)_n$ a sequence of probabilities on $\mathbb{R}^d$. We suppose that $(\mathbb{P}_n)_n$ is *uniformly tight* :

$$\forall\varepsilon > 0, \exists K \subset \mathbb{R}^d \text{ compact}, \forall n \in \mathbb{N}, \mathbb{P}_n(K) \geqslant 1 - \varepsilon.$$

Then $(\mathbb{P}_n)_n$ admits a subsequence that converges tightly.

In other words, if a sequence of random variables is tight, it converges in law, up to a subsequence. This part is dedicated to precise the estimations made in general in previous subsection, in the case where F is in a Wiener chaos. Let us begin by a theorem known as *method of moments*.

Theorem II.1 : Method of moments

Let $q \geqslant 2$, $\sigma > 0$ and $(F_n)_n \in (\mathfrak{H}_p)^{\mathbb{N}}$. We suppose that

$$\mathbb{E}\left[F_n^2\right] \xrightarrow[n\to+\infty]{} \sigma^2.$$

Then, the following assertions are equivalent :

(i) The sequence $(F_n)_n$ converges in law to $\mathcal{N}(0,\sigma^2)$;

(ii) If $N \sim \mathcal{N}(0,\sigma^2)$, then

$$\forall j \geqslant 3, \mathbb{E}[F_n^j] \xrightarrow[n\to+\infty]{} \mathbb{E}[N^j].$$

We have here a quite powerful way to determine if a sequence converges in law to a Gaussian random variable.

Proof : The hypothesis give a lot of information before starting proving the equivalence. First, since $(\mathbb{E}[F_n^2])_n$ is bounded, we have by hypercontractivity :

$$\forall\eta \geqslant 1, \sup_{n\in\mathbb{N}} \mathbb{E}\left[|F_n|^\eta\right] < +\infty.$$

Second, by Markov inequality, $(F_n)_n$ is tight. Indeed, let $K > 0$. Then :

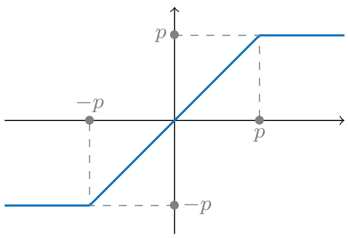
$$\mathbb{P}(|F_n| > K) \leqslant \frac{\sup_{n\in\mathbb{N}}\mathbb{E}[F_n^2]}{K^2}.$$

By Prohorov theorem, $(F_n)_n$ admits a subsequence that converges in law.

[$\Longrightarrow$] We suppose that $(F_n)_n$ converges in law to $N \sim \mathcal{N}(0, \sigma^2)$. By the continuous mapping theorem, this is the case of every F_n^j. So, for all $h : \mathbb{R} \longrightarrow \mathbb{R}$ continuous and bounded, we have

$$\mathbb{E}\left[h(F_n^j)\right] \xrightarrow[n \to +\infty]{} \mathbb{E}\left[h(N^j)\right].$$

Let $(h_p)_p$ a sequence of continuous bounded functions on $\mathbb{R}$ defined by :



Then $(h_p)_p$ converges pointwise to the identity map. We have the following domination :

$$\forall p \in \mathbb{N}, |h_p(F_n^j)| \leqslant |F_n|^j.$$

And we know by hypercontractivity that $F_n^j \in L^1$, so we have by dominated convergence :

$$\mathbb{E}\left[h_p(F_n^j)\right] \xrightarrow[p \to +\infty]{} \mathbb{E}[F_n^j].$$

By the same argument, we have

$$\mathbb{E}\left[h_p(N^j)\right] \xrightarrow[p \to +\infty]{} \mathbb{E}[N^j].$$

Hence, we can conclude about the expected convergence. We just need to cut in three :

$$\begin{aligned} & \left|\mathbb{E}[N^j] - \mathbb{E}[F_n^j]\right| \\ \leqslant \quad & \left|\mathbb{E}\left[N^j\right] - \mathbb{E}[h_p(N^j)]\right| + \left|\mathbb{E}\left[h_p(N^j)\right] - \mathbb{E}[h_p(F_n^j)]\right| \\ + \quad & \left|\mathbb{E}\left[h_p(F_n^j)\right] - \mathbb{E}[F_n^j]\right|. \end{aligned}$$

For p big enough, the previous dominated convergences shows that the first and last terms are as small as we want. The middle term is also small as we want for n big enough by the convergence in law. We conclude in our first implication.

[$\Longleftarrow$] By our remark in the beginning of the proof, there exists a subsequence ϕ and a random variable Y such that

$$F_{\phi(n)} \xrightarrow[n \to +\infty]{\text{law}} Y.$$

Then, Y admits moment at every order. Indeed, if we consider the same approximation $(h_p)_p$ that previously, we have by monotone convergence ($|h_p| \leqslant |h_{p+1}|$) that

$$\mathbb{E}\left[|h_p(Y^j)|\right] \xrightarrow[p \to +\infty]{} \mathbb{E}\left[|Y|^j\right].$$

But, we have $|h_p(x)| \leqslant |x|$ and using the convergence in law of $F_{\phi(n)}$ to Y, we have :

$$\forall p \in \mathbb{N}, \mathbb{E}\left[|h_p(Y^j)|\right] \leqslant \sup_{n \in \mathbb{N}} \mathbb{E}\left[|F_n|^j\right].$$

So, by letting $[p \to +\infty]$, we have :

$$\mathbb{E}\left[|Y|^j\right] \leqslant \sup_{n \in \mathbb{N}} \mathbb{E}\left[|F_n|^j\right] < +\infty.$$

So Y admits moments at every order. By dominated convergence (same argument as previously), for all $j \geqslant 2$:

$$\mathbb{E}\left[F_{\phi(n)}^j\right] \xrightarrow[n \to +\infty]{} \mathbb{E}[Y^j].$$

By unicity of the limit, we have :

$$\forall j \in \mathbb{N}, \mathbb{E}\left[Y^j\right] = \mathbb{E}\left[N^j\right].$$

Like we saw in Stein's lemma, it means that $Y \sim \mathcal{N}(0, \sigma^2)$. □

II.3 Fourth moment theorem

We can do better that this theorem, by using Stein's method. We will prove that the condition **(ii)** is *equivalent* to the same condition but for only $j = 4$. That is the fourth moment theorem.

Lemma II.3 : Product formula

Let $\mathcal{H} = L^2(T, \mathcal{B}, \mu)$, with μ an atomless measure on the σ-algebra $\mathcal{B}$. Let $p, j \in \mathbb{N}$, $f \in L^2_{\mathsf{S}}(T^p)$, $g \in L^2_{\mathsf{S}}(T^{p+j})$. For $r \leqslant p$, we name the *r-th contraction of f and g* the map $f \otimes_r g \in L^2(T^{2(p-r)+j})$ defined by :

$$\begin{aligned} f \otimes_r g(t_1, \cdots, t_{p-r}, s_1, \cdots, s_{p+j-r}) \overset{\text{def.}}{=} \quad & \int_{T^r} f(t_1, \cdots, t_{p-r}, u_1, \cdots, u_r) \\ \cdot \quad & g(s_1, \cdots, s_{p+j-r}, u_1, \cdots, u_r) \\ & \mathrm{d}\mu(u_1) \cdots \mathrm{d}\mu(u_r). \end{aligned}$$

Then, we can decompose in Wiener chaos the random variable $I_p(f) I_{p+j}(g)$:

$$I_p(f) I_{p+j}(g) = \sum_{r=0}^{p} r! \binom{p}{r} \binom{q}{r} I_{2(p-r)+j}\left(f \otimes_r g\right).$$

Note : If $f, g \in L^2_{\mathsf{S}}$, then $f \otimes_r g$ is not symmetric in general. We denote by $f \tilde{\otimes}_r g$ the symmetrization of $f \otimes_r g$.

Proof of the lemma : The proof is the same the one for Hermite polynomials in the one-dimensional case. By hypercontractivity, we know that $I_p(f) \in \cap_{p \geqslant 1} L^p(\mathbb{P})$. Moreover, by the characterization of the domain $\mathbb{D}^{p,2}$ with the series representation, we have in fact that $I_p(f) \in \mathbb{D}^{\infty,2}$, for every $p \in \mathbb{N}$ and $f \in L^2(T^p)$. We have even better : we have in fact $I_p(f) \in \mathbb{D}^{\infty,q}$, for every $q \in \mathbb{N}^*$. This is a consequence of Meyer's inequality, stating that the iterated divergence is continuous :

$$\delta^p : \mathbb{D}^{k,q}(L^2(T^p)) \longrightarrow \mathbb{D}^{k-p,q}.$$

We admit it here. Hence, it follows that the product $I_p(f)I_{p+j}(g)$ belongs to $\mathbb{D}^{\infty,2}$. Hence, we can apply the Strook formula :

$$I_p(f)I_{p+j}(g) = \sum_{s=0}^{+\infty} \frac{I_s\left(\mathbb{E}\left[\mathrm{D}^s_\bullet\left\{I_p(f)I_{p+j}(g)\right\}\right]\right)}{s!}.$$

We apply the Leibniz rule :

$$I_p(f)I_{p+j}(g) = \sum_{s=0}^{+\infty} \frac{I_s\left(\mathbb{E}\left[\sum_{k=0}^{s}\binom{s}{k}\mathrm{D}^k I_p(f)\otimes \mathrm{D}^{s-k}I_{p+j}(g)\right]\right)}{s!}.$$

The computation of the expectation is really the same as the case of Hermite polynomials, in section I. We refer to this to conclude the proof of this lemma. □

Remember that in the Wiener case, for $F = I_q(f)$, we have :

$$\left\langle \mathrm{D}F, -\mathrm{D}L^{-1}F\right\rangle_{\mathcal{H}} = \frac{\|\mathrm{D}F\|^2_{\mathcal{H}}}{q}.$$

Here is why we want to show the following inequality.

