WIENER CHAOS AND LIMIT THEOREMS

LIVIU I. NICOLAESCU

ABSTRACT. I survey the concepts of Gaussian Hilbert spaces, their chaos decomposition and the accompanying Malliavin calculus. I then describe how these ingredients fit in the recent central limit theorems of Nourdin and Peccati [26] in the Wiener chaos context.

CONTENTS

Date: Started October 15, 2013. Completed on November 18, 2015. Last modified on March 14, 2024.
These notes are for myself and whoever else reads this footnote.

NOTATION

- We set
$$\mathbb{N} := \{n \in \mathbb{Z};\ \ n > 0\},\ \ \mathbb{N}_0 := \{n \in \mathbb{Z};\ \ n \geq 0\}.$$
- If $\boldsymbol{V}$ is a real vector space, we denote by $\mathcal{L}(\boldsymbol{V})$ the space of linear operators $\boldsymbol{V} \to \boldsymbol{V}$.
- $\boldsymbol{1}_A$ denotes the characteristic function of a subset A of a set S,
$$\boldsymbol{1}_A : S \to \{0,1\},\ \ \boldsymbol{1}_A(a) = \begin{cases} 1, & a \in A, \\ 0, & a \in S \setminus A. \end{cases}$$
- For $k \in \mathbb{N}_0 \cup \{\infty\}$ and $n \in \mathbb{N}$ we denote by $C_b^k(\mathbb{R}^n)$ the space of C^k-functions with bounded derivatives of order $\leq k$.
- We will write $N \sim \mathcal{N}(m, v)$ to indicate that N is a normal random variable with mean m and variance v.
- If C is a symmetric, nonnegative definite $m \times m$ matrix, we write $N \sim \mathcal{N}(0, C)$ to indicate that N is an $\mathbb{R}^m$-valued Gaussian random vector with mean 0 and covariance form C.
- If $f : \mathbb{R}^m \to \mathbb{R}$ is a twice differentiable function, then we denote by $\mathrm{Hess}(f)$ its ***Hessian***.

INTRODUCTION

A few years ago I learned about central limit results of the type Breuer and Major [4] pioneered in the early 80s and I had the distinct feeling that such results would be helpful in my search of a CLT concerning the distribution of critical points of random functions. The work of Kratz and Léon [15, 16] convinced me that my initial suspicion had merits. I started these notes with the goal of learning Major's techniques and the subsequent developments.

The common set-up of the Breuer-Major type theorems is that of a Gaussian Hilbert space so the first part of these notes is devoted to this concept. I was greatly influenced by S. Janson's elegant and comprehensive book on this subject [12] and by Major's notes [19]. For a novice in probability such as myself, Major's book was extremely helpful since it discusses the type of Gaussian Hilbert spaces canonically associated to a random function on $\mathbb{R}^n$, and it offers a lot of intuition about the Wiener chaos decomposition of L^2, nonlinear functionals on Gaussian Hilbert spaces.

Part 1 of these notes discusses several important facts about such spaces, and in particular, it introduces the concept of multiple Ito integral and Wiener chaos decomposition. The diagram formula plays an important role in this story. I found Janson's approach in [12] the easiest to digest and it is the one I have included in Part 1.

While I was learning this subject, and coping with many other academic duties, I did not pay close attention to things that had developed or were developing in this area. I am talking here about the new approach to central limit theorems being developed by I. Nourdin, G. Peccati and others[1] based on a very elegant and convenient marriage between Malliavin calculus and the Stein method. The recent paper of Estrade and Léon [10] was a wake-up call. Luck would have it, Nourdin and Peccati published their excellent monograph [26] which made my learning job easier, dramatically changed my thinking and the way I thought about the organization of the paper.

Part 2 is devoted to Malliavin calculus. I follow [26] rather closely, though in some parts I followed points of view in Bogachev [3] and Janson [12] that looked more appealing to me. I have to say that Malliavin's magnum opus [20] has had a great influence on me. I found his monograph very difficult to penetrate due to its terseness and somewhat confusing notation. My struggle with [20] had a nontrivial impact though. Malliavin's elegance and concision are hard to match. Many of his approaches and points of view are very versatile and I have subconsciously adopted them.

[1] I apologize in advance to anyone whom I have omitted.

Part 3, is the reward for the theoretical foundations in the first two parts. Here I followed unashamedly Nourdin and Peccati, [26]. In Section 11 I discuss the basics of Stein's method method, while in Section 12 I describe how to blend the Malliavin calculus in Part 2 with the Stein method to produce the key central limit result, Theorem 12.15.

I make no claim of originality. These notes will certainly reflect the fact that my mathematical background is not that of a probabilist, but I hope they will convey my recent found fascination with the probabilistic thinking. These notes represent no substitute for the references from which they draw their inspiration, but may perhaps ease the way of a beginner such as myself into a more profound investigation. I worked on-off for two years on these notes. Typos and naiveté aside, I am overall pleased on how these notes turned out and I am posting them with the hope that somebody else will benefit from this effort. (I will update them, fix typos or clumsy explanations, enhance some arguments that do not seem as clear in hindsight.)

As I mentioned earlier, the motivation for learning these techniques came from my investigations of critical points of random functions. This learning process had an immediate payoff for me. In the span of two months I was able to solve some problems that resisted my efforts for several years, [23, 24]. That is a most enthusiastic endorsement of the power of the results in [26]!

PART 1. GAUSSIAN HILBERT SPACES AND CHAOS DECOMPOSITIONS

1. FINITE DIMENSIONAL GAUSSIAN MEASURES

1.1. **Gaussian measures and random vectors.** Denote by $\boldsymbol{\gamma}_1$ the standard Gaussian measure on $\mathbb{R}$ given by

$$\boldsymbol{\gamma}_1(dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

For any $v \geq 0$ and $\mu \in \mathbb{R}$ we denote by $\boldsymbol{\gamma}_{v,\mu}$ the probability measure on $\mathbb{R}$ given by

$$\boldsymbol{\gamma}_{v,\mu}(dx) = \begin{cases} \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-\mu)^2}{2v}} dx, & v > 0 \\ \delta_\mu, & v = 0, \end{cases}$$

where δ_μ denotes the Dirac measure supported at μ. When $\mu = 0$, we use the simpler notation $\boldsymbol{\gamma}_v := \boldsymbol{\gamma}_{v,\mu=0}$. The probability measure $\boldsymbol{\gamma}_{v,\mu}$ is called *the Gaussian measure on* $\mathbb{R}$ *with mean* μ *and variance* v. A real valued random variable X is called *Gaussian* if its probability distribution $\mathbb{P}_X$ is a Gaussian measure $\boldsymbol{\gamma}_{v(X),\mu(X)}$, i.e., for any Borel set $B \subset \mathbb{R}$

$$\mathbb{P}[X \in B] = \boldsymbol{\gamma}_{v(X),\mu(X)}(B).$$

The quantities $\mu(X)$ and $v(X)$ are respectively the mean and the variance of X. The Gaussian random variable is called *centered* if its mean is zero.

Remark 1.1. The above definition has one æsthetical flaw: it is "coordinate dependent". One can define the concept of Gaussian random variable without explicitly describing its probability distribution. More precisely an integrable random variable X is a mean zero Gaussian random variable if given two independent random variables Y, Z with the same distribution as X, then, for any $\theta \in [0, 2\pi]$, the random variable $(\cos\theta)Y + (\sin\theta)Z$ has the same distribution as X. For a proof we refer to [35, Sec. 2.2.1]. □

For $t \geq 0$ we denote by $\mathcal{R}_t$ the rescaling map $\mathbb{R} \to \mathbb{R}$ given by $\mathcal{R}_t(x) = tx$. Then

$$\boldsymbol{\gamma}_v = \left(\mathcal{R}_{\sqrt{v}}\right)_{\#} \boldsymbol{\gamma}_1,$$

where $(\mathcal{R}_t)_\#$ denotes the pushforward operation on measures induced by the map $\mathcal{R}_t$. For $\mu \in \mathbb{R}$ we denote by $\mathcal{T}_\mu$ the translation $\mathbb{R} \to \mathbb{R}$ given by $\mathcal{T}_\mu(x) = \mu + x$. Then

$$\boldsymbol{\gamma}_{v,\mu} = \big(\mathcal{T}_\mu\big)_\# \boldsymbol{\gamma}_v = \big(\mathcal{T}_\mu\big)_\# \big(\mathcal{R}_{\sqrt{v}}\big)_\# \boldsymbol{\gamma}_1.$$

Given a sequence $(v_k, \mu_k) \in \mathbb{R}_{\geq 0} \times \mathbb{R}$, the sequence of Gaussian measures $(\boldsymbol{\gamma}_{v_k,\mu_k})$ converges weakly if and only if the sequence (v_k, μ_k) converges to some (v, μ) as $k \to \infty$. If this happens, then

$$\boldsymbol{\gamma}_{v_k,\mu_k} \Rightarrow \boldsymbol{\gamma}_{v,\mu} \;\text{ as } k \to \infty.$$

where "$\Rightarrow$" denotes the weak convergence of probability measures.

The Fourier transform of the measure $\boldsymbol{\gamma}_{v,\mu}$ is the function $\widehat{\boldsymbol{\gamma}}_{v,\mu} : \mathbb{R} \to \mathbb{C}$ given by

$$\widehat{\boldsymbol{\gamma}}_{v,\mu}(\xi) = \int_{\mathbb{R}} e^{\boldsymbol{i}\xi x} \boldsymbol{\gamma}_{v,\mu}(dx) = e^{\boldsymbol{i}\mu\xi - \frac{v}{2}|\xi|^2}, \;\; \boldsymbol{i} := \sqrt{-1}. \tag{1.1}$$

Example 1.2. Suppose that X is a real valued, centered Gaussian random variable with variance v defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then,

$$\mathbb{E}\big[\, e^{tX} \,\big] = e^{\frac{vt^2}{2}}. \tag{1.2}$$

Indeed,

$$\begin{aligned}\mathbb{E}\big[\, e^{tX} \,\big] &= \frac{1}{\sqrt{2\pi v}} \int_{\mathbb{R}} e^{tx} e^{-x^2/(2v)} dt = e^{-\frac{x^2 - 2vtx + v^2t^2}{2v}} \cdot e^{tv^2/2} dx \\ &= \frac{e^{tv^2/2}}{\sqrt{2\pi v}} \int_{\mathbb{R}} e^{-\frac{(c-vt)^2}{2v}} dt = e^{tv^2/2}.\end{aligned}$$

In particular, we deduce that

$$e^{tX} \in L^p(\Omega, \mathcal{F}, \mathbb{P}), \;\; \forall p \in [1, \infty), \tag{1.3}$$

$$\mathbb{E}\big[\, X^n \,\big] = \begin{cases} 0, & n \equiv 1 \bmod 2 \\ v^k 1 \cdot 3 \cdots (2k-1), & n = 2k. \end{cases} \tag{1.4}$$

$\square$

Suppose that $\boldsymbol{X}$ is a finite dimensional real vector space. Denote by $\mathcal{B}_{\boldsymbol{X}}$ the σ-algebra of the Borel subsets of $\boldsymbol{X}$. A Gaussian measure on $\boldsymbol{X}$ is a Borel probability measure $\boldsymbol{\gamma}$ on $\boldsymbol{X}$ such that, for any $\xi \in \boldsymbol{X}^\vee := \operatorname{Hom}(\boldsymbol{X}, \mathbb{R})$, the pushforward $\xi_\# \boldsymbol{\gamma}$ is a Gaussian measure $\boldsymbol{\gamma}_{v(\xi),\mu(\xi)}$ on $\mathbb{R}$ with mean $\mu(\xi)$ and variance $v(\xi)$. More precisely

$$\mu(\xi) := \mathbb{E}[\xi] = \int_{\boldsymbol{X}} \xi(x) \boldsymbol{\gamma}(dx), \;\; v(\xi) = \int_{\boldsymbol{X}} \big(\, \xi(x) - \mu(\xi) \,\big)^2 \boldsymbol{\gamma}(dx).$$

The linearity of the expectation of a random variable implies that the map

$$\boldsymbol{X}^\vee \ni \xi \mapsto \mu(\xi)$$

is linear, and thus defines a point $\mu = \mu_{\boldsymbol{\gamma}} \in \boldsymbol{V}$ called the *mean* or *barycenter* of $\boldsymbol{\gamma}$.

Equivalently, after fixing a norm on $\boldsymbol{X}$, we can define μ_γ as the Bochner integral

$$\mu_{\boldsymbol{\gamma}} := \int_{\boldsymbol{X}} x \boldsymbol{\gamma}(dx).$$

The map

$$\boldsymbol{X}^\vee \ni \xi \mapsto v(\xi) \in \mathbb{R}$$

is a nonnegative definite quadratic form and thus defines a symmetric, nonnegative definite bilinear form

$$C = C_{\boldsymbol{\gamma}} : \boldsymbol{X}^\vee \times \boldsymbol{X}^\vee \to \mathbb{R},$$

called the *covariance form* of $\boldsymbol{\gamma}$. The Gaussian measure $\boldsymbol{\gamma}$ is called *centered* if $\mu_{\boldsymbol{\gamma}} = 0$.

The Fourier transform of $\boldsymbol{\gamma}$ is the function $\widehat{\boldsymbol{\gamma}} : \boldsymbol{X}^\vee \to \mathbb{C}$ given by

$$\widehat{\boldsymbol{\gamma}}(\xi) = \mathbb{E}\big[\, e^{\boldsymbol{i}\xi} \,\big] = \int_{\boldsymbol{X}} e^{\boldsymbol{i}\langle \xi, x\rangle} \boldsymbol{\gamma}(dx) = e^{\boldsymbol{i}\langle \xi, \mu_\gamma\rangle - \frac{1}{2}C_\gamma(\xi,\xi)}.$$

An $\boldsymbol{X}$-valued random variable is called *Gaussian* if its probability distribution is a Gaussian measure on $\boldsymbol{X}$.

Observe that if $(\boldsymbol{\gamma}_n)$ is a sequence of Gaussian measures on $\boldsymbol{X}$ with barycenters μ_n, covariance forms C_n, and converging weakly to a probability measure $\boldsymbol{\gamma}$, then for any $\xi \in \boldsymbol{X}^\vee$ the Gaussian measures $\xi_{\#}\boldsymbol{\gamma}_n$ on $\mathbb{R}$ converge weakly to $\xi_{\#}\boldsymbol{\gamma}$. Hence, $\xi_{\#}\boldsymbol{\gamma}$ is a Gaussian measure for any linear functional $\xi : \boldsymbol{X} \to \mathbb{R}$. Thus, the limiting measure $\boldsymbol{\gamma}$ is also Gaussian. Moreover

$$\mu_{\boldsymbol{\gamma}} = \lim_n \mu_{\boldsymbol{\gamma}_n}, \;\; C_{\boldsymbol{\gamma}} = \lim_n C_{\boldsymbol{\gamma}_n}.$$

If $\boldsymbol{Y}$ is another finite dimensional real vector space, $A : \boldsymbol{X} \to \boldsymbol{Y}$ is a linear map and $\boldsymbol{\gamma}$ is a Gaussian measure on $\boldsymbol{X}$, then the pushforward $A_{\#}\boldsymbol{\gamma}$ is a Gaussian measure on $\boldsymbol{Y}$ with mean

$$\mu_{A_{\#}\boldsymbol{\gamma}} = A\mu_{\boldsymbol{\gamma}},$$

and covariance form $C_{A_{\#}\boldsymbol{\gamma}}$ described by

$$C_{A_{\#}\boldsymbol{\gamma}}(\eta_1, \eta_2) = C_{\boldsymbol{\gamma}}(A^\vee \eta_1, A^\vee \eta_2), \;\; \forall \eta_1, \eta_2 \in \boldsymbol{Y}^\vee,$$

where $A^\vee : \boldsymbol{Y}^\vee \to \boldsymbol{X}^\vee$ is the dual or transpose of $A : \boldsymbol{X} \to \boldsymbol{Y}$.