Lemma II.4 : An estimation for the variance of the derivative of an element in a chaos

Let $q \geqslant 2$ and $F \in \mathfrak{H}_q$. Then

$$\mathrm{Var}\left(\frac{\|\mathrm{D}F\|^2_{\mathcal{H}}}{q}\right) \leqslant \frac{q-1}{3q}\left(\mathbb{E}\left[F^4\right] - 3\mathbb{E}\left[F^2\right]^2\right) \leqslant (q-1)\mathrm{Var}\left(\frac{\|\mathrm{D}F\|^2_{\mathcal{H}}}{q}\right).$$

Proof of the lemma : The idea is to expend every term in Wiener chaos and to compare those expansions.

- We have

$$\|\mathrm{D}F\|^2_{\mathcal{H}} = \int_T \mathrm{D}_t F^2 \,\mathrm{d}\mu(t).$$

Since $F = I_q(f)$, we have $\mathrm{D}_t F = qI_{q-1}(f(t,\cdot))$. Hence,

$$\|\mathrm{D}F\|^2_{\mathcal{H}} = q^2\int_T I_{q-1}(f(t,\cdot))^2 \,\mathrm{d}\mu(t).$$

By product formula :

$$\|\mathrm{D}F\|^2_{\mathcal{H}} = q^2\int_T \sum_{r=0}^{q-1} r!\binom{q-1}{r}^2 I_{2(q-1-r)}(f(t,\cdot)\otimes_r f(t,\cdot))\,\mathrm{d}\mu(t).$$

By linearity and continuity of I_{q-1} (it is an isometry), we have

$$\|\mathrm{D}F\|^2_{\mathcal{H}} = q^2\sum_{r=0}^{q-1} r!\binom{q-1}{r}^2 I_{2(q-1-r)}(f\otimes_{r+1} f).$$

By indexing the sum again :

$$\|\mathrm{D}F\|^2_{\mathcal{H}} = q^2\sum_{r=1}^{q} (r-1)!\binom{q-1}{r-1}^2 I_{2(q-r)}(f\otimes_{r} f).$$

We use the relation

$$q(r-1)!\binom{q-1}{r-1}^2 = \frac{r}{q}r!\binom{q}{r}^2,$$

to conclude that, we have our first expansion :

$$\|\mathrm{D}F\|^2_{\mathcal{H}} = \sum_{r=1}^{q} r\cdot r!\binom{q}{r}^2 I_{2(q-r)}(f\otimes_{r} f).$$

- We can compute its variance. Indeed, we have with the term $r = q$ that

$$\mathbb{E}\left[\|\mathrm{D}F\|^2_{\mathcal{H}}\right] = q^2(q!)^2\|f\|^4_{L^2(T^q)} = q^2\mathbb{E}[F^2]^2.$$

Then, by independence of the I_q :

$$\mathrm{Var}\left(\|\mathrm{D}F\|^2_{\mathcal{H}}\right) = \sum_{r=1}^{q-1} r^2(r!)^2\binom{q}{r}^4(2(q-r))!\left\|f\tilde{\otimes}_r f\right\|^2_{L^2(T^{2(q-r)})},$$

where $\tilde{\otimes}_r$ means the symmetrization of the tensor product. We keep this expression aside, and we will use it later for the comparison.

- Let us try to do the same kind of computations for F^4. Since $F\in\mathfrak{H}_q$, $LF = -qF$, so

$$\mathbb{E}\left[F^4\right] = -\frac{1}{q}\mathbb{E}\left[LF\cdot F^3\right].$$

By using $L = -\delta\mathrm{D}$, we have

$$\mathbb{E}\left[F^4\right] = \frac{1}{q}\mathbb{E}\left[\delta\mathrm{D}F\cdot F^3\right].$$

By duality :

$$\mathbb{E}\left[F^4\right] = \frac{1}{q}\mathbb{E}\left[\left\langle \mathrm{D}F, \mathrm{D}F^3\right\rangle_{\mathcal{H}}\right].$$

By chain rule, $\mathrm{D}F^3 = 3F^2\mathrm{D}F$ so :

$$\mathbb{E}\left[F^4\right] = \frac{3}{q}\mathbb{E}\left[F^2\|\mathrm{D}F\|^2_{\mathcal{H}}\right].$$

So, to conclude, we will use the expansion in Wiener chaos of each factor.

- We have, by product formula again :

$$F^2 = \sum_{r=0}^{q} r!\binom{q}{r}^2 I_{2(q-r)}(f\otimes_r f).$$

So,

$$\mathbb{E}\left[F^2\|\mathrm{D}F\|^2_{\mathcal{H}}\right] = \sum_{r=1}^{q} r(r!)^2\binom{q}{r}^2(2(q-r))!\left\|f\tilde{\otimes}_r f\right\|_{L^2(T^{2(q-r)})}.$$

And so

$$\mathbb{E}[F^4] = \frac{3}{q}\sum_{r=1}^{q} r(r!)^2\binom{q}{r}^2(2(q-r))!\left\|f\tilde{\otimes}_r f\right\|_{L^2(T^{2(q-r)})}.$$

By isolating the term $r = q$, we find $3\mathbb{E}[F^2]^2$. Hence, we have our second expression :

$$\begin{aligned}&\mathbb{E}[F^4]-3\mathbb{E}[F^2]^2\\ =\ &\frac{3}{q}\sum_{r=1}^{q-1} r(r!)^2\binom{q}{r}^2(2(q-r))!\left\|f\tilde{\otimes}_r f\right\|_{L^2(T^{2(q-r)})}.\end{aligned}$$

• Now that we have the expansion of every terms present in our expected inequality, we simply use :

◁ the fact that $r \leqslant q-1$ in the expression of $\mathrm{Var}\left(\|\mathrm{D}F\|^2_{\mathcal{H}}\right)$ to have

$$\mathrm{Var}\left(\|\mathrm{D}F\|^2_{\mathcal{H}}\right) \leqslant \frac{q(q-1)}{3}\left(\mathbb{E}[F^4]-3\mathbb{E}[F^2]^2\right).$$

◁ the fact that $r \leqslant r^2$ in the expansion of $\mathbb{E}[F^4]-3\mathbb{E}[F^2]^2$ to have :

$$\mathbb{E}[F^4]-3\mathbb{E}[F^2]^2 \leqslant \frac{3}{q}\mathrm{Var}\left(\|\mathrm{D}F\|^2_{\mathcal{H}}\right).$$

This concludes the proof of this inequality. □

Theorem II.2 : Fourth moment theorem

Let $q \geqslant 2$ and $(F_n)_n$ a sequence of elements of $\mathfrak{H}_q$. We suppose that

$$\mathbb{E}\left[F_n^2\right] \xrightarrow[n\to+\infty]{} \sigma^2 > 0.$$

Then the following assertions are equivalent.

(i) $(F_n)_n$ converges in law to $N \sim \mathcal{N}(0,\sigma^2)$.

(ii) The sequence $(\mathbb{E}[F_n^4])_n$ converges to $\mathbb{E}[N^4]$:

$$\mathbb{E}\left[F_n^4\right] \xrightarrow[n\to+\infty]{} 3\sigma^4 = \mathbb{E}[N^4].$$

(iii) The variance of the norm of the Malliavin derivative goes to zero :

$$\mathrm{Var}\left(\|\mathrm{D}F_n\|^2_{\mathcal{H}}\right) \xrightarrow[n\to+\infty]{} 0.$$

Proof : **[(i) $\Longrightarrow$ (ii)]** This implication is true by the moments method theorem, applied for $j=4$.

[(ii) $\Longleftrightarrow$ (iii)] This equivalence is true by the previous lemma.