Remark 1.3. If the vector space $\boldsymbol{X}$ is equipped with a Euclidean inner product $(-,-)$, and $\boldsymbol{\gamma}$ is a centered Gaussian measure on $\boldsymbol{X}$ with covariance form C, then we can identify $\boldsymbol{X}^\vee$ with $\boldsymbol{X}$ and the resulting symmetric bilinear form C with a nonnegative symmetric operator $S : \boldsymbol{X} \to \boldsymbol{X}$,

$$C(x_1, x_2) = (Sx_1, x_2), \;\; \forall x_1, x_2 \in \boldsymbol{X}.$$

We will denote by $\boldsymbol{\Gamma}_S$ this Gaussian measure.

If S happens to be invertible, then the measure $\boldsymbol{\Gamma}_S$ admits the more explicit description

$$\boldsymbol{\Gamma}_S(dx) = \frac{1}{\sqrt{\det 2\pi S}} e^{-\frac{1}{2}(S^{-1}x, x)} dx. \qquad \square$$

In general, if $\boldsymbol{\gamma}$ is a centered Gaussian measure on the n-dimension real vector space and

$$C : \boldsymbol{X}^\vee \times \boldsymbol{X}^\vee \to \mathbb{R}$$

is its covariance form, then we can choose a basis $\xi_1, \dots, \xi_n$ of $\boldsymbol{X}^\vee$ that diagonalizes C, i.e.,

$$\mathbb{E}\big[\, \xi_i \xi_j \,\big] = \begin{cases} 0, & i \neq j \\ v_i \in \{0, 1\}, & i = j. \end{cases}.$$

The basis $\{\xi_1, \dots, \xi_n\}$ of $\boldsymbol{X}^\vee$ determines a dual basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ of $\boldsymbol{X}$, the forms ξ_k are then the coordinates on $\boldsymbol{X}$ determined by the basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ and $\boldsymbol{\gamma}$ can be identified with the product measure on $\mathbb{R}^n$

$$\boldsymbol{\gamma}(d\xi) = \bigotimes_{k=1}^n \boldsymbol{\gamma}_{v_k}(d\xi_k). \tag{1.5}$$

$\square$

Definition 1.4. A *Gaussian random field* parametrized by the set S is a family $(X_s)_{s\in S}$ of real valued random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that, for any finite subset $F \subset S$ the $\mathbb{R}^F$-valued random vector

$$\Omega \ni \omega \mapsto \big(X_s(\omega) \big)_{s\in F} \in \mathbb{R}^F$$

is Gaussian. The random field is called *finite*, if the set S is finite, and *centered*, if all the random variables X_s, $s \in S$, are centered. □

Remark 1.5. Suppose that $(X_k)_{1\leq k\leq n}$ is a centered Gaussian random field such that

$$\mathbb{E}[X_i X_j] = 0, \;\; \forall i \neq j.$$

The discussion in Remark 1.3, especially (1.5), shows that the random variables $(X_k)_{1\leq k\leq n}$ are independent. In other words, in the Gaussian world, uncorrelated random variables are independent.□

1.2. **Wick's formula.** Suppose we are given a centered, finite Gaussian random field, $(X_k)_{1\leq k\leq n}$.

A *Feynman diagram* on these random variables is a graph Γ with n vertices labeled $X_1, \dots, X_n$ such that any vertex is connected to at most one other vertex. In other words, a Feynman diagram is a partial matching of the random variables X_k. The *weight* $w(\boldsymbol{e})$ of an edge $\boldsymbol{e} = [X_i, X_j]$ of a Feynman diagram is the correlation $w(\boldsymbol{e}) := \mathbb{E}[X_i X_j]$.

If Γ is a Feynman diagram, we denote by $\mathcal{E}(\Gamma)$ the set of edges of Γ, and by $\mathcal{I}(\Gamma)$ the set of *isolated* vertices. The number of edges of a Feynman diagram Γ is called the *rank* of the diagram and it is denoted by $r(\Gamma)$.

A Feynman diagram is called *complete* if it has no isolated vertices, i.e., each vertex is connected with exactly one other vertex. The *weight* of a Feynman diagram Γ is the *random variable*

$$w(\Gamma) := \left(\prod_{\boldsymbol{e}\in\mathcal{E}(\Gamma)} w(\boldsymbol{e}) \right) \prod_{k\in\mathcal{I}(\Gamma)} X_k.$$

When the diagram is complete, its weight is the real number

$$w(\Gamma) := \prod_{\boldsymbol{e}\in\mathcal{E}(\Gamma)} w(\boldsymbol{e}),$$

where $\mathcal{E}(\Gamma)$ denotes the set of edges of Γ.

Lemma 1.6. *Denote by $d_n(r)$ the number of Feynman diagram of rank r with n vertices.*

$$d_n(r) = \binom{n}{n-2r}(2r-1)!! = \frac{n!}{2^r (n-2r)! r!} \tag{1.6}$$

Proof. Let us first observe that

$$d_n(r) = \binom{n}{n-2r} d_{2r}(r).$$

Next observe that

$$d_{2r}(r) = (2r-1) d_{2r-2}(r-1).$$

Indeed, in a complete Feynman diagram with $2r$ vertices the vertex labelled 1 is connected with a unique other vertex. Therefore there are $(2r-1)$ way of producing an edge that has vertex 1 as one of his edges. After removing this edge we are left with a complete Feyman diagram with $(2r-2)$ edges. □

We have *Wick's formula*

$$\boxed{\mathbb{E}\big[\,X_1\cdots X_n\,\big]=\sum_{\Gamma}w(\Gamma)}, \tag{1.7}$$

where the summation is carried over all the complete Feynman diagrams on the variables $X_1,\dots,X_n$. Note that if n is odd, then the sum in the right-hand-side of (1.7) is trivial.

To prove (1.7) we first observe that

$$\mathbb{E}[X_1\cdots X_n]=\frac{1}{n!}\frac{\partial^n}{\partial t_1\cdots\partial t_n}\Big|_{t_1=\cdots=t_n=0}\mathbb{E}\big[\,(t_1X_1+\cdots+t_nX_n)^n\,\big].$$

Next, we observe that $t_1X_1+\cdots+t_nX_n$ is a centered Gaussian variable with variance

$$v(t_1,\dots,t_n)=\sum_{i,j=1}^n\mathbb{E}[X_iX_j]t_it_j=\sum_{j=1}^n\mathbb{E}[X_j^2]t_j^2+2\sum_{i<j}\mathbb{E}[X_iX_j]t_it_j.$$

If we let $n=2k$, we deduce from (1.4) that

$$\mathbb{E}\big[\,(t_1X_1+\cdots+t_{2k}X_{2k})^{2k}\,\big]=(2k-1)!!\left(\sum_{i,j=1}^{2k}\mathbb{E}[X_iX_j]t_it_j\right)^k.$$

so that

$$\begin{aligned}\mathbb{E}[X_1\cdots X_{2k}]&=\frac{(2k-1)!!}{(2k)!}\frac{\partial^{2k}}{\partial t_1\cdots\partial t_{2k}}\Big|_{t_1=\cdots=t_{2k}=0}\left(\sum_{i,j=1}^{2k}\mathbb{E}[X_iX_j]t_it_j\right)^k\\&=(2^kk!)\frac{(2k-1)!!}{(2k)!}\sum_{\Gamma}w(\Gamma)=\sum_{\Gamma}w(\Gamma).\end{aligned}$$

2. Gaussian Hilbert spaces and their Fock spaces

Definition 2.1. A *Gaussian linear space* is a vector space $\mathscr{X}$ of real random variables defined on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$ such that any random variable $X\in\mathscr{X}$ is centered Gaussian. If the vector space $\mathscr{X}$ is closed in $L^2(\Omega,\mathcal{F},\mathbb{P})$, then we say that $\mathscr{X}$ is a *Gaussian Hilbert space*. □

Example 2.2. If $\boldsymbol{X}$ is a finite dimensional real vector space equipped with a centered Gaussian measure, then its dual $\boldsymbol{X}^\vee$ is canonically a Gaussian Hilbert space. □

Example 2.3 (The Main Example)**.** Suppose that $(X_t)_{t\in T}$ is a centered Gaussian random field parameterized by the set T. Thus, there exists a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and a map

$$X:T\times\Omega\to\mathbb{R},\;\;T\times\Omega\ni(t,\omega)\mapsto X_t(\omega)\in\mathbb{R}$$

with the following properties.

(i) For any $t\in T$, the map $X_t:\Omega\to\mathbb{R}$ is measurable.
(ii) For any $n\in\mathbb{N}$ and any $t_1,\dots,t_n\in T$ the random vector $(X_{t_k})_{1\leq k\leq n}\in\mathbb{R}^m$ is centered Gaussian.

The closed subspace $\mathscr{X}\subset L^2(\Omega,\mathcal{F},\mathbb{P})$ generated by the collection $(X_t)_{t\in T}$ is called the *Gaussian Hilbert space associated to the random field* $(X_t)_{t\in T}$. We can then view the random field as a map

$$T\ni t\mapsto X_t\in\mathscr{X}.$$

The space $\mathbb{R}^T$ of functions $T \to \mathbb{R}$ is equipped with a natural σ-algebra, namely the smallest σ-algebra such that all the natural projection $\pi_t : \mathbb{R}^T \to \mathbb{R}$ are measurable. For every $\omega \in \Omega$ we denote by $X(\omega)$ the function $T \to \mathbb{R}$ given by $t \mapsto X_t(\omega)$. The resulting map

$$\Omega \ni \omega \mapsto X(\omega) \in \mathbb{R}^T$$

a probability measure on the space of functions $\mathbb{R}^T$. Thus, we can view a random field parametrized by T as a random function on T.

Conversely, any probability measure μ on $\mathbb{R}^T$ such that all the projections π_t are measurable, then we can regard $\mathbb{R}^T$ as a sample space and the projection π_t as a random function on T. Additionally, if for any finite subset $S \subset T$ the pushforward on $\mathbb{R}^S$ is a centered Gaussian measure, we get a tautological Gaussian random field parametrized by T,

$$X : T \times \mathbb{R}^T \to \mathbb{R}, \ \ (t, \mathbb{R}^T) \ni (t, f) \mapsto X_t(f) := f(t) \in \mathbb{R}. \qquad \square$$

Definition 2.4. An *isonormal Gaussian process* is a triplet $(H, \mathscr{X}, W)$ where $\mathscr{X}$ is a Gaussian Hilbert space, H is a Hilbert space and $W : H \to \mathscr{X}$ is an isomorphism of Hilbert spaces. The map W is called the *white noise map* of the isonormal process. □

Example 2.5. Suppose that H is a separable, real Hilbert space with inner product $(-,-)_H$. A *Gaussian measure on* H is a Borel probability measure $\boldsymbol{\Gamma}$ such that, for any $h \in H$, the linear functional $L_h : H \to \mathbb{R}$, $L_h(x) = (h, x)$, is a centered Gaussian random variable. In particular, the collection $(L_h)_{h\in H}$ is a Gaussian random field parameterized by H.

We denote by $C(h_1, h_2)$ the covariance of L_{h_1}, L_{h_2},

$$C(h_1, h_2) = \mathbb{E}\big[\, L_{h_1} L_{h_2} \,\big].$$

This defines an inner product on $H^* = \operatorname{Hom}(H, \mathbb{R})$. As explained in [8], there exists a symmetric, nonnegative *trace class* operator Q such that

$$C(h_1, h_2) = \big(\, Qh_1, h_2 \,\big)_H, \ \ \forall h_1, h_2 \in H.$$

Assume for simplicity that $\ker Q = 0$.

To this Gaussian measure we can associate the Gaussian Hilbert space $H^*_{\boldsymbol{\Gamma}}$ defined as the closure in $L^2(H, \boldsymbol{\Gamma})$ of the vector space spanned by $(L_h)_{h\in H}$. Note that we have a continuous map with dense image

$$L : H \to H^*_{\boldsymbol{\Gamma}}, \ \ h \mapsto L_h. \tag{2.1}$$

The Hilbert space $H^*_{\boldsymbol{\Gamma}}$ is canonically isomorphic with H as a Hilbert space. To construct this isomorphism consider the dense subspace $Q^{1/2}H$.

$$W : Q^{1/2}H \to L^2(H, \boldsymbol{\Gamma}), \ \ Q^{1/2}H \ni z \mapsto W_z := L_{Q^{-1/2}z}.$$

Clearly the image of W is equal to the image of the map L in (2.1). Observe that

$$\mathbb{E}\big[\, W_{z_1} W_{z_2} \,\big] = (z_1, z_2)_H, \ \ \forall z_1, z_2 \in Q^{-1/2}H.$$

This shows that the map W extends by continuity to an isometry $W : H \to H^*_{\boldsymbol{\Gamma}}$. This isomorphism of Hilbert spaces is called the *white noise map*. Observe that the triplet $(H, H^*_{\boldsymbol{\Gamma}}, W)$ is an isonormal Gaussian process.

The subspace $Q^{1/2}H \subset H$ is called the *Cameron-Martin space*. If we identify H with its topological dual we observe that $H^* = H \subset H^*_{\boldsymbol{\Gamma}}$. One could think of the elements of $H^*_{\boldsymbol{\Gamma}}$ as measurable linear functional $H \to \mathbb{R}$.

We fix an orthonormal $(e_k)_{k\in\mathbb{N}}$ (complete) basis of H consisting of eigenvectors of Q,

$$Qe_n = \lambda_n e_n, \ \ n \in \mathbb{N}.$$

The collection of linear functionals

$$W_{e_n} = \frac{1}{\sqrt{\lambda_n}} L_{e_n}, \quad n \in \mathbb{N}$$

is an orthonormal basis of the associated Gaussian Hilbert space $H^*_{\boldsymbol{\Gamma}}$. □

Suppose that $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Gaussian Hilbert space. We denote by $\widehat{\mathcal{F}}$ the σ-subalgebra of $\mathcal{F}$ generated by the collection of random variables $X \in \mathscr{X}$ and we define

$$\widehat{\mathscr{X}} := L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P}) \subset L^2(\Omega, \mathcal{F}, \mathbb{P}).$$

For reasons that will become clear a bit later, we will refer to $\widehat{\mathscr{X}}$ as the *Fock space*[2] of $\mathscr{X}$. To understand what happens when we pass from a Gaussian Hilbert space $\mathscr{X}$ to its Fock space $\widehat{\mathscr{X}}$ we consider first the simplest possible case, $\dim \mathscr{X} = 1$.

2.1. **Hermite polynomials.** Consider the standard Gaussian measure $\mathbb{P} = \boldsymbol{\gamma}_1$ on $\Omega = \mathbb{R}$,

$$\boldsymbol{\gamma}_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

As explained in Example 2.2, this tautologically defines a one-dimensional Gaussian Hilbert space $\mathscr{X}_1$ spanned by the identity function $\mathbb{1}_{\mathbb{R}}$. In this case $\mathcal{F}$ is the σ-algebra $\mathcal{B}_{\mathbb{R}}$ of the Borel subsets of $\mathbb{R}$ and $\widehat{\mathcal{F}} = \mathcal{B}_{\mathbb{R}}$. Moreover, we have

$$L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P}) = \widehat{\mathscr{X}_1} = L^2(\mathbb{R}, \gamma_1).$$

We see in this example that the Fock space $\widehat{\mathscr{X}_1}$ is much larger than $\mathscr{X}_1$. A convenient orthogonal basis of $\widehat{\mathscr{X}_1} = L^2(\mathbb{R}, \boldsymbol{\gamma}_1)$ is given by the *Hermite polynomials* $(H_n)_{n \geq 0}$, [20, V.1.3].

To define these polynomials we introduce the *creation operator* $\delta_x : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$,

$$\delta_x f(x) = -e^{\frac{x^2}{2}} \partial_x \big(e^{-\frac{x^2}{2}} f(x) \big) = -\partial_x f(x) + x f(x). \tag{2.2}$$

The creation operator is the formal adjoint with respect to the inner product in $L^2(\mathbb{R}, \boldsymbol{\gamma}_1)$ of the usual differential operator ∂_x , i.e.,

$$\int_{\mathbb{R}} f'(x) g(x) \boldsymbol{\gamma}_1(dx) = \int_{\mathbb{R}} f(x) \delta_x g(x) \boldsymbol{\gamma}_1(dx), \quad \forall f, g \in C_0^\infty(\mathbb{R}).$$

Then

$$H_n(x) = \delta_x^n 1. \tag{2.3}$$

Equivalently,

$$\partial_x^n \big(e^{-\frac{x^2}{2}} \big) = (-1)^n H_n(x) e^{-\frac{x^2}{2}}.$$

Let us observe that the operators ∂_x, δ_x satisfy the Heisenberg identity

$$\boxed{[\partial_x, \delta_x] = \mathbb{1}}.$$

Using this iteratively we deduce

$$\partial_x H_n(x) = n H_{n-1}(x), \quad \forall n \in \mathbb{N}, \tag{2.4a}$$

$$\delta_x \partial_x H_n(x) = n H_n(x), \quad \forall n \in \mathbb{N}. \tag{2.4b}$$

From the defining equation (2.3) we obtain the recurrence relation

$$H_{n+1}(x) = \delta_x H_n(x) = -H_n'(x) + x H_n(x), \quad \forall n \geq 0. \tag{2.5}$$

[2] The "Fock-space" terminology does not seem to be very used in probabilistic circles, but it is what it is.

For example,

$$H_0(x)=1,\ \ H_1(x)=x,\ \ H_2(x)=x^2-1,\ \ H_3(x)=x^3-3x,$$

$$H_4(x)=x^4-6x^2+3,\ \ H_5(x)=x^5-10\,x^3+15\,x,\ \ H_6(x)=x^6-15\,x^4+45\,x^2-15.$$

More generally

$$H_n(x)=n!\sum_{m=0}^{\lfloor\frac{n}{2}\rfloor}\frac{(-1)^m}{2^m m!(n-2m)!}x^{n-2m}\stackrel{(1.6)}{=}\sum_{m=0}^{\lfloor\frac{n}{2}\rfloor}(-1)^m d_n(m)x^{n-2m}. \tag{2.6}$$

Observe that the leading coefficient of $H_n(x)$ is 1. Note that for $|x|\geq 1$ we have

$$|H_n(x)|\leq\sum_{m=0}^{n}\binom{n}{m}|x|^{n-m}\left(\frac{1}{2}\right)^m=\left(\frac{1}{2}+|x|\right)^n,\ \ \forall|x|\leq 1, \tag{2.7a}$$

$$|H_n(x)|\leq\left(\frac{3}{2}\right)^n,\ \ \forall|x|\leq 1. \tag{2.7b}$$

From the equalities (2.3) and (2.4a) we deduce that the collection $(H_n)_{n\geq 0}$ is orthogonal in $L^2(\mathbb{R},\boldsymbol{\gamma}_1)$,

$$\int_{\mathbb{R}}H_m(x)H_n(x)\boldsymbol{\gamma}_1(dx)=\delta_{nm}n!. \tag{2.8}$$

Moreover, the collection $(H_n)_{n\geq 0}$ spans a dense subspace in $L^2(\mathbb{R},\boldsymbol{\gamma}_1)$ so that any $f\in L^2(\mathbb{R},\boldsymbol{\gamma}_1)$ admits a Hermite decomposition

$$f=\sum_{n\geq 0}c_n(f)H_n(x),\ \ c_n(f)=\frac{1}{n!}\int_{\mathbb{R}}f(x)H_n(x)\boldsymbol{\gamma}_1(dx).$$

Let us point out that if $g\in C^\infty(\mathbb{R})$ has the property that

$$g^{(k)}\in L^2(\mathbb{R},\boldsymbol{\gamma}_1),\ \ \forall k\geq 0,$$

then we have the following expansion in $L^2(\mathbb{R},\boldsymbol{\gamma}_1)$

$$g(x)=\sum_{n\geq 0}\frac{1}{n!}\mathbb{E}_{\boldsymbol{\gamma}_1}\big[\,g^{(n)}\,\big]H_n(x), \tag{2.9}$$

where $\mathbb{E}_{\gamma_1}$ denotes the expectation with respect to the probability measure $\boldsymbol{\gamma}_1$. If in the above equality we choose

$$g(x)=g_\lambda(x)=e^{\lambda x-\frac{\lambda^2}{2}},$$

Then

$$g_\lambda^{(n)}(x)=\lambda^n e^{\lambda x-\frac{\lambda^2}{2}},\ \ \mathbb{E}_{\boldsymbol{\gamma}_1}\big[\,g_\lambda^{(n)}\,\big]=\lambda^n e^{-\frac{\lambda^2}{2}}\int_{\mathbb{R}}e^{\lambda x}\boldsymbol{\gamma}_1(dx)\stackrel{(1.2)}{=}\lambda^n.$$

This proves that

$$\sum_{n\geq 0}H_n(x)\frac{\lambda^n}{n!}=e^{\lambda x-\frac{\lambda^2}{2}}=g_\lambda(x), \tag{2.10}$$

where the above series converges in $L^2(\mathbb{R},\boldsymbol{\gamma}_1)$ for any $\lambda\in\mathbb{C}$. The estimates (2.7a) and (2.7b) show that the above series also converges uniformly on the compacts of $\mathbb{R}\times\mathbb{C}$.

2.2. **Hermite decompositions.** Suppose that $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a *separable* Gaussian Hilbert space. Fix a complete orthonormal base $(X_n)_{n\geq 1}$ of $\mathscr{X}$. In particular, we have

$$\mathbb{E}\big[\, X_i X_j \,\big] = \delta_{ij} = \text{the Kronecker } \delta,$$

and thus the random variables $(X_n)_{n\geq 1}$ are independent. Additionally, each of the random variables X_n is Gaussian with mean zero and variance 1, and the σ-algebra $\sigma(X_1, X_2, \dots)$ generated by the collection (X_n) coincides with the σ-algebra $\widehat{\mathcal{F}}$.

Consider the space $\mathbb{R}^{\mathbb{N}}$, equipped with the product measure[3]

$$\boldsymbol{\gamma}_1^{\mathbb{N}} = \bigotimes_{n\in\mathbb{N}} \boldsymbol{\gamma}_1(dx_n),$$

defined on the Borel σ-algebra $\mathcal{B}^{\mathbb{N}}$ of the space $\mathbb{R}^{\mathbb{N}}$ equipped with the product topology.

We have a natural map

$$\vec{X} : \Omega \to \mathbb{R}^{\mathbb{N}}, \;\; \omega \mapsto \big(\, X_1(\omega), X_2(\omega), \dots, \,\big).$$

Then

$$\widehat{\mathcal{F}} = \sigma(X_1, X_2, \dots,) = \vec{X}^{-1}(\mathcal{B}^{\mathbb{N}})$$

and

$$\vec{X}_{\#}(\mathbb{P}) = \boldsymbol{\gamma}_1^{\mathbb{N}}.$$

This yields an isomorphism of Hilbert spaces

$$\widehat{\mathscr{X}} = L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P}) \to L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\gamma}_1^{\mathbb{N}}).$$

Moreover, a function $f : \Omega \to \mathbb{R}$ is $\sigma(X_1, X_2, \dots,)$-measurable if and only if there exists a $\mathcal{B}^{\mathbb{N}}$-meansurable function $F : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, such that

$$f(\omega) = F\big(\, X_1(\omega), X_2(\omega), \dots \big), \;\; \forall \omega \in \Omega.$$

Additionally $f \in \widehat{\mathscr{X}}$ iff $F \in L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\gamma}_1^{\mathbb{N}})$.

We can construct an orthonormal basis of $L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\gamma}_1^{\mathbb{N}})$ as follows. For any multi-index

$$\alpha = (\alpha_1, \alpha_2, \dots) \in \mathbb{N}_0^{\mathbb{N}}$$

such the $\alpha_k = 0$ for k sufficiently large, we consider the multi-variable polynomial

$$H_\alpha(\underline{x}) = \prod_{k\in\mathbb{N}} H_{\alpha_k}(x_k), \;\; \underline{x} = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}}.$$

The collection H_α thus obtained is a complete orthogonal basis of $L^2(\mathbb{R}^{\mathbb{N}}, \boldsymbol{\gamma}_1^{\mathbb{N}})$ and (2.8) shows that

$$\|\, H_\alpha \,\|^2_{L^2(\mathbb{R}^N, \boldsymbol{\gamma}_1^{\mathbb{N}})} = \alpha! := \prod_{k=1}^{\infty} \alpha_k!. \tag{2.11}$$

[3]For a construction of countable products of probability measures we refer to [14, Sec.14.3].

3. THE WIENER CHAOS DECOMPOSITION

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a separable Gaussian Hilbert space. We want to describe a coordinate independent orthogonal decomposition of the Fock space $\widehat{\mathscr{X}}$ that is closely related to the coordinate dependent Hermite decomposition described above.

Proposition 3.1. *The vector space*

$$\operatorname{span}_{\mathbb{R}}\{\xi_1 \cdots \xi_n;\quad n \in \mathbb{N},\ \ \xi_1, \dots, \xi_n \in \mathscr{X}\ \}$$

is dense in $\widehat{\mathscr{X}} = L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P})$.

Proof. Fix a complete orthonormal basis $X_1, X_2, \dots, X_n, \dots$ of $\mathscr{X}$. We will prove that

$$\mathcal{P}_{\mathbb{C}} := \operatorname{span}_{\mathbb{C}}\Big\{ X_1^{\alpha_1} \cdots X_n^{\alpha_n};\quad n \in \mathbb{N},\ \ \alpha_1, \dots, \alpha_n \in \mathbb{N}_0\ \Big\}$$

is dense in $\widehat{\mathscr{X}_{\mathbb{C}}} = \widehat{\mathscr{X}} \oplus \boldsymbol{i}\widehat{\mathscr{X}}$. We follow the approach in the proof of [12, Thm.2.6].

Denote by $\widehat{\mathcal{P}}_{\mathbb{C}}$ the closure of $\mathcal{P}_{\mathbb{C}}$ in $\widehat{\mathscr{X}_{\mathbb{C}}}$. We will prove prove that

$$\widehat{\mathscr{X}_{\mathbb{C}}} \subset \widehat{\mathcal{P}}_{\mathbb{C}}.$$

Set

$$\mathcal{V}_n = \operatorname{span}\Big\{ X_1, X_2, \dots, X_n, \Big\},\ \ \mathcal{V} = \bigcup_{n\geq 1} \mathcal{V}_n.$$

The result follows from the following two facts.

A. $e^{\boldsymbol{i}X} \in \widehat{\mathscr{X}_{\mathbb{C}}} = \widehat{\mathscr{X}} + \boldsymbol{i}\widehat{\mathscr{X}}$ for any $X \in \mathscr{X}$.
Proof. Let $X \in \mathscr{X}$. Then

$$e^{\boldsymbol{i}X} = \sum_{k=0}^{\infty} \frac{\boldsymbol{i}^k}{k!} X^k,$$

where the above series converges in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. This proves that $e^{\boldsymbol{i}X} \in \widehat{\mathscr{X}_{\mathbb{C}}}$.
B. If $Z \in \widehat{\mathscr{X}_{\mathbb{C}}}$ and $\mathbb{E}(Ze^{\boldsymbol{i}X}) = 0$, $\forall X \in \mathcal{V}$, then $Z = 0$.
Proof of B. We set

$$\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n),$$

so we get a filtration of σ-algebras

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$$

such that

$$\widehat{\mathcal{F}} = \bigcup_{n=1}^{\infty} \mathcal{F}_n. \tag{3.1}$$

Suppose that $Z \in \widehat{\mathscr{X}_{\mathbb{C}}}$ and $\mathbb{E}(Ze^{\boldsymbol{i}X}) = 0$, $\forall X \in \mathcal{V}$. We set

$$Z_n := \mathbb{E}\big[\, Z | \mathcal{F}_n \,\big].$$

The definition of conditional expectation implies that

$$\mathbb{E}[\, Z_n e^{\boldsymbol{i}X}\,] = 0,\ \ \forall X \in \mathcal{V}_n.$$

Now observe that since $Z_n \in L^2(\Omega, \mathcal{F}_n, \mathbb{P})$ we have

$$Z_n(\omega) = \varphi_n(\, X_1(\omega), \dots, X_n(\omega)\,)$$

for some $\varphi \in L^2(\mathbb{R}^n, \boldsymbol{\gamma}_1^n)$. We deduce that

$$\mathbb{E}\big[\, \varphi_n(X_1, \dots, X_n) e^{\boldsymbol{i}t_1 X_1 + \cdots + \boldsymbol{i}t_n X_n}\,\big] = 0,\ \ \forall t_1, \dots, t_n \in \mathbb{R}.$$

In other words, the Fourier transform of the complex valued measure

$$\varphi_n(x_1,\dots,x_n)\boldsymbol{\gamma}_1(dx_1)\cdots\boldsymbol{\gamma}_n(dx_n)$$

is trivial so that $\varphi_n=0$. Hence $Z_n=0$, $\forall n\in\mathbb{N}$, i.e.,

$$\mathbb{E}[\,Z|\mathcal{F}_n\,]=0,\;\;\forall n\in\mathbb{N}.$$

Using (3.1), we deduce from Doob's Martingale Convergence theorem[4]

$$Z=\mathbb{E}\big[\,Z|\widehat{\mathcal{F}}\,\big]=\lim_{n\to\infty}\mathbb{E}[\,Z|\mathcal{F}_n\,]=0.$$

□

For $n\in\mathbb{N}_0$ we define $\mathcal{P}_n(\mathscr{X})$ to be the closure in $\widehat{\mathscr{X}}$ of the subspace

$$\big\{\,p(\xi_1,\dots,\xi_m);\;\; m>0,\;\;\xi_1,\dots,\xi_m\in\mathscr{X},\;\;p\in\mathbb{R}[x_1,\dots,x_m],\;\;\deg p\leq n\,\big\}.$$

Proposition 3.1 shows that the vector space

$$\mathcal{P}(X)=\bigcup_{n\geq 0}\mathcal{P}_n(\mathscr{X}),$$

is dense in $\widehat{\mathscr{X}}$. Clearly $\mathcal{P}_{n-1}(\mathscr{X})\subset\mathcal{P}_n(\mathscr{X})$. We denote by $\mathscr{X}^{:n:}$ the orthogonal complement of $\mathcal{P}_{n-1}(\mathscr{X})$ in $\mathcal{P}_n(\mathscr{X})$. We deduce that

$$\widehat{\mathscr{X}}=\widehat{\bigoplus_{n\geq 0}}\mathscr{X}^{:n:},\tag{3.2}$$

where the direct sum in the right-hand-side indicates a Hilbert-complete direct sum, i.e.,

$$\xi\in\widehat{\bigoplus_{n\geq 0}}\mathscr{X}^{:n:}\Longleftrightarrow\xi=(\xi_n)_{n\geq 0},\;\;\xi_n\in\mathscr{X}^{:n:},\;\;\sum_{n\geq 0}\|\xi_n\|^2_{L^2}<\infty.$$

The decomposition (3.2) is called the *Wiener chaos decomposition* of $\widehat{\mathscr{X}}$. We will denote by Proj_n the orthogonal projection $\widehat{\mathscr{X}}\to\mathscr{X}^{:n:}$.