[(iii) $\Longrightarrow$ (i)] Here's the interesting part. Remember that we proved that

$$d_{\mathrm{Kol}}(F_n,N) \leqslant \frac{1}{\mathbb{E}[F_n^2]}\mathbb{E}\left[\left|\mathbb{E}\left[F_n^2\right]-\left\langle \mathrm{D}F_n, -\mathrm{D}L^{-1}F_n\right\rangle_{\mathcal{H}}\right|\right].$$

Since $F_n \in \mathfrak{H}_q$, we have by integration by parts formula for the pseudo inverse that

$$\mathbb{E}[F_n^2] = \mathbb{E}\left[\left\langle \mathrm{D}F_n, -\mathrm{D}L^{-1}F_n\right\rangle_{\mathcal{H}}\right].$$

So, we get by Cauchy-Schwarz inequality :

$$d_{\mathrm{Kol}}(F_n,N) \leqslant \frac{1}{\mathbb{E}[F_n^2]^2}\sqrt{\mathrm{Var}\left(\langle \mathrm{D}F_n, -\mathrm{D}L^{-1}F_n\rangle_{\mathcal{H}}\right)}.$$

But, we have in fact :

$$\left\langle \mathrm{D}F_n, -\mathrm{D}L^{-1}F_n\right\rangle_{\mathcal{H}} = \frac{1}{q}\|\mathrm{D}F_n\|^2_{\mathcal{H}}.$$

So, we have :

$$d_{\mathrm{Kol}}(F_n,N) \leqslant \frac{1}{\mathbb{E}[F_n^2]^2}\sqrt{\mathrm{Var}\left(\frac{1}{q}\|\mathrm{D}F_n\|^2_{\mathcal{H}}\right)}.$$

So the Kolmogorov distance between $(F_n)_n$ and N goes to zero, and the density function of N is continuous. We conclude that $(F_n)_n$ converges in law to N. □

We can prove the implication **(iii)** $\Longrightarrow$ **(i)** of the theorem by using the Prohorov theorem. We already used it for the method of moments.

Proof : We suppose that

$$\mathrm{Var}\left(\|\mathrm{D}F_n\|^2_{\mathcal{H}}\right) \xrightarrow[n\to+\infty]{} 0.$$

Let us show that $(F_n)_n$ converges in law to $\mathcal{N}(0,\sigma^2)$.

• First, let us see that if φ_n is the characteristic function of F_n, then φ_n is differentiable with

$$\varphi_n'(t) = \mathrm{i}\mathbb{E}\left[F_n e^{\mathrm{i}tF_n}\right].$$

By integration by parts formula for the pseudo inverse, we get :

$$\varphi_n'(t) = t\mathbb{E}\left[e^{\mathrm{i}tF_n}\left\langle \mathrm{D}F_n, -\mathrm{D}L^{-1}\right\rangle\right].$$

Since $F_n \in \mathfrak{H}_q$, we get :

$$\varphi_n'(t) = \frac{-t}{q}\mathbb{E}\left[e^{\mathrm{i}tF_n}\|\mathrm{D}F_n\|^2_{\mathcal{H}}\right].$$

• Since $\mathbb{E}[F_n^2] \xrightarrow[n\to+\infty]{} \sigma^2 > 0$, the sequence $(F_n)_n$ is tight (we just need to apply Markov's inequality). Hence, there exists a subsequence of $(F_n)_n$ that converges in law. To conclude, we just need to show that every subsequence of $(F_n)_n$ that converges in law admits $\mathcal{N}(0,\sigma^2)$ as limit (hence $\{F_n, n\in\mathbb{N}\}$ admits one adherence value and is relatively compact, so $(F_n)_n$ converges). We set a random variable Z and an increasing map $\psi : \mathbb{N} \longrightarrow \mathbb{N}$ such that

$$F_{\psi(n)} \xrightarrow[n\to+\infty]{\text{law}} Z.$$

Then, by using the same type of arguments than the one for the moment theorem, we have $\mathbb{E}[Z^2] < +\infty$, since $\mathbb{E}[F_n^2]$ goes to σ^2. Hence, φ_Z is C^1, with $\varphi'_{\psi(n)}$ converging pointwise to φ'_Z, since by continuous mapping theorem

$$F_{\psi(n)}e^{\mathrm{i}tF_{\psi}(n)} \xrightarrow[n\to+\infty]{\text{law}} Ze^{\mathrm{i}tZ}.$$

and $\mathbb{E}[Z^2] < +\infty$ (like in the moments method theorem, we approach the identity by its truncation to prove it). Moreover,

$$\varphi_n'(t) + \sigma^2 t\varphi_n(t) = t\mathbb{E}\left[e^{\mathrm{i}tF_n}\left(\frac{-\|\mathrm{D}F_n\|_{\mathcal{H}}^2}{q} + \sigma^2\right)\right].$$

So, since $\mathbb{E}\left[\|\mathrm{D}F_n\|_{\mathcal{H}}^2\right] = q\mathbb{E}[F_n^2]$, we have :

$$\left|\varphi_n'(t) + t\varphi_n(t)\right| \leqslant t\sqrt{\mathrm{Var}\left(\frac{\|\mathrm{D}F_n\|_{\mathcal{H}}}{q}\right)} + t\sqrt{\mathbb{E}[F_n^2] - \sigma^2}.$$

Since the variance goes to zero, we conclude that

$$\varphi_Z'(t) + \sigma^2 t\varphi_Z(t) = 0.$$

Since $\varphi_Z(0) = 1$, we conclude that $Z \sim \mathcal{N}(0, \sigma^2)$. □

Note : By the computations made on the lemma, the convergence **(i)** is also equivalent to :

$$\|f_n \otimes_r f_n\|_{L^2(T^{2(q-r)}} \xrightarrow[n\to+\infty]{} 0,$$

where $F_n = I_q(f_n)$.

We can express the same kind of theorem for a sequence in different Wiener chaos.

Corollary II.1 : Fourth moment theorem for various chaos

Let $(F_n)_n \in L^2(\mathbb{P})^{\mathbb{N}}$ a sequence of random variables such that $F_n \in \mathfrak{H}_{q(n)}$. We suppose that

$$\mathbb{E}\left[F_n^2\right] \xrightarrow[n\to+\infty]{} \sigma^2 > 0.$$

And we suppose that

$$\mathbb{E}\left[F_n^4\right] \xrightarrow[n\to+\infty]{} 3\sigma^4.$$

Then, $(F_n)_n$ converges in the Kolmogorov sens to $\mathcal{N}(0, \sigma^2)$, so converges in law.

Proof : Once again, we simply use the lemma. Indeed, we still have

$$d_{\mathrm{Kol}}(F_n, N) \leqslant \frac{2}{\mathbb{E}[F_n^2]}\sqrt{\mathrm{Var}\left(\frac{\|\mathrm{D}F_n\|_{\mathcal{H}}^2}{q(n)}\right)}.$$

By the lemma :

$$d_{\mathrm{Kol}}(F_n, N) \leqslant \frac{2}{\mathbb{E}[F_n^2]}\sqrt{\frac{q(n)-1}{3q(n)}\left(\mathbb{E}\left[F_n^4\right] - 3\mathbb{E}\left[F_n^2\right]^2\right)}.$$

Since the sequence $\left(\frac{n-1}{3n}\right)_n$ is bounded by $\frac{1}{3}$, we get :

$$d_{\mathrm{Kol}}(F_n, N) \leqslant \frac{2}{\sqrt{3}\mathbb{E}[F_n^2]}\sqrt{\left(\mathbb{E}\left[F_n^4\right] - 3\mathbb{E}\left[F_n^2\right]^2\right)},$$

which goes to zero when $[n \to +\infty]$. □

An application of this is the fact that we can have a second overview of the law of multiple integrals (first one is : their laws are absolutely continuous with respect to the Lebesgue measure). They are not Gaussian for chaos with order greater than 2.

Corollary II.2 : Non-Gaussian multiple integrals

Let $q \geqslant 2$ and $F \in \mathfrak{H}_q$. We suppose that $\mathbb{E}[F^2] = \sigma^2 > 0$. Then $\mathbb{E}[F^4] > 3\sigma^4$. Hence, F is not a Gaussian random variable.

Proof : By the lemma :

$$\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 \geqslant \frac{q(q-1)}{3}\mathrm{Var}\left(\|\mathrm{D}F\|_{\mathcal{H}}^2\right) \geqslant 0.$$

Suppose that we have equality. Then $\mathrm{D}F = 0$ almost everywhere and almost surely so that $F = \mathbb{E}[F] = 0$ almost surely. This contradicts the fact that $\mathbb{E}[F^2] > 0$, so we don't have equality, and we have $\mathbb{E}[F^4] > 3\mathbb{E}[F^2]^2$. Since every Gaussian random variable satisfies the equality, F cannot be a Gaussian random variable. □

II.4 Estimation on general case : second order Poincaré inequality

We want to use the fourth moment theorem to find new conditions to assure a convergence in law to $\mathcal{N}(0, 1)$. We begin by a consequence of our inequalities of distance in terms of Malliavin calculus.

Proposition II.3 : A first condition for converging in law to a Gaussian law

Let $(F_n)_n \subset \mathbb{D}^{1,2}$ such that $\mathbb{E}[F_n] = 0$. We suppose that

$$\mathbb{E}\left[F_n^2\right] \xrightarrow[n\to+\infty]{} \sigma^2 > 0.$$

Then, if

$$\mathbb{E}\left[\left|\sigma^2 - \left\langle \mathrm{D}F_n, -\mathrm{D}L^{-1}F_n\right\rangle_{\mathcal{H}}\right|\right] \xrightarrow[n\to+\infty]{} 0,$$

then $(F_n)_n$ converges in law to $\mathcal{N}(0,\sigma^2)$.

The following result is the key for the condition we will derive later. It generalizes the one we saw in the one-dimensional case. Remember that we chose $\mathcal{H} = L^2(T,\mathcal{B},\mu)$, with μ a atomless measure.

Theorem II.3 : Second order Poincaré inequality

Let $F \in \mathbb{D}^{2,4}$ such that $\mathbb{E}[F] = 0$ and $\mathbb{E}[F^2] = \sigma^2 > 0$. Then,

$$d_{\mathrm{TV}}(F,\mathcal{N}(0,\sigma^2)) \leqslant \frac{3}{\sigma^2}\mathbb{E}\left[\left\|\mathrm{D}^2F\right\|_{\mathsf{op}}^4\right]^{\frac{1}{4}} \mathbb{E}\left[\left\|\mathrm{D}^2F\right\|_{L^2(T^2)}^4\right]^{\frac{1}{4}},$$

where

$$\left\|\mathrm{D}^2F\right\|_{\mathsf{op}} \overset{\text{def.}}{=} \sup_{\|f\|_{\mathcal{H}}=1}\left[\int_T\left|\int_T f(s)\mathrm{D}^2_{s,t}F \,\mathrm{d}\mu(s)\right|^2 \,\mathrm{d}\mu(t)\right]^{\frac{1}{2}}$$

is the operator norm of $f \longmapsto \int_T f(s)\mathrm{D}_{\bullet,s}F\,\mathrm{d}\mu(s) \in L^2(T)$. We can moreover have an estimation of it :

$$\mathbb{E}\left[\left\|\mathrm{D}^2F\right\|_{\mathsf{op}}^4\right] \leqslant \mathbb{E}\left[\left\|\mathrm{D}^2F \otimes_1 \mathrm{D}^2F\right\|_{L^2(T^2)}^2\right].$$

Recall that if $f,g \in L^2(T^2)$:

$$f \otimes_1 g(x,y) \overset{\text{def.}}{=} \int_T f(x,t)g(y,t)\,\mathrm{d}\mu(t).$$

With it, here comes the criterion we look for.

Proposition II.4 : Convergence in law with Malliavin calculus

Let $(F_n)_n \subset \mathbb{D}^{2,4}$ such that $\mathbb{E}[F_n] = 0$. We suppose that :

(i) $\mathbb{E}[F_n^2] \xrightarrow[n\to+\infty]{} \sigma^2 > 0$;

(ii) $\sup_{n\geqslant 1}\mathbb{E}\left[\|\mathrm{D}F_n\|_{L^2(T)}^4\right] < +\infty$;

(iii) $\mathbb{E}\left[\left\|\mathrm{D}^2F_n\right\|_{\mathsf{op}}^4\right] \xrightarrow[n\to+\infty]{} 0.$

Then $(F_n)_n$ converges in law (and in total variation distance) to $\mathcal{N}(0,\sigma^2)$.

Note : To prove **(iii)**, it is enough to prove that

$$\mathbb{E}\left[\left\|\mathrm{D}^2F_n \otimes_1 \mathrm{D}^2F_n\right\|_{L^2(T^2)}^2\right] \xrightarrow[n\to+\infty]{} 0.$$

Proof : Let us note $\sigma_n^2 = \mathbb{E}[F_n^2]$. Then :

$$d_{\mathrm{TV}}(F_n,\mathcal{N}(0,\sigma^2)) \leqslant \begin{array}{ll} & d_{\mathrm{TV}}(F_n,\mathcal{N}(0,\sigma_n^2)) \\ + & d_{\mathrm{TV}}(\mathcal{N}(0,\sigma_n^2),\mathcal{N}(0,\sigma^2)). \end{array}$$

By Poincaré's inequality, and by the estimation of the total variation for Gaussian random variables, we have

$$d_{\mathrm{TV}}(F_n,\mathcal{N}(0,\sigma^2)) \leqslant \begin{array}{ll} & \frac{3}{\sigma_n^2}\mathbb{E}\left[\left\|\mathrm{D}^2F_n\right\|_{\mathsf{op}}^4\right]^{\frac{1}{4}}\mathbb{E}\left[\left\|\mathrm{D}^2F_n\right\|_{L^2(T^2)}^4\right]^{\frac{1}{4}} \\ + & \frac{2|\sigma^2-\sigma_n^2|}{\max\{\sigma_n^2,\sigma^2\}}. \end{array}$$

Hence, the total variation goes to zero, so as the Kolmogorov distance. Since the law $\mathcal{N}(0,\sigma^2)$ has a continuous density, we conclude that $(F_n)_n$ converges in law to $\mathcal{N}(0,\sigma^2)$. □

III Multivariate Stein's lemma and multivariate estimations

Here, we will generalize all the concepts we saw in previous section in multidimensional case. We will be quick on the proofs, the ideas are the same, it is simply the popping up of the differential calculus that will make the difference between the previous section and this one. The song remains the same.

III.1 Multidimensional Stein's lemma, Stein's equation, Stein's bound

Lemma III.1 : Multidimensional Stein's lemma

Let $N = (N_1, \cdots, N_d)$ a random vector, $C \in \mathcal{S}_n^+(\mathbb{R})$. The following are equivalent :

- **i.** The random vector N follows $\mathcal{N}_d(0, C)$;
- **ii.** For all $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ C^2_{b},

$$\mathbb{E}\left[\langle N, \nabla f(N)\rangle\right] = \mathbb{E}\left[\mathrm{Tr}\left(C \ \mathrm{H} f(N)\right)\right],$$

where $\mathrm{H}f(x)$ is the Hessian matrix of f taken in $x \in \mathbb{R}^d$.

We can prove it by the same way than the one dimensional case. We could use the partial differential equation satisfied by the characteristic function of a $\mathcal{N}(0, C)$:

$$\nabla\varphi(t) = \varphi(t)Ct.$$

Or use the fact that the $\mathcal{N}(0, C)$ is uniquely known by knowing the joint moments. In the following, we consider the *Hilbert-Schmidt norm* :

$$\forall A, B \in \mathcal{M}_d(\mathbb{R}), \langle A, B\rangle_{\mathrm{HS}} \overset{\text{def.}}{=} \mathrm{Tr}\left({}^tAB\right).$$

Definition III.1

Let $C \in \mathcal{S}_n^+(\mathbb{R})$, $N \sim \mathcal{N}_d(0, C)$ and $h : \mathbb{R}^d \longrightarrow \mathbb{R}$ such that $\mathbb{E}[|h(N)|] < +\infty$. We call *Stein's equation* the partial differential equation

$$\mathrm{Tr}(C \ \mathrm{H}f(x)) - \langle x, \nabla f(x)\rangle = h(x) - \mathbb{E}[h(N)],$$

where $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ is the unknown.

Proposition III.1 : A solution in the Lipschitz case

Let $C \in \mathcal{S}_n^+(\mathbb{R})$, $N \sim \mathcal{N}_d(0, C)$ and $h : \mathbb{R}^d \longrightarrow \mathbb{R}$ being K-Lipschitz. Then

$$f_h(x) \overset{\text{def.}}{=} \int_0^{+\infty} \mathbb{E}\left[h(N) - h\left(e^{-t}x + \sqrt{1 - e^{-2t}}N\right)\right] \ \mathrm{d}t$$

is a solution of the Stein's equation. Moreover, f_h satisfies

$$\sup_{x \in \mathbb{R}^d} \|\mathrm{H}f_h(x)\|_{\mathrm{HS}} \leqslant \sqrt{d}H\left\|C^{-1}\right\|_{\mathsf{op}}\|C\|_{\mathsf{op}}^{\frac{1}{2}},$$

where for all $A \in \mathcal{M}_d(\mathbb{R})$:

$$\|A\|_{\mathsf{op}} = \sup_{\|x\|=1} \|Ax\|.$$

Proposition III.2 : Estimation of the Wasserstein distance

Let $F \in L^2$ a random vector. Then, if $N \sim \mathcal{N}(0,C)$:

$$d_{\mathrm{W}}(F,N) \leqslant \sup_{f \in \mathcal{F}_{\mathrm{W}}^d(C)} \left| \mathbb{E}\left[\mathrm{Tr}(C \ \mathrm{H}f(F)) - \langle F, \nabla f(F) \rangle \right] \right|,$$

where

$$\mathcal{F}_{\mathrm{W}}^d(C) \overset{\text{def.}}{=} \left\{ f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}), \sup_{x \in \mathbb{R}^d} \|\mathrm{H}f(x)\|_{\mathrm{HS}} \leqslant \sqrt{d} H \left\| C^{-1} \right\|_{\mathsf{op}} \|C\|_{\mathsf{op}}^{\frac{1}{2}} \right\}.$$

III.2 Estimations on general case with Malliavin calculus

Here's the start of game for multi-dimensional case.

Proposition III.3 : Start of the multi dimensional game.