Example 3.2. Suppose that $\mathscr{X}$ is the 1-dimensional Gaussian Hilbert space generated by a standard Gaussian random variable ξ with mean 0 and variance 1. In this case

$$\mathcal{P}_n(\mathscr{X})=\operatorname{span}_{\mathbb{R}}\{H_k(\xi);\;\;k\leq n\}.$$

Since $\mathbb{E}[H_j(\xi)H_k(\xi)]=0$ for $j\neq k$, we deduce that

$$\mathscr{X}^{:n:}=\operatorname{span}\{H_n(\xi)\,\}.$$

Moreover, (2.9) implies that, $\forall n\geq 0$ we have

$$\xi^n=\sum_{k=0}^{n}\binom{n}{k}\mathbb{E}[\xi^{n-k}]H_k(\xi)=\sum_{j=0}^{\lfloor n/2\rfloor}\binom{n}{2j}\mathbb{E}[\xi^{2j}]H_{n-2j}(\xi)\tag{3.3}$$

$$\stackrel{(1.4)}{=}\sum_{j=0}^{\lfloor n/2\rfloor}(2j-1)!!\binom{n}{2j}H_{n-2j}(\xi).$$

In particular,

$$\operatorname{Proj}_n(\xi^n)=H_n(\xi).\tag{3.4}$$

[4]This is a very special case of the martingale convergence theorem. However, even this special case is rather nontrivial because it implies (a form of) Lebesgue's differentiation theorem. We refer to [34, Cor. 5.2.4] for a proof.

If $\mathscr{X}$ is a separable Gaussian Hilbert space and $\underline{X} = (X_n)_{n\geq 1}$ is a complete orthonormal basis of $\mathscr{X}$, then the computations in Example 2.2 show that $\mathcal{P}_n(\mathscr{X})$ is the closure of the subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ spanned by the "polynomials"

$$H_\alpha(\underline{X}) = H_\alpha(X_1, X_2, \dots),$$

where $\alpha \in \mathbb{N}_0^{\mathbb{N}}$ is any multi-index such that $|\alpha| \leq n$, where we recall that

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots. \qquad \square$$

From the above example we obtain the following useful consequence.

Corollary 3.3. *Suppose that $\underline{X} = (X_k)_{k\geq 1}$ is a complete orthonormal basis of $\mathscr{X}$. Then the collection*

$$H_\alpha(\underline{X}), \;\; \alpha \in \mathbb{N}_0^{\mathbb{N}}, \;\; |\alpha| = n,$$

is an orthogonal basis of $\mathscr{X}^{:n:}$. $\square$

4. Wick products and the diagram formula

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a separable real Gaussian Hilbert space $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $\widehat{\mathscr{X}}$ the Fock space of $\mathscr{X}$, and consider the Wiener chaos decomposition

$$\widehat{\mathscr{X}} = \widehat{\bigoplus_{n\geq 0}} \mathscr{X}^{:n:}.$$

As usual, we denote by Proj_n the orthogonal projection $\widehat{\mathscr{X}} \to \mathscr{X}^{:n:}$. We have bilinear maps

$$\mathscr{X}^{:m:} \times \mathscr{X}^{:n:} \to \mathscr{X}^{:(m+n):}, \;\; \mathscr{X}^{:m:} \times \mathscr{X}^{:n:} \ni (\xi, \eta) \mapsto \xi \bullet \eta := \mathrm{Proj}_{m+n}(\xi\eta).$$

Remark 4.1. If $\underline{X} = (X_k)_{k\geq 1}$ is a complete orthonormal basis of $\mathscr{X}$, and $\alpha, \beta \in \mathbb{N}_0$ are such that $|\alpha| = m$, $|\beta| = n$, then

$$H_\alpha(\underline{X}) \bullet H_\beta(\underline{X}) = H_{\alpha+\beta}(\underline{X}). \tag{4.1}$$

Indeed

$$H_\alpha(\underline{X}) \bullet H_\beta(\underline{X}) = \sum_{|\gamma|=m} c_\gamma H_\gamma(\underline{X}).$$

Now observe that for any multi-index γ such that $|\gamma| = m + n$, and $\gamma \neq \alpha + \beta$ the coefficient of $\underline{X}^{\alpha+\beta}$ in $H_\gamma(\underline{X})$ is 0, while the coefficient of $\underline{X}^{\alpha+\beta}$ in $H_{\alpha+\beta}(\underline{X})$ is 1. $\square$

Definition 4.2. Fix n random variables $\xi_1, \dots, \xi_n \in \mathscr{X}$ and a polynomial $P \in \mathbb{R}[x_1, \dots, x_n]$ of degree m. The random variable $\mathrm{Proj}_m P(\xi_1 \cdots \xi_n)$ is called the *Wick polynomial* associated to P and it is denoted by $: P(\xi_1 \cdots \xi_n) :$. $\square$

Theorem 4.3. *Let $\xi_1, \dots, \xi_n \in \mathscr{X}$. Then*

$$\boxed{: \xi_1 \cdots \xi_n : \; = \sum_\Gamma (-1)^{r(\Gamma)} w(\Gamma)}, \tag{4.2}$$

where the summation is over all Feynman diagrams with vertices labelled by $\xi_1, \dots, \xi_n$.

Proof. Denote by $L(\xi_1,\dots,\xi_m)$ the left-hand-side of (4.2) and by $R(\xi_1,\dots,\xi_m)$ the right-hand side. Observe that both L and R are symmetric multi-linear forms in the variables $\xi_1,\dots,\sigma_m$ and thus

$$L(\xi_1,\dots,\xi_m)=R(\xi_1,\dots,\xi_m),\quad \forall\xi_1,\dots,\xi_m)\Longleftrightarrow L(\underbrace{\xi,\dots,\xi}_{m})=L(\underbrace{\xi,\dots,\xi}_{m}),\ \ \forall\xi.$$

Let$\xi\in\mathscr{X}$ such that $\operatorname{Var}\big[\,\xi\,\big]=1$. Then

$$L(\underbrace{\xi,\dots,\xi}_{m})=:\xi^m:\stackrel{(3.4)}{=}H_m(\xi).$$

We set $\mathbb{I}_m=\{1,\dots,m\}$. For any finite set S we denote $D_*(S)$ the set of complete Feynmans diagram with vertices in S. Then

$$R(\underbrace{\xi,\dots,\xi}_{m})=\sum_{r\geq 0}(-1)^r\left(\sum_{\substack{S\subset\mathbb{I}_m\\ \#S=2r}}\ \sum_{\Gamma\in D_*(S)}\right)\xi^{m-2r}$$

Now observe

$$d_m(r)=\left(\sum_{\substack{S\subset\mathbb{I}_m\\ \#S=2r}}\ \sum_{\Gamma\in D_*(S)}\right),$$

where $d_m(r)$ denotes the number of diagrams of rank r on a set of m vertices. Hence

$$R(\underbrace{\xi,\dots,\xi}_{m})=\sum_{r\geq 0}(-1)^r d_m(r)\xi^{m-2r}\stackrel{(2.6)}{=}H_m(\xi)=L(\underbrace{\xi,\dots,\xi}_{m}).$$

□

The equality (4.2) has the following immediate consequence.[5]

Corollary 4.4. *Consider n-random variables $\xi_1,\dots,\xi_n\in\mathscr{X}$. For any nonnegative integer m we define*

$$\mathcal{P}_m(\xi_1,\dots,x_m):=\operatorname{span}\big\{\xi^\alpha:=\xi_1^{\alpha_1}\cdots\xi_n^{\alpha_m};\ \ \alpha\in\mathbb{Z}^n_{\geq 0},\ \ |\alpha|\leq m\,\big\}\subset\mathcal{P}_m(\mathscr{X}).$$

and we denote by $\mathscr{X}^{:m:}(\xi_1,\dots,\xi_m)$ the orthogonal complement of $\mathcal{P}_{m-1}(\xi_1,\dots,\xi_n)$ in the space $\mathcal{P}_m(\xi_1,\dots,\xi_n)$.

If $P\in\mathbb{R}[x_1,\dots,x_n]$ is a real polynomial of degree m, then $:P(\xi_1,\dots,\xi_n):$ is equal to the orthogonal projection of $P(\xi_1,\dots,\xi_n)$ on $\mathscr{X}^{:m:}(\xi_1,\dots,\xi_n)$. □

Corollary 4.4 and (3.3) imply the following result.

Corollary 4.5. *Suppose that $\xi_1,\dots,\xi_n\in\mathscr{X}$ is an orthonormal system, i.e.,*

$$\mathbb{E}[\xi_i\xi_j]=\delta_{ij},\ \ \forall i,j.$$

Then for any $\alpha\in\mathbb{N}_0^n$ we have

$$:\xi_1^{\alpha_1}\cdots\xi_n^{\alpha_n}:=H_\alpha(\xi_1,\dots,\xi_n).$$

[5] For a more direct proof of Corollary 4.4 we refer to [19, Prop. 2.2].

Proof. Set $m := |\alpha|$. Since the random variables $\xi_1, \dots, \xi_n$ are independent, we deduce that

$$\mathbb{E}[\, H_\beta(\xi_1, \dots, \xi_n) H_\gamma(\xi_1, \dots, \xi_n)\,] = \prod_{j=1}^n \mathbb{E}[\, H_{\beta_j}(\xi_j) H_{\gamma_j}(\xi_j)\,], \;\; \forall \beta, \gamma \in \mathbb{Z}^n_{\geq 0}.$$

We deduce from the orthogonality of the Hermite polynomials that the collection

$$\big(\, H_\beta(\xi_1, \dots, \xi_n)\,\big)_{|\beta| \leq m}$$

is an orthogonal basis of $\mathcal{P}_m(\xi_1, \dots, \xi_n)$. In particular, we have a unique linear decomposition

$$\xi^\alpha = \sum_{|\beta| \leq m} c_\beta H_\beta(\xi), \tag{4.3a}$$

$$:\xi^\alpha := \sum_{|\beta| = m} c_\beta H_\beta(\xi) \tag{4.3b}$$

For any multi-index β such that $|\beta| = m$, the coefficient of ξ^β in the right-hand-side of (4.3a) is c_β. We deduce that $c_\beta = 0$ for all β such that $|\beta| = m$ and $\beta \neq \alpha$. The conclusion of Corollary 4.5 is now obvious. □

Corollary 4.6. *The space*

$$\operatorname{span}\big\{\, (:\xi_1\xi_2 \cdots \xi_n:)\,;\;\; \xi_1, \dots, \xi_n \in \mathscr{X}\,\big\}$$

is dense in $\mathscr{X}^{:n:}$.

Proof. Follows from Example 3.2 and Corollary 4.5. □

Theorem 4.7 (Diagram Formula)**.** *Consider an array of random variables*

$$(\xi) = \big\{\xi_{ij} \in \mathscr{X};\;\; 1 \leq i \leq k,\;\; 1 \leq j \leq \ell_i\,\big\}.$$

Denote by $D'(\xi)$ the collection of ***Feynman diagrams compatible with the array***

$$(i,j),\;\; 1 \leq i \leq k,\;\; 1 \leq j \leq \ell_i.$$

This means that

- *its vertices are labeled by the variables ξ_{ij}, and*
- *no edge connects variables situated on the same row of the array ξ.*

We let $D'_c(\xi)$ to denote the sub collection of $D'(\xi)$ consisting of complete diagrams. For $i = 1, \dots, k$ we set

$$Y_i := \operatorname{Proj}_{\ell_i}\left(\prod_{j=1}^{\ell_i} \xi_{ij}\right),\;\; \ell = \ell_1 + \cdots + \ell_k.$$

In other words, Y_i is the Wick product of the variables situated on the i-th row. Then

$$\boxed{\mathbb{E}[Y_1 \cdots Y_k] = \sum_{\Gamma \in D'_c(\xi)} w(\Gamma)}, \tag{4.4a}$$

$$\boxed{Y_1 \cdots Y_k = \sum_{\Gamma \in D'(\xi)} (:w(\Gamma):) = \sum_{\Gamma \in D'(\xi)} \operatorname{Proj}_{\ell - 2r(\Gamma)} w(\Gamma)}. \tag{4.4b}$$

Proof. We follow the approach in [12, Thm. 3.12, 3.15]. Denote by D_i the Feynman diagrams with vertices on the i-th row of the array and by $D_c(\xi)$ the collection of all complete Feynman diagrams with vertices in the array. Theorem 4.3 implies

$$Y_i = \sum_{\Gamma_i \in D_i} (-1)^{r(\Gamma_i)} w(\Gamma_i),$$

$$\begin{aligned} Y_1 \cdots Y_k &= \prod_{i=1}^{k} \left(\sum_{\Gamma_i \in D_i} (-1)^{r(\Gamma_i)} w(\Gamma_i) \right) \\ &= \sum_{(\Gamma_1,\dots,\Gamma_k) \in D_1 \times \cdots \times D_k} (-1)^{\sum_{i=1}^k r(\Gamma_i)} \prod_{i=1}^{k} w(\Gamma_i), \end{aligned}$$

so that

$$\mathbb{E}[Y_1 \cdots Y_k] = \sum_{(\Gamma_1,\dots,\Gamma_k) \in D_1 \times \cdots \times D_k} (-1)^{\sum_{i=1}^k r(\Gamma_i)} \mathbb{E}\left[\prod_{i=1}^{k} w(\Gamma_i) \right].$$

Given $(\Gamma_1, \dots, \Gamma_k) \in D_1 \times \cdots \times D_k$ we denote by $D(\Gamma_1, \dots, \Gamma_k)$ the subcollection of $D_c(\xi)$ that contains $\Gamma_1 \cup \cdots \cup \Gamma_k$ as a sub diagram. We deduce from Wick's formula (1.7) that

$$\mathbb{E}\left[\prod_{i=1}^{k} w(\Gamma_i) \right] = \sum_{\Gamma' \in D(\Gamma_1,\dots,\Gamma_k)} w(\Gamma')$$

Hence

$$\mathbb{E}[Y_1 \cdots Y_k] = \sum_{\Gamma'} w(\Gamma') \underbrace{\left(\sum_{\substack{(\Gamma_1,\dots,\Gamma_k) \in D_1 \times \cdots \times D_k \\ \Gamma_1 \cup \cdots \cup \Gamma_k \subset \Gamma'}} (-1)^{\sum_{i=1}^k r(\Gamma_i)} \right)}_{=:S(\Gamma')}$$

Now observe that

$$S(\Gamma') = \prod_{i=1}^{k} \underbrace{\left(\sum_{\substack{\Gamma_i \in D_i \\ \Gamma_i \subset \Gamma'}} (-1)^{r(\Gamma_i)} \right)}_{=:S_i(\Gamma')}$$

Now observe that $S_i(\Gamma') = 0$ if Γ' has edges connecting vertices on the i-th row, and it is $= 1$ otherwise. Thus

$$S(\Gamma') = \begin{cases} 1, & \Gamma' \in D_c'(\xi), \\ 0, & \Gamma' \in D(\xi) \setminus D_c'(\xi). \end{cases}$$