Let $F = (F_1, \cdots, F_n)$ a random vector with entries taking values in $\mathbb{D}^{1,4}$. Then for all $i,j \in [\![1,d]\!]$, $\left\langle \mathrm{D}F_j, -\mathrm{D}L^{-1}F_i \right\rangle_{\mathcal{H}} \in L^2(\mathbb{P})$. Moreover, for all $C \in \mathcal{S}_d^+(\mathbb{R})$:

$$d_{\mathrm{W}}(F, \mathcal{N}(0,C)) \leqslant \sqrt{d} \left\| C^{-1} \right\|_{\mathsf{op}} \|C\|_{\mathsf{op}}^{\frac{1}{2}} \sqrt{\sum_{i,j=1}^d \mathbb{E}\left[\left(C_{i,j} - \langle \mathrm{D}F_j, -\mathrm{D}L^{-1}F_i \rangle_{\mathcal{H}} \right)^2 \right]}.$$

This proposition gives a first criterion, which looks like the one we derived at the end of the previous section.

Proposition III.4 : A criterion for convergence in law with multi dimensional Malliavin calculus

Let $(F_n)_n$ a sequence of centered random vectors such that for every $i \in [\![1,d]\!]$, $F_n(i) \in \mathbb{D}^{2,4}$. Let $C \in \mathcal{S}_d^+(\mathbb{R})$. If we suppose :

(i) For all $i,j \in [\![1,d]\!]$,

$$\mathbb{E}[F_n(i)F_n(j)] \xrightarrow[n \to +\infty]{} C_{i,j}.$$

(ii) For all $i \in [\![1,d]\!]$:

$$\sup_{n \in \mathbb{N}} \mathbb{E}\left[\|\mathrm{D}F_n(i)\|_{\mathcal{H}}^4 \right] < +\infty.$$

(iii) For all $i \in [\![1,d]\!]$:

$$\mathbb{E}\left[\left\| \mathrm{D}^2 F_n(i) \otimes_1 \mathrm{D}^2 F_n(i) \right\|_{\mathcal{H}^{\otimes 2}} \right] \xrightarrow[n \to +\infty]{} 0.$$

Then $(F_n)_n$ converges in law to $\mathcal{N}(0,C)$.

III.3 Estimations on Wiener chaos

Until now, besides the differential calculus, we did not learn something new about estimations here. We will explain here the new things we have for random vectors with entries in a Wiener chaos.

Theorem III.1 : Convergence in law in Wiener chaos

Let $d \geqslant 1$, $q_1, \cdots, q_d \geqslant 1$ and $F = (F(1), \cdots, F(d)) \in \mathfrak{H}_{q_1} \times \cdots \times \mathfrak{H}_{q_d}$ a centered random vector. Let $C = (\mathbb{E}[F_iF_j])_{i,j} \in \mathcal{S}_d(\mathbb{R})$. If we suppose that $C \in \mathcal{S}_d^{++}(\mathbb{R})$, then there exists a map $\psi : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}$ depending only on the sequence $(q_i)_i$ such that

$$\psi(x,y) \xrightarrow[x\to 0]{} 0,$$

and if we set

$$m(F) \overset{\text{def.}}{=} \psi\left(\begin{pmatrix} \mathbb{E}\left[F(1)^4\right] - 3\mathbb{E}\left[F(1)^2\right]^2 \\ \vdots \\ \mathbb{E}\left[F(d)^4\right] - 3\mathbb{E}\left[F(d)^2\right]^2 \end{pmatrix}, \begin{pmatrix} \mathbb{E}\left[F(1)^2\right] \\ \vdots \\ \mathbb{E}\left[F(d)^2\right] \end{pmatrix}\right),$$

then

$$d_{\mathrm{W}}(F,N) \leqslant \sqrt{d}\left\|C^{-1}\right\|_{\mathsf{op}}\|C\|_{\mathsf{op}}^{\frac{1}{2}} m(F).$$

Note that the constant $m(F)$ only depends on the vector F, and not on C. We can give an explicit expression for ψ, we could deduce it by expending in Wiener chaos the powers of each component of F.

Corollary III.1 : Joint convergence in law

Let $F_n = (F_n(1), \cdots, F_n(d)) \in \mathfrak{H}_{q_1} \times \cdots \times \mathfrak{H}_{q_d}$ a sequence of centered random vectors such that there exists $\sigma_1^2, \cdots, \sigma_d^2 > 0$ satisfying :

$$\forall i \in [\![1,d]\!], F_n(i) \xrightarrow[n\to+\infty]{\text{law}} \mathcal{N}(0,\sigma_i^2)$$

Then there exists a matrix $C \in \mathcal{S}_n^{++}(\mathbb{R})$ such that

$$F_n \xrightarrow[n\to+\infty]{\text{law}} \mathcal{N}(0,C).$$

In other words, *the convergence in law component by component to some Gaussian variables is equivalent in Wiener chaos to the convergence in law of the global vector to some Gaussian vector.*

Moreover, the entries of C are given by the limit of $\mathbb{E}[F_n(i)F_n(j)]$. The next subsection is about a consequence of this strong fact (indeed, in general, converging component by component does not give the convergence of the vector).

III.4 A way to prove central limit theorems

Theorem III.2 : An access to central limit theorem

Let $(F_n)_n$ a sequence of square integrable centered random variables, having its Wiener chaos decomposition :

$$F_n = \sum_{q=1}^{+\infty} I_q(f_{q,n}).$$

We suppose :

a. For all $q \geqslant 1$, there exists $\sigma_q^2 > 0$ such that

$$q!\|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}} \xrightarrow[n\to+\infty]{} \sigma_q^2.$$

b. The series $\sum_q \sigma_q^2$ is convergent. We note σ^2 its sum.

c. For all $q \geqslant 2$ and $r \in [\![1, q-1]\!]$:

$$\|f_{n,q} \otimes_r f_{n,q}\|_{\mathcal{H}^{\otimes(2(q-r))}} \xrightarrow[n\to+\infty]{} 0.$$

d. We have

$$\lim_{N\to+\infty}\left\{\sup_{n\geqslant 1}\sum_{q=N+1}^{+\infty} q!\|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}}\right\}=0.$$

Then, we have

$$F_n \xrightarrow[n\to+\infty]{\text{law}} \mathcal{N}(0,\sigma^2).$$

Let us prove this theorem, which is more useful than it could look like, in the sense that the hypothesis are not so hard that we could think. Indeed, this theorem is the heart of the theorem we present in the following and last section.

Proof : Let $X\sim\mathcal{N}(0,\sigma^2)$. Let us prove the convergence by using the characteristic function. The idea is to consider the truncated sum in the decomposition of F_n is Wiener chaos, proving that the rest will be sufficiently close to zero in order to prevent the convergence in law to N. We consider for $N\in\mathbb{N}^*$:

$$F_{n,N}\overset{\text{def.}}{=}\sum_{q=1}^{N} J_q F_n=\sum_{q=1}^{N} I_q(f_{n,q}),$$

and $X_N\sim\mathcal{N}\left(0,\sum_{q=1}^N\sigma_q^2\right)$. Let $\varepsilon>0$. We cut in three the difference of the characteristic function :

$$\begin{aligned} & \left|\mathbb{E}\left[e^{\mathrm{i}tF_n}\right]-\mathbb{E}\left[e^{\mathrm{i}tX}\right]\right| \\ \leqslant & \left|\mathbb{E}\left[e^{\mathrm{i}tF_n}\right]-\mathbb{E}\left[e^{\mathrm{i}tF_{n,N}}\right]\right| \\ + & \left|\mathbb{E}\left[e^{\mathrm{i}tF_{n,N}}\right]-\mathbb{E}\left[e^{\mathrm{i}tX_N}\right]\right| \\ + & \left|\mathbb{E}\left[e^{\mathrm{i}tX_N}\right]-\mathbb{E}\left[e^{\mathrm{i}tX}\right]\right|. \end{aligned}$$

We denote the sum as (A) + (B) + (C).

A. By factorizing and using $\left|e^{\mathrm{i}\alpha}-1\right|\leqslant|\alpha|$, we have :

$$\text{(A)}\leqslant|t|\;\mathbb{E}\left[\left|F_n-F_{n,N}\right|\right].$$

By Cauchy-Schwarz inequality :

$$\text{(A)}\leqslant|t|\;\mathbb{E}\left[\left|F_n-F_{n,N}\right|^2\right]^{\frac12}.$$

Hence, we have by computing this expectation :

$$\text{(A)}\leqslant|t|\left(\sum_{q=N+1}^{+\infty} q!\|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}}\right)^{\frac12}.$$

Hence, by making appear the hypothesis **d.** :

$$\text{(A)}\leqslant|t|\left(\sup_{n\in\mathbb{N}}\left\{\sum_{q=N+1}^{+\infty} q!\|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}}\right\}\right)^{\frac12}.$$

Since this bound is independent of n, we conclude that

$$\sup_{n\in\mathbb{N}}\text{(A)}\leqslant|t|\left(\sup_{n\in\mathbb{N}}\left\{\sum_{q=N+1}^{+\infty} q!\|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}}\right\}\right)^{\frac12}.$$

This goes to zero when $[N\to+\infty]$ by hypothesis **d.**. We set N big enough (let us say $N\geqslant N_1$) such that (A) $\leqslant\frac{\varepsilon}{3}$.

B. We apply the fourth moment theorem and then the corollary about joint convergence in Wiener chaos. Let $q\in[\![1,N]\!]$. The sequence $(I_q(f_{n,q}))_{n\geqslant1}$ belongs to the q-th chaos $\mathfrak{H}_q$, and satisfies the two following convergence :

$$\mathbb{E}\left[I_q(f_{n,q})^2\right]\xrightarrow[n\to+\infty]{}\sigma_q^2,$$

by hypothesis **a.**, and by **c.** :

$$\|f_{n,q}\otimes_r f_{n,q}\|_{\mathcal{H}^{\otimes q}}\xrightarrow[n\to+\infty]{}0.$$

This gives that $\mathrm{Var}\left(\|\mathrm{D}I_q(f_{n,q})\|^2_{\mathcal{H}}\right)$ goes to zero when $[n\to+\infty]$. Hence, by fourth moment theorem, we have :

$$\forall q\in[\![1,N]\!],I_q(f_{n,q})\xrightarrow[n\to+\infty]{\text{law}}\mathcal{N}(0,\sigma_q^2).$$

But, since $I_q(f_{n,q})$ belongs to a set chaos, by the equivalence between convergence component by component and convergence of the vector to a Gaussian law, we have the existence of $C\in\mathcal{S}_n^{++}(\mathbb{R})$ such that

$$\left(I_1(f_{1,n}),\cdots,I_N(f_{N,n})\right)\xrightarrow[n\to+\infty]{\text{law}}\mathcal{N}(0,C).$$

The entries of C is given by the limit of the covariances of each term. Since two variables in two different chaos have a zero covariance, we conclude that C is the diagonal matrix whose entries are the σ_i^2. By the continuous mapping theorem, we conclude that

$$\sum_{q=1}^N I_q(f_{q,n})\xrightarrow[n\to+\infty]{\text{law}}\mathcal{N}\left(0,\sum_{q=1}^N\sigma_q^2\right).$$

Hence,

$$\mathbb{E}\left[e^{\mathrm{i}tF_{n,N}}-e^{\mathrm{i}tX_N}\right]\xrightarrow[n\to+\infty]{}0,$$

meaning that for n large enough, let us say bigger than a rank n_1, (B) $\leqslant\frac{\varepsilon}{3}$.

C. This is in fact the easiest term to treat since this is the difference between two centered Gaussian random variable. We have

$$\text{(C)}=e^{\frac{-t^2}{2}\sum_{q=1}^N\sigma_q^2}\left|1-\exp\left(\frac{-t^2}{2}\sum_{q=N+1}^{+\infty}\sigma_q^2\right)\right|.$$

By using the fact that $e^{-u}\leqslant1$ when $u\geqslant0$ and $1-e^x\leqslant x$ for $x\in\mathbb{R}$, we have :

$$\text{(C)}\leqslant\frac{t^2}{2}\sum_{q=N+1}^{+\infty}\sigma_q^2.$$

By the hypothesis **b.**, this is the reminder of a convergent series. Hence, since the right hand side is independent of n, we conclude that for N big enough, let us say $N\geqslant N_2$, we have (C) $\leqslant\frac{\varepsilon}{3}$.

CCL. We set $N\geqslant\max\{N_1,N_2\}$. Then, we proved that there exists a rank n_1 such that for every $n\geqslant n_1$:

$$|\varphi_{F_n}(t)-\varphi_X(t)|\leqslant\varepsilon.$$

This means by Paul Lévy's theorem that $(F_n)_n$ converges in law to $\mathcal{N}(0,\sigma^2)$. □

The unfortunate fact here is that we don't have a speed of convergence.

IV Breuer-Major theorem

We will derive a powerful way to prove central limit theorems for a Gaussian stationary process just by knowing its covariance function. We need to introduce a notation before going in.

Definition IV.1

Let $f \in L^2_0(\gamma)$, where γ is the standard Gaussian measure on $\mathbb{R}$, and $L^2_0(\gamma)$ is the subset of $f \in L^2(\gamma)$ such that $\int_{\mathbb{R}} f \, \mathrm{d}\gamma = 0$. We call *Hermite rank of f* by :

$$\text{Hermite rank of } f = \inf\left\{d \in \mathbb{N}^*, \int_{\mathbb{R}} f(x)H_d(x) \, \mathrm{d}\gamma(x) = 0\right\}.$$

If $f = 0$, we set it equal to 0.

In other words, the Hermite rank is the first non zero term in Hermite decomposition of f.

Theorem IV.1 : Breuer-Major theorem

Let $(X_n)_{n\in\mathbb{Z}}$ a centered stationary Gaussian sequence, with covariance function ρ, such that $\mathbb{E}[X_0^2] = 1$. Let $f \in L^2_0(\gamma)$ and we denote by d its Hermite rank. Then, if we suppose that

$$\sum_{v\in\mathbb{Z}} |\rho(n)|^d < +\infty,$$

then

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} f(X_k) \xrightarrow[n\to+\infty]{\text{law}} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = \sum_{q=d}^{+\infty} q! a_q^2 \sum_{v\in\mathbb{Z}} \rho(v)^q,$$

and $a_q = \langle f, H_q \rangle_{L^2(\gamma)}$.

Note : The convergence of the series defining σ is assured by the integrability condition on ρ.
Note : By taking $\rho = \mathbf{1}_{\{0\}}$, we deal with the first central limit theorem, with the hypothesis of independence and distribution.

Lemma IV.1 : Existence of a Hilbert space where we can do Malliavin calculus

Let $(X_n)_{n\in\mathbb{Z}}$ a centered stationary Gaussian sequence with covariance function ρ. Then, there exists a real separable Hilbert space $\mathcal{H}$, a family $(\varepsilon_k)_{k\in\mathbb{Z}}$ of elements of $\mathcal{H}$ and an isonormal Gaussian process X such that :

1. The Hilbert space $\mathcal{H}$ is generated by $(\varepsilon_k)_{k\in\mathbb{Z}}$;
2. For all $k \in \mathbb{Z}$, $X_k = X(\varepsilon_k)$;
3. For all $k, l \in \mathbb{Z}$: $\rho(k - l) = \langle \varepsilon_k, \varepsilon_l \rangle_{\mathcal{H}}$.

Proof of the lemma : • Let $\mathcal{E} \subset \mathbb{R}^{\mathbb{Z}}$ the set of almost null sequences. We consider for all $h, g \in \mathcal{E}$:

$$\langle h, g \rangle_{\mathcal{H}} \stackrel{\text{def.}}{=} \sum_{k,l\in\mathbb{Z}} g_k h_l \rho(k - l).$$

• We define $\mathcal{H}$ as the closure of $\mathcal{E}$ for this norm. $\mathcal{H}$ is real separable Hilbert space. For the separability, we consider $\mathcal{E}_n^{\mathbb{Q}} \subset \mathcal{H}$ the set of almost null sequences such that for every $h \in \mathcal{E}$, we have $h_k \in \mathbb{Q}$ for all $k \in \mathbb{Z}$ and for all $|k| \geqslant n + 1$, $h_k = 0$. Then $\mathcal{E}_n^{\mathbb{Q}}$ is in bijection with $\mathbb{Q}^{2n+1}$ so is countable. So is $\bigcup_{n\in\mathbb{N}} \mathcal{E}_n^{\mathbb{Q}}$. But, $\mathcal{E}$ is included in this set. By taking the closure, we conclude that $\mathcal{H}$ is the closure of a countable set, so $\mathcal{H}$ is indeed separable.

• We set $(\varepsilon_k)_k$ as

$$\varepsilon_k \stackrel{\text{def.}}{=} \mathbf{1}_{\{k\}}.$$

Then, since $\mathcal{E}$ is the algebraic span of the $(\varepsilon_k)_k$, we conclude that

$\mathcal{H}$ is the (Hilbert) span of it. So, we get **(i)**. Moreover, we have indeed by definition that

$$\langle \varepsilon_k, \varepsilon_l \rangle_{\mathcal{H}} = \rho(k-l).$$

- We define for all $h = \sum_{i\in\mathbb{Z}} h_i \varepsilon_i \in \mathcal{E}$:

$$X(h) \overset{\text{def.}}{=} \sum_{i\in\mathbb{Z}} h_i X_i.$$

Note that this is well defined since the sum is finite. Then, we can already observe that X is an isometry on $\mathcal{E}$:

$$\mathbb{E}[X(h)^2] = \sum_{i,j\in\mathbb{Z}} h_i h_j \mathbb{E}[X_i X_j].$$

By definition of the covariance function, since the process is stationary :

$$\mathbb{E}[X(h)^2] = \sum_{i,j\in\mathbb{Z}} h_i h_j \rho(i-j) = \|h\|_{\mathcal{H}}^2.$$

- We extend X on $\mathcal{H}$ by taking the L^2-limit of $(X(h_n))_n$, where $(h_n)_n \subset \mathcal{E}$ is a sequence converging to h.