This proves (4.4a). Denote by L, respectively R the left-hand-side respectively the right-hadn-side of the equality (4.4b). The (4.4a) which implies that for any random variables

$$\eta_1, \dots, \eta_m \in \operatorname{span}\big\{ \xi_{ij} \in \mathscr{X},;\;\; 1 \le i \le k,\;\; 1 \le j \le \ell_i \big\}$$

we have

$$\mathbb{E}[LZ] = \mathbb{E}[RZ], \;\; Z := (: \eta_1 \cdots \eta_m :).$$

The equality (4.4b) now follows from Corollary 4.4. □

Example 4.8. Let us apply the diagram formula in the special case when the array (ξ) consists of two rows and $\xi_{ij} = \xi$, $\forall (ij)$, $\mathbb{E}[\xi^2] = 1$. Assume that $\ell_1 \geq \ell_2$. Then

$$Y_i = H_{\ell_i}(\xi)$$

and we deduce

$$H_{\ell_1}(\xi)H_{\ell_2}(\xi) = \sum_{\Gamma \in D'(\xi)} H_{\ell_1+\ell_2-2r(\Gamma)}(\xi) = \sum_{r=0}^{\ell_2} r!\binom{\ell_1}{r}\binom{\ell_2}{r} H_{\ell_1+\ell_2-2r}(\xi).$$

More generally, assume the array has two lines, but the variables on the first line are equal to ξ_1, while the variables on the second line are equal to ξ_2, $\mathbb{E}[\xi_1^2] = \mathbb{E}[\xi_2]^2$. Then, if we set $c := \mathbb{E}[\xi_1\xi_2]$, we deduce

$$H_{\ell_1}(\xi_1)H_{\ell_2}(\xi_2) = \sum_{r=0}^{\ell} r!\binom{\ell_1}{r}\binom{\ell_2}{r} c^r \operatorname{Proj}_{\ell_1+\ell_2-2r}(\xi_1^{\ell_1-r}\xi_2^{\ell_2-r}). \tag{4.5}$$

If $\ell_1 = \ell_2 = \ell$, then (4.4a) Implies that

$$\mathbb{E}\big[\, H_{\ell_1}(\xi_1)H_{\ell_2}(\xi_2)\,\big] = \ell!\binom{2\ell}{\ell} c^{\ell}. \tag{4.6}$$

□

The equality (4.4b) implies[6] that for any positive integer n there exists a constant $C(n) > 0$ such that for any $X \in \mathcal{P}_n(\mathscr{X})$ we have

$$\|X\|_{L^4} \leq C(n)\|X\|_{L^2}.$$

In particular, this shows that the bilinear map

$$\mathscr{X}^{:m:} \times \mathscr{X}^{:n:} \ni (X, Y) \mapsto X \bullet Y := \operatorname{Proj}_{m+n}(XY) \in \mathscr{X}^{:m+n:}$$

is continuous. Corollary 3.3 now implies that the multiplication $\bullet$ satisfies the associativity property

$$(\xi \bullet \eta) \bullet \zeta = \xi \bullet (\eta \bullet \zeta), \;\; \forall \xi \in \mathscr{X}^{:\ell:},\; \eta \in \mathscr{X}^{:m:},\; \zeta \in \mathscr{X}^{:n:}, \;\; \forall \ell, m, n \in \mathbb{N}_0. \tag{4.7}$$

Indeed, (4.1) shows that the above equality is true for

$$\xi, \eta, \zeta \in \big\{\, H_\alpha(\underline{X});\;\; \alpha \in \mathbb{Z}^{\mathbb{N}}_{\geq 0},\;\; |\alpha| < \infty \,\big\}.$$

The general case follows from the multi-linearity and continuity of (4.7) in ξ, η, ζ. A similar argument shows that

$$\xi \bullet \eta = \eta \bullet \xi, \;\; \forall \xi \in \mathscr{X}^{:m:},\;\; \eta \in \mathscr{X}^{:n:}. \tag{4.8}$$

We thus obtain a structure of commutative and associative $\mathbb{R}$-algebra on $\mathscr{X}$ called the *Wick algebra* of $\mathscr{X}$. The product $\bullet$ is called the *Wick product*. In general, if

$$\xi = \sum_{n\geq 0} \xi_n, \;\; \eta = \sum_{n \geq 0} \eta_n, \;\; \xi_n, \eta_n \in \mathscr{X}^{:n:},$$

then

$$\xi \bullet \eta := \sum_{n \geq 0} \left(\sum_{j+k=n} \xi_j \bullet \eta_k \right).$$

[6] See [12, Lemma 3.44] for details.

5. Tensor products and the Fock space

5.1. **Tensor products of separable Hilbert spaces.** The tensor product of two Hilbert spaces H_1, H_2 is defined as follows. Construct the *algebraic* tensor product $H_1 \otimes H_2$. The universality property of the tensor product implies that there exists a unique inner product $(-,-)_{H_1\otimes H_2}$ of $H_1 \otimes H_2$ such that, for any $x_i, y_i \in H_i$, $i = 1, 2$, we have

$$\big(\, x_1 \otimes x_2\,,\; y_1 \otimes y_2 \,\big)_{H_1\otimes H_2} = (x_1, y_1)_{H_1} \cdot (x_2, y_2)_{H_2}.$$

We denote by $H_1 \hat{\otimes} H_2$ the completion of $H_1 \otimes H_2$ with respect to the norm defined by the above inner product. The Hilbert space $H_1 \hat{\otimes} H_2$ is called the (analytic) *tensor product* of the Hilbert spaces H_1, H_2.

For example if, $(M_j, \mathcal{M}_j, \mu_j)$, $i = 1, 2$, are two measured spaces such that $H_j = L^2(M_j, \mathcal{M}_j \mu_j)$, then there exists a unique isomorphism of Hilbert space

$$H_1 \hat{\otimes} H_2 \to L^2(M_1 \times M_2, \mu_1 \otimes \mu_2),$$

such that

$$f_1 \otimes f_2 \mapsto (f_1 \otimes f_2 : M_1 \times M_2 \to \mathbb{R}),\;\; f_1 \otimes f_2(x_1, x_2) = f_1(x_1) f_2(x_2).$$

For details we refer to [31, Thm. II.10] or the original source [22].

The tensor product of two separable Hilbert spaces H_1, H_2 can also be realized as the space of Hilbert-Schmidt bilinear functionals $u : H_1 \times H_2 \to \mathbb{R}$. This means that for any complete orthonormal bases $(e_m)_{m\geq 1}$ of H_1 and $(f_n)_{n\geq 1}$ of H_2 we have

$$\sum_{m,n\geq 1} |u(e_m, f_n)|^2 < \infty.$$

The tensor product of Hilbert spaces enjoys the usual commutativity and associativity properties

$$H_1 \hat{\otimes} H_2 \cong H_2 \hat{\otimes} H_1,\;\; (H_1 \hat{\otimes} H_2) \hat{\otimes} H_3 \cong H_1 \hat{\otimes} (H_2 \hat{\otimes} H_3).$$

Given a Hilbert space H we denote by $H^{\odot n}$ its *algebraic* n-th symmetric product, i.e., the subspace of $H^{\otimes n}$ consisting of elements fixed by the obvious action of the symmetric group $\mathfrak{S}_n$. The closure of $H^{\odot n}$ in $H^{\hat{\otimes} n}$ is denoted by $H^{\hat{\odot} n}$ and it is called the analytic *n-th symmetric power* of H.

Note that we have a natural projector $\mathbf{Sym} : H^{\otimes n} \to H^{\odot n}$ defined by

$$\mathbf{Sym}[x_1 \otimes \cdots \otimes x_n] := \frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)},\;\; \forall x_1, \dots, x_n \in H.$$

For $x_1, \dots, x_n \in H$ we set

$$x_1 \odot \cdots \odot x_n := \sqrt{n!}\,\mathbf{Sym}[x_1 \otimes \cdots \otimes x_n] = \frac{1}{\sqrt{n!}} \sum_{\sigma\in\mathfrak{S}_n} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}. \tag{5.1}$$

Note that

$$x^{\odot n} = \underbrace{x \odot \cdots \odot x} = \sqrt{n!} x^{\otimes n}$$

and

$$\|x^{\odot n}\|^2 = n! \|x\|^{2n}. \tag{5.2}$$

This is a manifestation of a more general phenomenon.

Lemma 5.1. *If $e_1, \dots, e_n$ is an orthonormal system and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, then*

$$\|e^{\odot\alpha}\|^2 := \|e_1^{\odot\alpha_1} \odot \cdots \odot e_n^{\odot\alpha_n}\|^2 = \alpha! = (\alpha_1!) \cdots (\alpha_n!) = \|H_\alpha\|^2. \tag{5.3}$$

Proof. Consider the multiset

$$E := \{\underbrace{e_1,\dots,e_1}_{\alpha_1},\dots,\underbrace{e_n,\dots,e_n}_{\alpha_n}\},$$

Define a permutation π of the multiset E to be a map

$$\pi : \{1,\dots,|\alpha|\} \to \{1,\dots,n\}$$

such that $\#\pi^{-1}(k) = \alpha_k$, $\forall k = 1,\dots,n$. (The entries of E equal to e_k are moved to occupy the positions in the set $\pi^{-1}(k)$.) We denote by $\mathcal{P}(E)$ the set of permutations of the multiset E. Then

$$\#\mathcal{P}(E) = \binom{|\alpha|}{\alpha_1,\dots,\alpha_n} := \frac{|\alpha|!}{\alpha_1!\cdots\alpha_n!},$$

To each $\pi \in \mathcal{P}(E)$ we associate the element

$$e_\pi := e_{\pi(1)} \otimes \cdots \otimes e_{\pi(|\alpha|)} \in H^{\otimes|\alpha|}.$$

Let us observe that

$$(e_\pi, e_{\pi'}) = \delta_{\pi\pi'},\ \ \forall \pi,\pi' \in \mathcal{P}(E).$$

Then

$$e^{\odot\alpha} = \frac{\prod_{k=1}^n \alpha_k!}{\sqrt{|\alpha|!}} \sum_{\pi\in\mathcal{P}(E)} e_\pi,$$

so that

$$\|e^{\odot\alpha}\|^2 = \left(\frac{\prod_{k=1}^n \alpha_k!}{\sqrt{|\alpha|!}}\right)^2 \binom{|\alpha|}{\alpha_1,\dots,\alpha_n} = (\alpha_1!)\cdots(\alpha_n!).$$

□

The above lemma implies[7] that we have continuous bilinear map

$$\odot : H^{\odot m} \times H^{\odot n} \to H^{\odot(m+n)}$$

defined by

$$\boxed{X \odot Y := \sqrt{\binom{n+m}{m}}\,\mathbf{Sym}\big(X \otimes Y\big),\ \ \forall X \in H^{\odot m},\ \ Y \in H^{\odot n}}.$$

We obtain in this fashion a graded associative and commutative algebra

$$\bigoplus_{n\geq 0} H^{\hat{\odot} n}.$$

Its completion

$$\widehat{\bigoplus_{n\geq 0}} H^{\hat{\odot} n}$$

is called the *Fock space* of H and it is denoted by $\boldsymbol{F}(H)$.

[7]The details are straightforward and not particularly illuminating.

5.2. **The Fock space of a Gaussian Hilbert space.** Suppose that $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a separable Gaussian Hilbert space. We have a linear map

$$\Theta_n : \mathscr{X}^{\odot n} \to \mathscr{X}^{:n:}, \tag{5.4}$$

naturally determined by the correspondences

$$\xi_1 \odot \cdots \odot \xi_n \mapsto \xi_1 \bullet \cdots \bullet \xi_n = :\xi_1 \cdots \xi_n:$$

Corollary 4.5, (2.11) and (5.3) imply that if $\xi_1, \dots, \xi_n$ is an orthonormal system in $\mathscr{X}$, then

$$\sqrt{n!}\| \,\mathbf{Sym}[\xi_1 \otimes \cdots \otimes \xi_n] \,\| = \Big\|\xi_1 \odot \cdots \odot \xi_n\| = \| :\xi_1 \cdots \xi_n: \|. \tag{5.5}$$

We obtain isometries $\Theta_n : \mathscr{X}^{\odot n} \to \mathscr{X}^{:n:}$ and thus an isomorphism of graded Hilbert spaces

$$\Theta : \boldsymbol{F}(\mathscr{X}) \to \widehat{\mathscr{X}}.$$

The associativity (4.7) shows that Ψ is actually an isomorphism of algebras.

If $\mathscr{X}_1, \mathscr{X}_2 \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ are two Gaussian Hilbert spaces, then any bounded linear operator $A : \mathscr{X}_1 \to \mathscr{X}_2$ induces bounded linear operators

$$A^{\odot n} : \mathscr{X}_1^{\odot n} \to \mathscr{X}_2^{\odot n}, \;\; n \in \mathbb{N}_0$$

uniquely determined by the requirements

$$A^{\odot n}(\,\xi_1 \odot \cdots \odot \xi_n\,) = (\,A\xi_1\,) \odot \cdots \odot (\,A\xi_n\,), \;\; \forall \xi_1, \dots, \xi_n \in \mathscr{X}_1.$$

In particular $A^{\odot 0} = \mathbb{1}$. Moreover

$$\|A^{\odot n}\| = \|A\|^n.$$

If $\|A\| \leq 1$, the operators $A^{\odot n}$ combine to a bounded linear operator

$$\boldsymbol{F}(A) : \boldsymbol{F}(\mathscr{X}_1) \to \boldsymbol{F}(\mathscr{X}_2).$$

We deduce that if $\|A\| \leq 1$, then A induces a bounded linear operator $\widehat{A} : \widehat{\mathscr{X}_1} \to \widehat{\mathscr{X}_2}$ uniquely defined by the equalities

$$\widehat{A}(\,\xi_1 \bullet \cdots \bullet \xi_n\,) = (\,\widehat{A}\xi_1\,) \bullet \cdots \bullet (\,\widehat{A}\xi_n\,), \;\; \forall \xi_1, \dots, \xi_n \in \mathscr{X}_1.$$

Note that

$$\|\widehat{A}\| = 1.$$

In particular, a unitary isomorphism $T : \mathscr{X}_1 \to \mathscr{X}_1$ induces a canonical unitary isomorphism

$$\widehat{T} : \widehat{\mathscr{X}_1} \to \widehat{\mathscr{X}_1}$$

which preserves the Wick algebra structure.

Example 5.2. Suppose that $H = L^2(M, \mathcal{M}, \mu)$. Observe that we have a Hilbert space isomorphism

$$H^{\otimes m} \cong L^2(M^m, \mathcal{M}^{\otimes m}, \mu^{\otimes m}), \;\; f_1 \otimes \cdots \otimes f_m(x_1, \dots, x_m) = f_1(x_1) \cdots f_m(x_m).$$

Denote by $\mathcal{M}^{\circledcirc n} \subset \mathcal{M}^{\otimes n}$ the σ-algebra consisting of $\mathcal{S}_n$-invariant $\mathcal{M}^{\otimes n}$-measurable subsets.

Let us observe that a function $f : M^n \to \mathbb{R}$ is $\mathcal{M}^{\odot n}$-measurable if and only if it is $\mathfrak{S}_n$-invariant.[8] We set

$$\mu^{\circledcirc n} := \frac{1}{n!}\mu^{\otimes n}.$$

For any $f_1, \dots, f_n \in L^2(M, \mathcal{M}, \mu)$ define

$$f_1 \circledcirc \cdots \circledcirc f_n : M^n \to \mathbb{R}, \;\; (f_1 \circledcirc \cdots \circledcirc f_n)(x_1, \dots, x_n) = \sum_{\varphi \in \mathfrak{S}_n} \prod_{k=1}^n f_k(x_{\varphi(k)}).$$

[8] For a proof of this fact we refer to [7, Thm. II.4.4].