▷ This limit exists since the sequence $(X(h_n))_n$ is a Cauchy sequence by isometry property : for all $n, p \in \mathbb{N}$:

$$\mathbb{E}\left[|X(h_{n+p}) - X(h_n)|^2\right] = \|h_{n+p} - h_n\|_{\mathcal{H}}^2.$$

▷ Let us prove that it does not depend on the choice of the sequence. Let $(h_n)_n$ and $(g_n)_n$ two sequences of $\mathcal{E}$ such that both converge to h for $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and such that there exists $H, G \in L^2(\mathbb{P})$ such that $(X(h_n))_n$ converges to H in $L^2(\mathbb{P})$ and $(X(g_n))_n$ to G. Let us show that $H = G$. To do that, we simply use the isometry property :

$$\mathbb{E}\left[|X(h_n) - X(g_n)|^2\right] = \|h_n - g_n\|_{\mathcal{H}}^2 \xrightarrow[n\to+\infty]{} 0.$$

Hence $H = G$, so X is well-defined on $\mathcal{H}$.

▷ Let us finally prove that X is an isonormal Gaussian process. We already have the isometry property, and the linearity. We just need to check that $X(h) \sim \mathcal{N}\left(0, \|h\|_{\mathcal{H}}^2\right)$. It is true for $h \in \mathcal{E}$, since $X(h)$ is in this case linear combination of $(X_k)_k$ which is a Gaussian process. But, by using for instance the estimation in Kolmogorov distance between two centered Gaussian random variables, we can conclude that if $(\sigma_n^2)_n$ converges to σ^2 then $\mathcal{N}(0, \sigma_n^2)$ converges in Kolmogorov distance to $\mathcal{N}(0, \sigma^2)$. It means that $X(h) \sim \mathcal{N}\left(0, \|h\|_{\mathcal{H}}^2\right)$. Hence, X is indeed an isonormal Gaussian process. □

Let us begin the proof of the Breuer-Major theorem.

Proof : The plan is quite straightforward, but the computations are very technical. We set the Hilbert space $\mathcal{H}$ as in the previous lemma, and note

$$S_n(f) \overset{\text{def.}}{=} \frac{1}{\sqrt{n}} \sum_{k=1}^n f(X_k).$$

1. We decompose $S_n(f)$ in Wiener chaos.
2. We check every point of the theorem about convergence in law in Wiener chaos explained in previous section.

Let us begin.

1. Since $f \in L^2(\gamma)$, we can expend it in $L^2(\gamma)$ in the Hermite basis :

$$f = \sum_{q=d}^{+\infty} a_q H_q.$$

Hence, in $L^2(\mathbb{P})$:

$$S_n(f) = \frac{1}{\sqrt{n}} \sum_{k=1}^n \sum_{q=d}^{+\infty} a_q H_q(X_k).$$

Hence, we have :

$$S_n(f) = \sum_{q=d}^{+\infty} \frac{a_q}{\sqrt{n}} \sum_{k=1}^n H_q(X_k).$$

Since $\frac{a_q}{\sqrt{n}} \sum_{k=1}^n H_q(X_k) \in \mathfrak{H}_q$, we have our decomposition in Wiener chaos. We can write it with multiple integrals : $S_n(f) = \sum_{q=d}^{+\infty} I_q(f_{n,q})$, with :

$$f_{n,q} \overset{\text{def.}}{=} \frac{a_q}{\sqrt{n}} \sum_{k=1}^n \varepsilon_k^{\otimes q},$$

where $(\varepsilon_k)_k$ is given by the previous lemma. In particular, $f_{n,q} = 0$ for $q \leqslant d-1$.

2. We check one by one every hypothesis of the theorem.

a. Let us show that

$$\forall q \geqslant 1, q! \|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 \xrightarrow[n\to+\infty]{} 0.$$

We simply compute this norm by expending it, and expanding in product the inner product of tensorized functions.

$$\|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 = \frac{a_q^2}{n} \sum_{k,l=1}^n \langle \varepsilon_k, \varepsilon_l \rangle_{\mathcal{H}}^q.$$

By definition of $(\varepsilon_k)_k$:

$$\|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 = \frac{a_q^2}{n} \sum_{k,l=1}^n \rho(k-l)^q.$$

Then, by changing variables and switching (finite) sums, we have :

$$\|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 = \frac{a_q^2}{n} \sum_{l\in\mathbb{Z}} \rho(l)^q \sharp [\![1,n]\!] \cap [\![l+1, l+n]\!].$$

This cardinal can be computed if $l < 0$ or $l > 0$. We have :

$$\begin{aligned} \|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 = & \quad \frac{a_q^2}{n} \sum_{l=0}^{n-1} \rho(l)^q \sharp [\![l+1, n]\!]. \\ & + \frac{a_q^2}{n} \sum_{l=-(n-1)}^{-1} \rho(l)^q \sharp [\![1, n+l]\!]. \end{aligned}$$

So :

$$\begin{aligned} \|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 = & \quad \frac{a_q^2}{n} \sum_{l=0}^{n-1} \rho(l)^q (n-l). \\ & + \frac{a_q^2}{n} \sum_{l=-(n-1)}^{-1} \rho(l)^q (n+l). \end{aligned}$$

In other words :

$$\|f_{n,q}\|_{\mathcal{H}^{\otimes q}}^2 = \frac{a_q^2}{n} \sum_{l=-(n-1)}^{n-1} \rho(l)^q (n - |l|).$$

We write as :

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

But, we have the domination :

$$\left|\rho(l)^q\left(1-\frac{|l|}{n}\right)\mathbf{1}_{\{|l|<n\}}\right| \leqslant |\rho(l)|^q,$$

which is the general term of a convergent series. Hence, by dominated convergence, we conclude that

$$q!\|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}} \xrightarrow[n\to+\infty]{} q!a_q^2\sum_{l\in\mathbb{Z}}\rho(l)^q \stackrel{\text{def.}}{=} \sigma_q^2.$$

b. Let us prove that $\sum_{q=d}^{+\infty}\sigma_q^2 < +\infty$. To do that, we observe first that since $\rho(0)=1$, we have by Cauchy-Schwarz inequality :

$$\forall v\in\mathbb{Z}, |\rho(v)| \leqslant \left(\mathbb{E}[X_0^2]\mathbb{E}[X_v^2]\right)^{\frac{1}{2}} = \rho(0) = 1.$$

So, we have

$$\sum_{q=d}^{+\infty}\sigma_q^2 \leqslant \sum_{q=d}^{+\infty} q!a_q^2\sum_{l\in\mathbb{Z}}|\rho(l)|^d < +\infty.$$

So, $\sigma^2 \stackrel{\text{def.}}{=} \sum_{q=d}^{+\infty}\sigma_q^2$ is finite.

c. Let us prove that for all $q\geqslant d$ and $r\in[\![1,q-1]\!]$:

$$\|f_{n,q}\otimes_r f_{n,q}\|_{\mathcal{H}^{\otimes(2(q-r))}} \xrightarrow[n\to+\infty]{} 0.$$

• First, let us give an expression for this contraction. We have :

$$f_{n,q}\otimes_r f_{n,q} = \left(\frac{a_q}{\sqrt{n}}\right)^2\left(\sum_{k=1}^n\varepsilon_k^{\otimes q}\right)\otimes_r\left(\sum_{l=1}^n\varepsilon_l^{\otimes q}\right).$$

We expend, knowing that :

$$\varepsilon_k^{\otimes q}\otimes_r\varepsilon_l^{\otimes q} = \langle\varepsilon_k,\varepsilon_l\rangle^r_{L^2(T)}\varepsilon_k^{\otimes(q-r)}\otimes\varepsilon_l^{\times(q-r)}.$$

So, we get :

$$f_{n,q}\otimes_r f_{n,q} = \left(\frac{a_q}{\sqrt{n}}\right)^2\sum_{k,l=1}^n\rho(k-l)^r\varepsilon_k^{\otimes(q-r)}\otimes\varepsilon_l^{\otimes(q-r)}.$$

We can compute the norm of it. We call

$$(\text{A}) = \|f_{n,q}\otimes_r f_{n,q}\|^2_{\mathcal{H}^{\otimes(2(q-r))}}.$$

Then, by expanding, we get :

$$\begin{aligned}(\text{A}) = \quad & \frac{a_q^4}{n^2}\sum_{k,i=1}^n\sum_{l,j=1}^n\rho(k-j)^r\rho(i-l)^r \\ & \cdot\left\langle\varepsilon_k^{\otimes(q-r)}\otimes\varepsilon_j^{\otimes(q-r)},\varepsilon_i^{\otimes(q-r)}\otimes\varepsilon_l^{\otimes(q-r)}\right\rangle_{\mathcal{H}^{\otimes(2(q-r))}}.\end{aligned}$$

By the definition of the inner product on a tensor product, we have

$$\begin{aligned}(\text{A}) = \quad & \frac{a_q^4}{n^2}\sum_{k,i=1}^n\sum_{l,j=1}^n\rho(k-j)^r\rho(i-l)^r \\ & \cdot\ \langle\varepsilon_k,\varepsilon_i\rangle^{q-r}_{\mathcal{H}}\langle\varepsilon_j,\varepsilon_l\rangle^{q-r}_{\mathcal{H}}.\end{aligned}$$

And by definition of $(\varepsilon_k)_k$, we get :

$$\begin{aligned}(\text{A}) = \quad & \frac{a_q^4}{n^2}\sum_{k,i=1}^n\sum_{l,j=1}^n\rho(k-j)^r\rho(i-l)^r \\ & \cdot\ \rho(k-i)^{q-r}\rho(l-j)^{q-r}.\end{aligned}$$