Clearly, $f_1 \circledcirc \cdots \circledcirc f_n$ is $\mathfrak{S}_n$-invariant. Arguing as in the proof of Lemma 5.1 we deduce that if $\{f_1, \dots, f_n\} \subset L^2(M, \mu)$ is an orthonormal system and $\alpha_1, \dots, \alpha_n \in \mathbb{N}$, then

$$\| \underbrace{f_1 \circledcirc \cdots \circledcirc f_1}_{\alpha_1} \circledcirc \cdots \circledcirc \underbrace{f_n \circledcirc \cdots \circledcirc f_n}_{\alpha_n} \|^2_{L^2(M^{|\alpha|}, \mu^{\circledcirc |\alpha|})} = \prod_{k=1}^n \alpha_k!.$$

We thus obtain a Hilbert space isomorphism

$$\Psi_n : L^2(M, \mathcal{M}, \mu)^{\odot n} \to L^2(M^n, \mathcal{M}^{\circledcirc n}, \mu^{\circledcirc n}) \tag{5.6}$$

uniquely determined by the the requirement

$$\Psi_n(f_1 \odot \cdots \odot f_n) = f_1 \circledcirc \cdots \circledcirc f_n, \;\; \forall f_1, \dots, f_n \in L^2(M, \mathcal{M}, \mu). \tag{5.7}$$

Observe that $L^2(M, \mathcal{M}, \mu)^{\odot n}$ can be identified with the closed subspace

$$L^2(M^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})_{\mathfrak{S}_n} \subset L^2(M^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})$$

consisting of symmetric L^2-functions $F : M^n \to \mathbb{R}$.

The orthogonal projection onto $L^2(M^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})_{\mathfrak{S}_n}$ is the symmetrization operator

$$F \mapsto \mathbf{Sym}[F], \;\; \mathbf{Sym}[F](x_1, \dots, x_n) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} F(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Note that

$$\frac{1}{n!} \int_{M^n} F d\mu^{\otimes n} = \frac{1}{n!} \int_{M^n} \mathbf{Sym}(F) d\mu^{\otimes n} = \int_{M^n} \mathbf{Sym}(F) d\mu^{\circledcirc n}.$$

□

Remark 5.3. Let I denote the unit interval, $\mathcal{B}$ σ-algebra of the Borel subset of I and $\boldsymbol{\lambda}$ the Lebesgue measure on $\mathcal{B}$. For any positive integer n we denote by Δ_n the simplex

$$\Delta_n = \big\{ (x_1, \dots, x_n) \in I^n; \;\; x_1 \leq x_2 \leq \cdots \leq x_n \big\}.$$

Observe that the space $L^2\big(\Delta_n, \mathcal{B}^{\otimes n}, \boldsymbol{\lambda}^{\otimes n} \big)$ is *isometric* to the subspace

$$L^2\big(I^n, \mathcal{B}^{\otimes n}, \boldsymbol{\lambda}^{\circledcirc n}\big)_{\mathfrak{S}_n} \;=\; L^2\big(I^n, \mathcal{B}^{\otimes n}, \frac{1}{n!}\boldsymbol{\lambda}^{\otimes n}\big)_{\mathfrak{S}_n}.$$

□

Example 5.4. Consider the one-dimensional Gaussian Hilbert space $\mathscr{X}$ spanned by a standard normal random variable ξ. In this case

$$\widehat{\mathscr{X}} = L^2(\mathbb{R}, \boldsymbol{\gamma}_1).$$

Any linear operator $\mathscr{X} \to \mathscr{X}$ has the form $r\mathbb{1}$, and it is a contraction provided $|r| \leq 1$.

Any $f \in L^2(\mathbb{R}, \boldsymbol{\gamma}_1)$ has the form

$$f(\xi) = \sum_{n \geq 0} f_n H_n(\xi), \;\; f_n = \frac{1}{n!} \mathbb{E}\big[f(\xi) H_n(\xi) \big] = \frac{1}{n!} \int_{\mathbb{R}} f(x) H_n(x) \boldsymbol{\gamma}_1(dx).$$

Since : $\xi^n := H_n(\xi)$ we deduce that $\widehat{r\mathbb{1}} H_n(\xi) = r^n H_n(\xi)$ and

$$\widehat{r\mathbb{1}} f = \widehat{r\mathbb{1}} \left(\sum_{n \geq 0} f_n H_n(\xi) \right) = \sum_{n \geq 0} f_n r^n H_n(\xi).$$

The operator $\widehat{r\mathbb{1}} : L^2(\mathbb{R}, \boldsymbol{\gamma}_1) \to L^2(\mathbb{R}, \boldsymbol{\gamma}_1)$, $|r| \leq 1$ is called the *Mehler transform*. It is an integral operator with kernel

$$\mathcal{M}_r(x,y) = \sum_{n\geq 0} H_n(x)H_n(y)\frac{r^n}{n!} \in L^2(\mathbb{R}^2, \boldsymbol{\gamma}_1^{\otimes 2}).$$

The above series converges uniformly for (x, y, r) on the compacts of $\mathbb{R}^2 \times (-1, 1)$. We denote by $\mathcal{M}_r$ the integral operator $\widehat{r\mathbb{1}}$. Consider the function

$$g_\lambda(x) = \sum_{n\geq 0} H_n(x)\frac{\lambda^n}{n!} = e^{\lambda x - \frac{\lambda^2}{2}}.$$

Observe that for $|r| \leq 1$ we have $\mathcal{M}_r g_\lambda = g_{r\lambda}$. This equality determines $\mathcal{M}_r(x, y)$ uniquely. Consider the function

$$M_r(x,y) = \frac{1}{\sqrt{1-r^2}} \exp\left(-\frac{(rx)^2 - 2rxy + (ry)^2}{2(1-r^2)}\right).$$

A direct but tedious computation shows that[9]

$$\int_{\mathbb{R}} M_r(x,y)g_\lambda(y)d\gamma_1(dy) = g_{r\lambda}(x)$$

so that

$$\sum_{n\geq 0} H_n(x)H_n(y)\frac{r^n}{n!} = \frac{1}{\sqrt{1-r^2}} \exp\left(-\frac{(rx)^2 - 2rxy + (ry)^2}{2(1-r^2)}\right), \quad \forall |r| < 1.$$

The function $\mathcal{M}_r(x,y) = M_r(x,y)$ is called the *Mehler kernel*.

The family of operators $T_t := \widehat{e^{-t}\mathbb{1}}$, $t \geq 0$ is called the *Ornstein-Uhlenbeck semigroup*. □

6. Gaussian noise and the Wiener-Ito integral

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space space.

Definition 6.1. A Gaussian *Wiener-Ito integral* is defined by a measured space $(M, \mathcal{M}, \mu)$ and a Hilbert space isometry

$$\boldsymbol{J} : L^2(M, \mathcal{M}, \mu) \to L^2(\Omega, \mathcal{F}, \mathbb{P}),$$

whose image $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a (real) Gaussian Hilbert space.

A *complex* Gaussian Wiener-Ito integral is defined by a measured space $(M, \mathcal{M}, \mu)$ and a complex Hilbert space isometry

$$\boldsymbol{J} : L^2_{\mathbb{C}}(M, \mathcal{M}, \mu) \to L^2_{\mathbb{C}}(\Omega, \mathcal{F}, \mathbb{P}),$$

whose image is a complex Gaussian Hilbert space. □

Remark 6.2. We see that a Wiener-Ito integral is a special isonormal Gaussian process $(H, \mathscr{X}, \boldsymbol{I})$ (see Definition 2.4) where $H = L^2(M, \mathcal{M}, \mu)$. □

Example 6.3. Suppose that $(X_t)_{t\in\mathbb{R}^n}$ is a centered stationary Gaussian random field defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The stationarity implies that there exists a function $C : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathbb{E}\big(X_t X_s\big) = C(t-s), \quad \forall t, s \in \mathbb{R}. \tag{6.1}$$

[9]For a more conceptual approach we refer to [12, Example 4.18], [20, V.1.5] or Lemma 9.3.

Assume that $t \mapsto C(t)$ is continuous. Then Bochner's theorem [2, 11] implies that there exists a finite positive Borel measure μ on $\mathbb{R}$ such that

$$C(t) = \int_{\mathbb{R}^n} e^{\boldsymbol{i}(t,\xi)}\mu(d\xi). \tag{6.2}$$

The measure σ is the so called *spectral measure* of the stationary field. The exponentials

$$e_t(\xi) = e^{\boldsymbol{i}(t,\xi)},\;\; t \in \mathbb{R}^n, \tag{6.3}$$

span a subspace dense in $L^2(\mathbb{R}^n, \mu(d\xi))$. The equalities (6.1,6.2,6.3) show that the map

$$\boldsymbol{J} : \operatorname{span}_{\mathbb{C}}\big\{ e_t;\;\; t \in \mathbb{R}^n \big\} \to \operatorname{span}_{\mathbb{C}}\big\{ X_t;\;\; t \in \mathbb{R}^n \big\}, e_t \mapsto X_t,$$

is an isometry. This induces an isometry

$$\boldsymbol{J} : L^2_{\mathbb{C}}(\mathbb{R}^n, \mu(d\xi)\,) \to \mathscr{X}_{\mathbb{C}},$$

where $\mathscr{X}_{\mathbb{C}}$ is the complex Gaussian Hilbert space generated by the random field $(X_t)_{t\in\mathbb{R}^n}$. This proves that Gaussian stationary fields on $\mathbb{R}^n$ with continuous covariance kernel come equipped with a canonical complex Wiener-Ito integral. □

Example 6.4 (Malliavin)**.** Denote by $\boldsymbol{\lambda}$ the Lebesgue measure on $I = [0, 1]$ by $\mathcal{B}$ the σ-algebra of Borel subsets of I. Let us explain how to construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a Gaussian $\boldsymbol{\lambda}$-noise

$$W : \mathcal{B} \to L^2(\Omega, \mathcal{F}, \mathbb{P}).$$

For $n = 0, 1, 2, \dots$ we denote by $\mathcal{B}_n \subset \mathcal{B}$ the σ-algebra generated by the closed intervals

$$I_{n,k} = \big[\, (k-1)/2^n, k/2^n\,\big),\;\; k = 1, \dots, 2^n.$$

Note that we have inclusions

$$q_n : L^2(I, \mathcal{B}_n, \lambda) \hookrightarrow L^2(I, \mathcal{B}_{n+1}, \boldsymbol{\lambda}),\;\; n = 0, 1, 2, \dots$$

and orthogonal projections

$$\mathbb{E}\big[-|\mathcal{B}_n\big] : L^2(I, \mathcal{B}_{n+1}, \boldsymbol{\lambda}) \to L^2(I, \mathcal{B}_n, \boldsymbol{\lambda}).$$

From Doob's martingale convergence theorem we deduce that

$$\bigcup_{n\geq 0} L^2(I, \mathcal{B}_n, \boldsymbol{\lambda})$$

is dense in $L^2(I, \mathcal{B}, \boldsymbol{\lambda})$.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ containing a family of Gaussian random variables

$$W(I_{n,k}) \in L^2(\Omega, \mathcal{F}, \mathbb{P}),\;\; n = 0, 1, \dots,\;\; k = 1, \dots, 2^n,$$

such that, for all $n = 0, 1, 2, \dots$, we have

$$\mathbb{E}\big[\, W(I_{n,k})W(I_{n,j})\,\big] = \boldsymbol{\lambda}(I_{n,k} \cap I_{n,j}) = \big(\,\boldsymbol{1}_{I_{n,j}}, \boldsymbol{1}_{I_{n,k}}\,\big)_{L^2(\boldsymbol{\lambda})},\;\; \forall 1 \leq k, j \leq 2^n, \tag{6.4a}$$

$$W(I_{n,k}) = W(I_{n+1,2k-1}) + W(I_{n+1,2k}),\;\; \forall 1 \leq k \leq 2^n. \tag{6.4b}$$

The collection

$$\boldsymbol{1}_{I_{n,k}},\;\; k = 1, \dots, 2^n$$

is an orthogonal basis of $L^2(I, \mathcal{B}_n, \boldsymbol{\lambda})$ and the equality (6.4a) shows that we have an isometry

$$W_n : L^2(\boldsymbol{I}, \mathcal{B}_n, \boldsymbol{\lambda}) \hookrightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}).$$

We denote by $\mathscr{X}_n$ the image of this isometry. The equality (6.4b) shows that we have a commutative diagram

$$\begin{tikzcd} L^2(\boldsymbol{I}, \mathcal{B}_n, \boldsymbol{\lambda}) \arrow[r, hook, "q_n"] \arrow[dr, "W_{n+1}"'] & L^2(\boldsymbol{I}, \mathcal{B}_{n+1}, \boldsymbol{\lambda}) \arrow[d, "W_{n+1}"] \\ & L^2(\Omega, \mathcal{F}, \mathbb{P}) \end{tikzcd}$$

For any function $f \in L^2(I, \mathcal{B})$ we set

$$f_n = \mathbb{E}[f|\mathcal{B}_n] = \sum_{1 \leq k \leq 2^n} f(I_{n,k}) \mathbf{1}_{I_{n,k}} \in L^2(I, \mathcal{B}_n, \boldsymbol{\lambda}), \;\; f(I_{n,k}) := \int_{I_{n,k}} f(x) \boldsymbol{\lambda}(dx),$$

For $m < n$ we have, $f_n - f_m \in L^2(\boldsymbol{I}, \mathcal{B}_n, \boldsymbol{\lambda})$ and

$$\|W_n(f_n) - W_m(f_m)\|_{L^2} = \|W_n(f_n) - W_n(f_m)\|_{L^2} = \|f_n - f_m\|_{L^2}.$$

Thus the sequence $W_n(f_n)$ converges in L^2 to a Gaussian random variable $W(f)$. The resulting map

$$W : L^2(I, \mathcal{B}, \boldsymbol{\lambda}) \to L^2(\Omega, \mathcal{F}, \mathbb{P}), \;\; f \mapsto W(f),$$

is a Gaussian Wiener integral. The stochastic process

$$B(t) = W(\boldsymbol{I}_{[0,t]}), \;\; t \in [0,1]$$

is, up to a modification, the Brownian motion started at the origin. □

The notion of Wiener-Ito integral is intimately related to the notion of *Gaussian noise*, or *Gaussian stochastic measure*.

Definition 6.5. Suppose that $(M, \mathcal{M}, \mu)$ is a finite measured space. A (complex) *Gaussian μ-noise* consists of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and map

$$W : \mathcal{M} \to L^2(\Omega, \mathcal{F}, \mathbb{P}),$$

with the following propertite.

- $\forall A \in \mathcal{M}$, $W(A)$ is a centered (complex) Gaussian random variable.
- $\forall A, B \in \mathcal{M}$
$$\mathbb{E}\big(W(A) \cdot \overline{W(B)} \big) = \mu(A \cap B).$$
- $\forall A, B \in \mathcal{M}$, $A \cap B = \emptyset$,
$$W(A \cup B) = W(A) + W(B), \;\; \text{a.s..}$$

□

Observe that if $\boldsymbol{J} : L^2(M, \mathcal{M}, \mu) \to \mathscr{X}$ is a real Wiener-Ito integral, then the correspondence

$$\mathcal{M} \ni A \mapsto W(A) = W_{\boldsymbol{J}}(A) := \boldsymbol{J}(\mathbf{1}_A) \in \mathscr{X}$$

is a Gaussian μ-noise. Similarly, if $\boldsymbol{J} : L^2_{\mathbb{C}}(M, \mathcal{M}, \mu) \to \mathscr{X}_{\mathbb{C}}$ is a complex Wiener-Ito integral, then the correspondence

$$\mathcal{M} \ni A \mapsto W_{\boldsymbol{J}}(A) := \boldsymbol{J}(\mathbf{1}_A) \in \mathscr{X}$$

is a complex Gaussian μ-noise.

Conversely, suppose that $\big(W(A) \big)_{A \in \mathcal{M}}$ is a Gaussian μ-noise, $W(A) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then for any $A \in \mathcal{M}$ we set

$$\boldsymbol{J}_W(\mathbf{1}_A) := W(A).$$

We define a *simple function* on M to be a measurable function $f : M \to \mathbb{R}$ with *finite* range. Denote by $\mathcal{S}(M, \mathcal{M})$ the vector space of simple functions. For $f \in \mathcal{S}(M, \mathcal{M})$ We define

$$\boldsymbol{J}_W : \mathcal{S}(M, \mu) \to L^2(\Omega, \mathcal{F}, \mathbb{P}), \quad \boldsymbol{J}_W(f) := \sum_{t \in \mathbb{R}} tW\big(f^{-1}(t)\big).$$

From the properties of a Gaussian μ-noise we deduce that the map $\boldsymbol{J}_\mu$ is linear and

$$\|\boldsymbol{J}_W(f)\|_{L^2} = \|f\|_{L^2}.$$

Thus, $\boldsymbol{J}_W$ extends to an isometry

$$\boldsymbol{J}_W : L^2(M, \mathcal{M}, \mu) \to L^2(\Omega, \mathcal{F}, \mathbb{P})$$

whose image is a Gaussian Hilbert space. Traditionally one uses the notation

$$\boldsymbol{J}_W(f) := \int_M f(x) Z_\mu(dx), \quad \forall f \in L^2(M, \mathcal{M}, \mu).$$

Example 6.6 (Brownian motion)**.** Consider the Brownian motion $(B_t)_{t \geq 0}$ on $[0, \infty)$. It is known that

$$\boldsymbol{cov}(B_t, B_s) = \min(t, s).$$

The "differential" $W(dt) = dB_t$ is a $\boldsymbol{\lambda}$-white noise on $[0, \infty)$, where $\boldsymbol{\lambda}$ denotes the Lebesgue measure. The usual Ito integral [13, 17] associates to each function $L^2(\mathbb{R}_{\geq 0}, \boldsymbol{\lambda})$ a Gaussian random variable

$$\int_0^\infty f(s) dB_s$$

with mean zero and variance $\|f\|_{L^2}^2$. We denote by $\mathscr{X}$ the Gaussian Hilbert space generated by $(B_t)_{t \geq 0}$. The correspondence

$$L^2(\mathbb{R}_{\geq 0}, \boldsymbol{\lambda}) \ni f \mapsto \boldsymbol{J}(f) := \int_0^\infty f(s) dB_s$$

is an isometry onto $\mathscr{X}$, and

$$B_t = \boldsymbol{J}(\boldsymbol{1}_{[0,t]}) = \int_0^t dB_s. \qquad \square$$

Remark 6.7. We see that there exists a bijective correspondence between Wiener-Ito integrals and Gaussian white noise measures on $(M, \mathcal{M}, \mu)$. If the measured space is sufficiently regular, say M is a smooth manifold, $\mathcal{M} =$ the σ-algebra of Borel sets and μ the measure induced by a Riemann metric on M, then on can describe Gaussian μ-noises as certain Gaussian measures on the space $C^{-\infty}(M)$, the space of generalized functions on M. For details we refer to the beautiful presentation in [11, Sec. III.4]. $\square$

Suppose that $\boldsymbol{J} : L^2(M, \mathcal{M}, \mu) \to \mathscr{X}$ is a Gaussian Wiener-Ito integral with associated Gaussian μ-noise W. The isometries (5.4) yield isometries

$$L^2(M, \mathcal{M}, \mu)^{\hat{\odot} n} \to \mathscr{X}^{:n:}, \quad \forall n \in \mathbb{N}_0.$$

Using the isometries (5.6) in Example 5.2 we obtain *isometries*

$$\boldsymbol{J}_n : L^2(M^n, \mathcal{M}^{\odot n}, \mu^{\odot n}) \xrightarrow{\Psi_n^{-1}} L^2(M, \mu)^{\hat{\odot} n} \xrightarrow{\Theta_n} \mathscr{X}^{:n:}.$$

For example if

$$F = \mathbf{Sym}[f_1(x_1) \cdots f_n(x_n)] = \frac{1}{n!} f_1 \circledcirc \cdots \circledcirc f_n(x_1, \dots, x_n) \in L^2(M^n, \mathcal{M}^{\odot n}, \mu^{\odot n}),$$

then

$$\Psi_n^{-1}(F) = \frac{1}{n!} f_1 \odot \cdots \odot f_n, \quad \boldsymbol{J}_n[F] = \frac{1}{n!} \boldsymbol{J}(f_1) \bullet \cdots \bullet \boldsymbol{J}(f_n.)$$

For $F \in L^2(M^n, \mathcal{M}^{\otimes n}, \mu^{\otimes n})$ we define its *multiple Wiener-Ito integral* to be the *random variable*

$$\boldsymbol{I}_n[F] := n! \boldsymbol{J}_n(\mathbf{Sym}[F]).$$

Often one uses the integral notation

$$\int_{M^n} F dW^n = \int_{M^n} F(x_1, \dots, x_n) W(dx_1) \cdots W(dx_n) := \boldsymbol{I}_n[F], \quad \forall F : M^n \to \mathbb{R}.$$

In particular, if $F \in L^2(M^n, \mathcal{M}^{\otimes n})$ is *symmetric*, then

$$\boxed{\boldsymbol{I}_n[F] = n! \boldsymbol{J}_n[F] \Longleftrightarrow \boldsymbol{J}_n(F) = \frac{1}{n!} \int_{M^n} F dW^n.}$$

Note that if $F(x_1, \dots, x_n) = f_1(x_1) \cdots f_n(x_n)$, $f_1, \dots, f_n \in L^2(M, \mathcal{M}, \mu)$, then we obtain the important equality

$$\boxed{\boldsymbol{I}_n[f_1(x_1) \cdots f_n(x_n)] = \boldsymbol{I}_n\big[\,\mathbf{Sym}[\,f_1(x_1) \cdots f_n(x_n)]\,\big] = :\boldsymbol{J}(f_1) \cdots \boldsymbol{J}(f_n):}. \tag{6.5}$$

This equality uniquely determines the multiple Ito integral.

Note that since $\boldsymbol{J}_n$ is an isometry we deduce that for any $F \in L^2(M^n, \mathcal{M}^n, \mu^{\otimes n})$ have

$$\mathbb{E}\Big[\,|\,\boldsymbol{I}_n[F]\,|^2\,\Big] = \mathbb{E}\Big[\,\big|\,\boldsymbol{I}_n\big[\,\mathbf{Sym}[F]\,\big]\,\big|^2\,\Big] = \mathbb{E}\Big[\,\big|\,n! \boldsymbol{J}_n\big[\,\mathbf{Sym}[F]\,\big]\,\big|^2\,\Big]$$

$$= \|n!\,\mathbf{Sym}[F]\|^2_{L^2(M^n, \mu^{\odot n})} = n! \|\,\mathbf{Sym}[F]\|^2_{L^2(M^n, \mu^{\otimes n})} \le n! \|F\|^2_{L^2(M^n, \mu^{\otimes n})}.$$

We observe that any $X \in \widehat{\mathscr{X}}$ has a unique orthogonal decomposition

$$X = \sum_{n \ge 0} \boldsymbol{I}_n[F_n] = \sum_{n \ge 0} \int_{M^n} F_n dW^n,$$

where $F_n : M^n \to \mathbb{R}$ are symmetric L^2-functions. Moreover

$$\mathbb{E}[X^2] = \sum_{n \ge 0} n! \|F_n\|^2_{L^2(M^n, \mu^{\otimes n})} = \sum_{n \ge 0} (n!)^2 \|F_n\|^2_{L^2(M^n, \mu^{\odot n})}.$$

Remark 6.8. There are many normalization conditions involving the multiple Ito integrals and there is danger of confusion. In [28], the Hermite polynomials have a different normalization than the one we use in these notes which is the more commonly used. If $F : M^n \to \mathbb{R}$ symmetric function, then $\boldsymbol{I}_n(F)$, as defined in [28] coincides with the multiple integral $\boldsymbol{I}_n[F]$ defined above.

The operator $\boldsymbol{J}_n$ that we have described in this section coincides with the operator $\boldsymbol{I}_G$ in [19]. The correspondence $L^2(M, \mu) \ni f \mapsto \boldsymbol{I}_1(f) = \boldsymbol{J}(f) \in \mathscr{X}$ is the white-noise map which explains why, in many books it is denoted denoted by $f \mapsto W(f)$.

When the measure μ has no atoms, the multiple Ito integral can be given a constructive description that justifies the terminology *integral*. Fo details refer to [21, §VI.2] or [28, §1.1.2]. □

Suppose that we are given a (real) Gaussian Wiener integral $\boldsymbol{J} : L^2(M, \mathcal{M}, \mu) \to \mathscr{X}$ with associated μ-noise W. If $F : M^2 \to \mathbb{R}$ is an integrable function we define the *contraction*

$$CF := \int_M F(x, x) \mu(dx).$$

More generally, if $F : M^n \to \mathbb{R}$ and $1 \leq i < j \leq n$ we define the *contraction* $C_{ij}F : M^{n-2} \to \mathbb{R}$ to be

$$C_{ij}F := \int_M F(x_1, \dots, x_n)_{x_i = x_j = x}\, \mu(dx)$$

Given a Feynman diagram Γ with vertices labelled by $\{1, \dots, n\}$ we define

$$C_\Gamma := \prod_{\boldsymbol{e} \in \mathcal{E}(\Gamma)} C_{\boldsymbol{e}},$$

where for any edge $\boldsymbol{e} = (i, j)$ of Γ we set $C_{\boldsymbol{e}} = C_{ij}$.

Lemma 6.9. *Suppose we are given functions $F_i \in L^2(M^{n_i})$, $i = 1, \dots, k$. We define*

$$F = F_1 \otimes \cdots \otimes F_k : M^n \to \mathbb{R}, \;\; n = \sum_{i=1}^k n_i,$$

$$F_1 \otimes \cdots \otimes F_k(x_{11}, \dots, x_{1n_1}; \dots ; x_{k1}, \dots, x_{kn_k}) := \prod_{j=1}^k F_j(x_{j1}, \dots, x_{jn_k}).$$

For any Feynman diagram Γ compatible with the array

$$(i, j), \;\; 1 \leq i \leq k, \;\; 1 \leq j \leq n_i,$$

we have

$$\| C_\Gamma(F_1 \otimes \cdots \otimes F_k)\|_{L^2(M^{n-2r(\Gamma)})} \leq \prod_{j=1}^k \|F_j\|_{L^2}.$$

Proof. We use induction on k. The case $k = 1$ is trivial since Γ has no wedge and thus $C_\Gamma(F) = F$.

For $k = 2$, we can assume, after relabeling the variables, that the $r(\Gamma)$ edges of Γ connectd the vertices $(1, j)$ and $(2, j)$, $j = 1, \dots, r$. Then for

$$x' \in M^{n_1 - r}, \;\; x'' \in M^{n_2 - r}, \;\; y \in M^r$$

we have

$$C_\Gamma F(x', x'') = \int_{M^r} F_1(y, x')F_2(y, x'')\mu^r(dy)$$

and thus, by Cauchy-Schwarz

$$|C_\Gamma F(x', x'')|^2 \leq \left(\int_{M^r} F_1(y, x')^2 \mu^r(dy)\right)\left(\int_{M^r} F_2(y, x'')^2\right)\mu^r(dy)$$

Integrating the remaining variables (x', x'') we deduce

$$\|C_\Gamma(F)\|^2_{L^2} \leq \|F_1\|^2_{L^2} \cdot \|F_2\|^2_{L^2}.$$

This disposes of the case $k = 2$.

For $k > 2$ we set

$$F_2' = f_2 \otimes \cdots \otimes F_k.$$

Denote by Γ_1 the subdiagram of Γ consisting of the edges that have one vertex on the first row, $(1, j)$, $1 \leq j \leq n_1$, and denote by Γ_2 the subdiagram of Γ determined by the edges of Γ that connect points on rows different from the first row.

We than have

$$C_\Gamma(F) = C_{\Gamma_1}\big(\, F_1 \otimes C_{\Gamma_2}(F_2')\,\big).$$

Thus, using the inequality established for $k = 2$ and the induction assumption we reach the desired conclusion. $\square$

Suppose we are given functions $F_i \in L^2(M^{n_i})$, $i = 1, \dots, k$. We set

$$Y_i = \boldsymbol{I}_{n_i}[F_i] = \boldsymbol{I}_{n_i}\big[\,\mathbf{Sym}[F_i]\,\big] \in \mathscr{X}^{:n_i:}, \;\; 1 \le i \le k.$$

From the Diagram Formula (Theorem 4.7) we deduce rather easily (see [12, Thm. 7.33] for details) the following important result.

Theorem 6.10. *We set $n := n_1 + \cdots + n_k$. Then*

$$Y_1 \cdots Y_k = \sum_{\Gamma \in D} \boldsymbol{I}_{n-2r(\Gamma)}[C_\Gamma(F_1 \otimes \cdots \otimes F_k)\,], \tag{6.6}$$

where D denotes the collection of Feynman diagrams compatible with the array

$$(i,j),\;\; 1 \le i \le k,\;\; 1 \le j \le n_i.$$

In particular

$$\mathbb{E}[Y_1 \cdots Y_k] = \sum_{\Gamma \in D^c} C_\Gamma(F_1 \otimes \cdots \otimes F_k), \tag{6.7}$$

where $D^c \subset D$ denotes the collection of complete *Feynman diagrams compatible with the above array.*

Proof. Lemma 6.9 shows that all the contractions and stochastic integrals are well defined. From the definition of the multiple Wiener-Ito integrals we deduce that the right-hand side of (6.6) defines a continuous multilinear map

$$\prod_{i=1}^k L^2(M^{n_i}, \mu^{n_i}) \to \widehat{\mathscr{X}}.$$

Thus it suffice to verify the equality (6.6) in the special case when each f_i is a monomial

$$f_i(x_1, \dots, x_{m_i}) = f_{i1}(x_1) \cdots f_{im_i}(x_{m_i}).$$

This special case follows immediately from the Diagram Formula. □

Remark 6.11. Theorem 6.10 corresponds to [19, Thm. 5.3] where it is referred to as the *Diagram Formula*. □

Corollary 6.12. *Suppose that $f \in L^2\big(\,M^n, \mu^{\otimes n}\,\big)$, $h \in L^2(M, \mu)$. Define*

$$f \times_k h\big(\,x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n\,\big) := \int_M f(x_1, \dots, x_n) h(x_k) \mu(dx_k).$$

Then

$$\boldsymbol{I}_n[f]\boldsymbol{I}_1[h] = \boldsymbol{I}_{n+1}[f \otimes h] + \sum_{k=1}^n \boldsymbol{I}_{n-1}[f \times_k h]. \qquad \square$$

PART 2. MALLIAVIN CALCULUS

7. THE MALLIAVIN GRADIENT AND GAUSSIAN SOBOLEV SPACES

Suppose H is a real Hilbert space, and $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \boldsymbol{P})$ is a real Gaussian Hilbert space.