But,

$$|\rho(k-i)^{q-r}\rho(k-j)^r| \leqslant |\rho(k-j)|^q + |\rho(k-i)|^q.$$

Hence, we have :

$$\begin{aligned}(\text{A}) \leqslant \quad & \frac{a_q^4}{n^2}\sum_{k,i=1}^n\sum_{l,j=1}^n|\rho(k-j)|^q|\rho(i-l)|^r|\rho(l-j)|^{q-r} \\ & + \frac{a_q^4}{n^2}\sum_{k,i=1}^n\sum_{l,j=1}^n|\rho(k-i)|^q|\rho(i-l)|^r|\rho(l-j)|^{q-r}.\end{aligned}$$

By isolating the sum in k, we get :

$$(\text{A}) \leqslant \frac{2a_q^4}{n^2}\sum_{k\in\mathbb{Z}}|\rho(k)|^q\cdot\sum_{i,j,l=1}^n|\rho(i-l)|^r|\rho(l-j)|^{q-r}$$

By writing :

$$\begin{aligned}&\sum_{i,j,l=1}^n|\rho(i-l)|^r|\rho(l-j)|^{q-r} \\ &= \sum_{l=1}^n\left\{\left(\sum_{i=1}^n|\rho(i-l)|^r\right)\left(\sum_{j=1}^n|\rho(j-l)|^{q-r}\right)\right\},\end{aligned}$$

we can have a nice estimation, using the fact that $|i-l|<n$ and $|j-l|<n$:

$$(\text{A}) \leqslant \frac{2a_q^4}{n}\sum_{k\in\mathbb{Z}}|\rho(k)|^q\sum_{|i|<n}|\rho(i)|^r\sum_{|j|<n}|\rho(j)|^{q-r}.$$

We finally write it as :

$$\begin{aligned}(\text{A}) \leqslant \quad & 2a_q^4\left(\sum_{k\in\mathbb{Z}}|\rho(k)|^q\right) \\ & \cdot\left(\frac{1}{n^{1+\frac{r}{q}}}\sum_{|i|<n}|\rho(i)|^r\right)\left(\frac{1}{n^{1+\frac{q-r}{q}}}\sum_{|j|<n}|\rho(j)|^{q-r}\right).\end{aligned}$$

• So, to conclude in our convergence in zero, it is enough to check that for all $r\in[\![1,q-1]\!]$:

$$\frac{1}{n^{1-\frac{r}{q}}}\sum_{|j|<n}|\rho(j)|^r \xrightarrow[n\to+\infty]{} 0.$$

To do this, we introduce a parameter $\delta\in]0,1[$, and cut the sum as :

$$\sum_{|j|<n} = \sum_{|j|\leqslant\lfloor\delta n\rfloor} + \sum_{\lfloor\delta n\rfloor<|j|<n}.$$

Let $\varepsilon>0$.

▷ For the first sum, we apply Jensen's inequality for $x\longmapsto x^{\frac{q}{r}}$, knowing that there is $2\lfloor\delta n\rfloor+1$:

$$\frac{1}{n^{1-\frac{r}{q}}}\sum_{|j|\leqslant\lfloor\delta n\rfloor}|\rho(j)|^r \leqslant \left(\frac{2\lfloor\delta n\rfloor+1}{n}\right)^{1-\frac{r}{q}}\left[\sum_{|j|\leqslant\lfloor\delta n\rfloor}|\rho(j)|^q\right]^{\frac{r}{q}}.$$

We write it as, by Jensen inequality :

$$\begin{aligned}&\frac{1}{n^{1-\frac{r}{q}}}\sum_{|j|\leqslant\lfloor\delta n\rfloor}|\rho(j)|^r \\ &\leqslant \frac{1}{2^{\frac{r}{q}}}\left((2\delta)^{1-\frac{r}{q}}+\frac{1}{n^{1-\frac{r}{q}}}\right)\left[\sum_{j\in\mathbb{Z}}|\rho(j)|^q\right]^{\frac{r}{q}}.\end{aligned}$$

So, there exists $\delta_0\in(0,1)$ and $N_1\in\mathbb{N}$ such that for every $\delta<\delta_0$ and $n\geqslant N_1$:

$$\frac{1}{n^{1-\frac{r}{q}}}\sum_{|j|\leqslant\lfloor\delta n\rfloor}|\rho(j)|^r \leqslant \frac{\varepsilon}{2}.$$

▷ For the second, we use Jensen inequality again, but with $n - \lfloor \delta n \rfloor + 1$:

$$\frac{1}{n^{1-\frac{r}{q}}} \sum_{\lfloor \delta n \rfloor < |j| < n} |\rho(j)|^r \leqslant \left(\frac{2(n - \lfloor \delta n \rfloor + 1)}{n} \right)^{1-\frac{r}{q}} \left[\sum_{\lfloor \delta n \rfloor < |j| < n} |\rho(j)|^q \right]^{\frac{r}{q}}.$$

So :

$$\frac{1}{n^{1-\frac{r}{q}}} \sum_{\lfloor \delta n \rfloor < |j| < n} |\rho(j)|^r \leqslant 2^{1-\frac{r}{q}} \left(1 - \delta + \frac{2}{n} \right)^{1-\frac{r}{q}} \left[\sum_{\lfloor \delta n \rfloor < |j| < n} |\rho(j)|^q \right]^{\frac{r}{q}}.$$

And since $\delta \in]0,1[$:

$$\frac{1}{n^{1-\frac{r}{q}}} \sum_{\lfloor \delta n \rfloor < |j| < n} |\rho(j)|^r \leqslant 2^{2-\frac{r}{q}} \left[\sum_{\lfloor \delta n \rfloor < |j| < n} |\rho(j)|^q \right]^{\frac{r}{q}}.$$

Since $\sum_{j \in \mathbb{Z}} |\rho(j)|^q < \infty$, there exists a rank $N_2 \in \mathbb{N}$ such that

$$2^{2-\frac{r}{q}} \left[\sum_{\lfloor \delta n \rfloor < |j|} |\rho(j)|^q \right]^{\frac{r}{q}} \leqslant \frac{\varepsilon}{2}.$$

▷ In conclusion, for all $\delta \in]0, \delta_0[$ and $n \geqslant \max\{N_1, N_2\}$, we have

$$\frac{1}{n^{1-\frac{r}{q}}} \sum_{|j|<n} |\rho(j)|^r \leqslant \varepsilon.$$

Since this property is in fact independent of δ, we conclude in the expected convergence to zero. So is $\|f_{n,q} \otimes_r f_{n,q}\|^2_{\mathcal{H}^{\otimes(2(q-r))}}$.

d. Let us finally show that

$$\lim_{N \to +\infty} \sup_{n \geqslant 1} \sum_{q=N+1}^{+\infty} q! \|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}} = 0.$$

To do that, we use the first expended expression of $\|f_{n,q}\|^2$ derived in **a.**. Let $N \in \mathbb{N}$. We call

$$(\mathrm{B}) \overset{\text{def.}}{=} \sum_{q=N+1}^{+\infty} q! \|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}}.$$

Then the expansion gives

$$(\mathrm{B}) = \sum_{q=N+1}^{+\infty} \frac{a_q^2 q!}{n} \sum_{k,l=1}^{n} \rho(k-l)^q.$$

By isolating one of the sum, we can conclude to :

$$(\mathrm{B}) \leqslant \sum_{q=N+1}^{+\infty} a_q^2 q! \sum_{k \in \mathbb{Z}} |\rho(k)|^d.$$

This estimation is independent of n, so the supremum is also bounded by this term. The sum in q is going to zero since the sum of it is $\|f\|^2_{L^2(\gamma)}$. Hence, we have

$$\lim_{N \to +\infty} \sup_{n \geqslant 1} \sum_{q=N+1}^{+\infty} q! \|f_{n,q}\|^2_{\mathcal{H}^{\otimes q}} = 0.$$

CCL. We proved the four points of the theorem to conclude that we have

$$S_n(f) \xrightarrow[n \to +\infty]{\text{law}} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \sum_{q=d}^{+\infty} q! |a_q|^2 \sum_{l \in \mathbb{Z}} |\rho(l)|^q$. □

To improve this theorem, we could wonder how to get a speed of convergence, for the Wasserstein distance for instance, like we get in the central limit theorem by Berry-Esséen. Remember that to have it for the CLT, we had to suppose that the variables have a moment with order 3. Here, it is possible that we need to have less smooth hypothesis that the one we have here. The Stein's method gives here a positive answer to this : for quite stronger hypothesis, it is indeed possible to have a rate of convergence, and better, we can even find the optimal one.

References

— *Normal approximations with Malliavin calculus : from Stein's method to universality*, Ivan Nourdin & Giovanni Peccati, 2012 ;

— *A Second Course in Probability*, Sheldon M. Ross, Erol A. Peköz, 2007 ;

— *Central limits theorems for multiples stochastic integrals and Malliavin calculus*, David Nualart and Salvador Ortiz-Latorre, 2007.