- We denote by $L^0_{\mathscr{X}}(\Omega)$ the space of $\widehat{\mathcal{F}}$-measurable functions $f : \Omega \to \mathbb{R}$.
- For $p \in [1, \infty]$ we denote by $L^p_{\mathscr{X}}(\Omega)$ the subspace of $L^0_{\mathscr{X}}(\Omega)$ consisting of p-integrable functions equipped with the usual L^p-norm. Note that $\widehat{\mathscr{X}} = L^2_{\mathscr{X}}(\Omega)$.

- If H is a Hilbert space, then we denote by $L^0_{\mathscr{X}}(\Omega, H)$ the space of $\widehat{\mathcal{F}}$-measurable maps $f : \Omega \to H$, and by $L^p_{\mathscr{X}}(\Omega, H)$ the subspace $L^0_{\mathscr{X}}(\Omega, H)$ consisting of maps $f : \Omega \to H$ such that $\|f\| \in L^p_{\mathscr{X}}(\Omega)$. The norm in this space is
$$\mathbb{E}\big[\,\|f\|^p_H\,\big]^{\frac{1}{p}}.$$
The space $L^0_{\mathscr{X}}(\Omega, H)$ is equipped with a bilinear map
$$(-,-)_H : L^0_{\mathscr{X}}(\Omega, H) \times L^0_{\mathscr{X}}(\Omega, H) \to L^0_{\mathscr{X}}(\Omega),$$
$$(f, g)_H(\omega) := \big(\, f(\omega), g(\omega)\,\big)_H, \;\; \forall f, g \in L^0_{\mathscr{X}}(\Omega, H).$$
- A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *admissible* if it is smooth and its derivatives, of any order, have at most polynomial growth.

We will construct various Banach subspaces $L^0_{\mathscr{X}}(\Omega)$. These depend only on $\mathscr{X}$.

Definition 7.1. (a) We denote by $\mathcal{S}(\mathscr{X}) \subset L^0_{\mathscr{X}}(\Omega)$ the set of random variables of the form $f(X_1, \dots, X_m)$ where $f : \mathbb{R}^m \to \mathbb{R}$ is an admissible function.

(b) We denote by $\mathcal{P}(\mathscr{X}) \subset \mathcal{S}(\mathscr{X})$ the set of random variables of the form $P(X_1, \dots, X_m)$, where $P : \mathbb{R}^m \to \mathbb{R}$ is a polynomial in m variables with real coefficients. □

Given $f(X_1, \dots, X_n) \in \mathcal{S}(\mathscr{X})$, define $Df(X_1, \dots, X_m) \in L^0_{\mathscr{X}}(\Omega, \mathscr{X})$
$$Df(X_1, \dots, X_m)(\omega) = \sum_j \frac{\partial f}{\partial x_j}\big(\, X_1(\omega), \dots, X_m(\omega)\,\big) X_j(-). \tag{7.1}$$

Remark 7.2. (a) For $X \in \mathscr{X}$ we set
$$D_X f(X_1, \dots, X_m)(\omega) := \sum_j \frac{\partial f}{\partial x_j}\big(\, X_1(\omega), \dots, X_m(\omega)\,\big)(X_j, X), \tag{7.2}$$
where $(-,-)$ denotes the inner product in $\mathscr{X}$, $(X, Y) = \mathbb{E}[\, XY\,]$. We have the almost sure equality
$$D_X f := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big(\, f\big(\, X_1(\omega) + \varepsilon(X_1, X), \dots, X_m(\omega) + \varepsilon(X_m, X)\,\big) - f\big(\, X_1(\omega), \dots, X_m(\omega)\,\big)\,\Big).$$
(b) From the definition (7.1) it is not clear whether the equality
$$f(X_1, \dots, X_m) = g(Y_1, \dots, Y_n) \in \mathcal{S}(\mathscr{X}),$$
where f and g are admissible, implies that
$$Df(X_1, \dots, X_m) = Dg(Y_1, \dots, Y_n) \in L^0_{\mathscr{X}}(\Omega, \mathscr{X}).$$
This is indeed the case, but the proof is more involved and requires an alternate definition of D based on the *Cameron-Martin shift*. For a proof we refer to [12, Thm.14.1& Def.15.26]. The resulting operator
$$f(X_1, \dots, X_m) \mapsto Df(X_1, \dots, X_m)$$
is called the *Malliavin gradient* or *derivative*. □

Example 7.3. Let $X \in \mathscr{X}$. Then DX is the constant map $\Omega \to \mathscr{X}$, $\omega \mapsto X$. For this reason we will rewrite (7.1) in the form
$$\boxed{Df(X_1, \dots, X_m)(\omega) = \sum_j \frac{\partial f}{\partial x_j}\big(\, X_1(\omega), \dots, X_m(\omega)\,\big) DX_j}. \tag{7.3}$$

This notation better conveys the nature of the two factors $\frac{\partial f}{\partial x_j}\big(X_1(\omega),\dots,X_m(\omega)\big)$ and DX_j. The first is a *scalar*, while the second is an element of $\mathscr{X}$. Note also that

$$D_X F = (DF, DX)_{\mathscr{X}} \in L^0_{\mathscr{X}}(\Omega). \qquad \square$$

For a positive integer p and $f(X_1,\dots,X_m)$ as above we define

$$D^p f(X_1,\dots,X_m) \in L^0_{\mathscr{X}}\big(\Omega, \mathscr{X}^{\hat{\odot} p}\big)$$

by setting

$$D^p f(X_1,\dots,X_m)(\omega) = \sum_{i_1,\dots,i_p=1}^m \frac{\partial^p f}{\partial x_{i_1}\cdots\partial x_{i_p}}\big(X_1(\omega),\dots,X_m(\omega)\big) DX_{i_1}\otimes\cdots\otimes DX_{i_p}.$$

Remark 7.4. Arguing as in Lemma 5.1 we deduce that

$$\begin{aligned}
&\sum_{i_1,\dots,i_p=1}^m \frac{\partial^p f}{\partial x_{i_1}\cdots\partial x_{i_p}}\big(X_1(\omega),\dots,X_m(\omega)\big) DX_{i_1}\otimes\cdots\otimes DX_{i_p} \\
&= \sqrt{p!}\sum_{\alpha\in\mathbb{N}_0^m;|\alpha|=p} \frac{1}{\alpha!}\partial_x^\alpha f\big(X_1(\omega),\dots,X_m(\omega)\big)(DX_1)^{\odot\alpha_1}\odot\cdots\odot(DX_m)^{\odot\alpha_m} \\
&= \sum_{\alpha\in\mathbb{N}_0^m;|\alpha|=p} \frac{1}{\alpha!}\partial_x^\alpha f\big(X_1(\omega),\dots,X_m(\omega)\big)\,\mathbf{Sym}\big[\,(DX_1)^{\otimes\alpha_1}\otimes\cdots\otimes(DX_m)^{\otimes\alpha_m}\,\big]. \qquad \square
\end{aligned}$$

Observe that the class $\mathcal{S}$ contains the algebra generated by the polynomials $H_n(X)$, $X\in\mathscr{X}$. Arguing as in the proof of Proposition 3.1 we deduce that $\mathcal{S}$ is dense in $L^q_{\mathscr{X}}(\Omega,\mathbb{P})$, $\forall q\in(1,\infty)$.

Proposition 7.5. *Let $p\in\mathbb{N}$ and $q\in(1,\infty)$. Then the operator*

$$D^p : \mathcal{S}(\mathscr{X}) \subset L^q_{\mathscr{X}}(\Omega) \to L^q_{\mathscr{X}}\big(\Omega, \mathscr{X}^{\hat{\odot} p}\big)$$

is closable.

Proof. We follow closely the proof of [26, Prop.2.3.4]. We consider only the case $p=1$.

Let $F,G\in\mathcal{S}(\mathscr{X})$ and $X\in\mathscr{X}$ such that $\|X\|_{L^2}=1$. Note that $FG\in\mathcal{S}(X)$. We can assume that

$$FG = f(X_1,\dots,X_n),$$

where $\{X_1,\dots,X_n\}\subset\mathscr{X}$ is an orthonormal system, $X_1=X$, and f is admissible. Then

$$\mathbb{E}\big[\,(D(FG),X)_{\mathscr{X}}\,\big] = (2\pi)^{-n/2}\int_{\mathbb{R}^n}\frac{\partial f}{\partial x_1}(x_1,\dots,x_n)e^{-\frac{x_1^2+\cdots+x_n^2}{2}}dx$$

(integrate by parts along the x_1-direction)

$$= (2\pi)^{-n/2}\int_{\mathbb{R}^n} x_1 f(x_1,\dots,x_n)e^{-\frac{x_1^2+\cdots+x_n^2}{2}}dx = \mathbb{E}\big[\,XFG\,\big].$$

Clearly $D(FG)=G(DF)+F(DG)$. We deduce the following Gaussian integration by parts formula

$$\boxed{\mathbb{E}\big[\,G(DF,X)_{\mathscr{X}}\,\big] = -\mathbb{E}\big[\,F(DG,X)_{\mathscr{X}}\,\big] + \mathbb{E}\big[\,XFG\,\big]}. \tag{7.4}$$

Using the notation (7.2) we can rewrite the above equality in the more suggestive form

$$\boxed{\mathbb{E}\big[\,GD_XF\,\big] = \mathbb{E}\big[\,F(-D_X+X)G\,\big]}. \tag{7.5}$$

The above equation extends by linearity to all $X\in\mathscr{X}$, not necessarily of L^2-norm 1.

Now let (F_n) be a sequence in $\mathcal{S}(\mathscr{X})$ such that the following hold.

(i) $F_n \to 0$ in $L^q_{\mathscr{X}}(\Omega)$.

(ii) The sequence DF_n converges in the norm of $L^q_{\mathscr{X}}(\Omega, \mathscr{X})$ to some $\eta \in L^q_{\mathscr{X}}(\Omega, \mathscr{X})$.

We have to show that $\eta = 0$ a.s.. Let $X \in \mathscr{X}$, $G \in \mathcal{S}(\mathscr{X})$. Since $F_n \to 0$ in L^q and XG and $(DG, X)_{\mathscr{X}}$ belong to $L^{\frac{q}{q-1}}$ we deduce from (7.4) that

$$\mathbb{E}\big[\, G(\eta, X)\,\big] = \lim_{n\to\infty} \mathbb{E}\big[\, G(DF_n, X)_{\mathscr{X}}\,\big] = -\lim_{n\to\infty} \mathbb{E}\big[\, F_n(DG, X)_{\mathscr{X}}\,\big] + \lim_{n\to\infty} \mathbb{E}\big[\, XF_nG\,\big] = 0.$$

Thus

$$\mathbb{E}\big[\, G(\eta, X)_{\mathscr{X}}\,\big] = 0, \;\; \forall G \in \mathcal{S}(\mathscr{X}), \;\; X \in \mathscr{X}.$$

Since $\mathcal{S}(\mathscr{X})$ is dense in any L^r, $r \in [1, \infty)$, we deduce that

$$\forall X \in \mathscr{X}, \;\; (\eta, X)_{\mathscr{X}} = 0 \;\; a.s..$$

Thus, if $(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $\mathscr{X}$, there exists a negligible set $\mathcal{N} \subset \Omega$ such that

$$(\eta(\omega), e_n)_{\mathscr{X}} = 0, \;\; \forall n \in \mathbb{N}, \;\; \omega \in \Omega \setminus \mathcal{N}.$$

Thus $\eta = 0$ a.s. □

Definition 7.6. Let $p \in \mathbb{N}$ and $q \in [1, \infty)$. We define the Gaussian Sobolev space $\boldsymbol{D}^{p,q}(\mathscr{X})$ to be the closure of $\mathcal{S}(\mathscr{X})$ with respect to the norm

$$\|F\|_{\mathbb{D}^{p,q}} := \left(\sum_{k=0}^{p} \mathbb{E}\Big[\, \|D^kF\|^q_{\mathscr{X}^{\hat{\odot} k}}\,\Big] \right)^{\frac{1}{q}}. \qquad \square$$

According to Proposition 7.5, the operator D^p can be consistently extended as a continuous operator

$$D^p : \mathbb{D}^{p,q}(\mathscr{X}) \to L^q_{\mathscr{X}}(\Omega, \mathscr{X}^{\hat{\odot} p}).$$

Remark 7.7. (a) For any $\varepsilon \geq 0$, any $m \in \mathbb{N}_0$, any $p \in \mathbb{N}$ and any $q \in (1, \infty)$ we have

$$\mathbb{D}^{p,q+\varepsilon} \subset \mathbb{D}^{p+m,q}.$$

(b) The space $\mathbb{D}^{p,2}(\mathscr{X})$ is a Hilbert space with inner product

$$(F, G)_{\mathbb{D}^{p,2}} = \sum_{j=0}^{p} \mathbb{E}\big[\, (D^jF, D^jG)_{\mathscr{X}^{\hat{\odot} j}}\,\big].$$

(c) The space $\mathcal{S}(X)$ is dense in $\mathbb{D}^{p,q}(\mathscr{X})$, $\forall p \geq 0$, $q \in [1, \infty)$. □

Example 7.8. Suppose that $\mathscr{X}$ is a finite dimensional Gaussian Hilbert space, $\dim \mathscr{X} = n$. Fix an orthonormal basis $X_1, \dots, X_n$. Then

$$L^2_{\mathscr{X}}(\Omega) \cong L^2\big(\mathbb{R}^n, \boldsymbol{\Gamma}_{\mathbb{1}}(dx)\big), \;\; \boldsymbol{\Gamma}_{\mathbb{1}}(dx) = (2\pi)^{-n/2} e^{-\frac{|x|^2}{2}} dx.$$

If $f \in C^\infty(\mathbb{R}^n)$ is a function such that is derivatives of any order have at most polynomial growth, then the Malliavin gradient $Df(X_1, \dots, X_n)$ corresponds to the differential of f

$$df = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} dx_k.$$

Furthermore, the Gaussian Sobolev space corresponds to the weighted Sobolev space $W^{p,q}(\mathbb{R}^n, \mathbf{\Gamma})$ equipped with the norm

$$\|f\|_{\mathbb{D}^{p,q}} = \left(\sum_{|\alpha| \leq p} \int_{\mathbb{R}^n} |\partial_x^\alpha f(x)|^q \, \mathbf{\Gamma}_1(dx) \right)^{\frac{1}{q}}. \qquad \square$$

Proposition 7.9. *Let $f \in \widehat{\mathscr{X}}$, $p \in \mathbb{N}$. Recall that Proj_n denotes the orthogonal projection onto the n-th chaos $\mathscr{X}^{:n:}$. The following statements are equivalent.*

(i) $F \in \mathbb{D}^{p,2}(\mathscr{X})$.

(ii)

$$\sum_{n \geq 0} n^p \| \operatorname{Proj}_n F\|^2 < \infty.$$

Outline of proof. Fix an orthonormal basis $\underline{X} = (X_k)_{k \geq 1}$ of $\mathscr{X}$. We have

$$\operatorname{Proj}_n = \sum_{\alpha \in \mathbb{N}_0, \ |\alpha| = n} c_\alpha(f) H_\alpha(\underline{X}).$$

From the equality (2.4b) we deduce that

$$\int_{\mathbb{R}^N} \partial_j H_\alpha(x) \partial_j H_\beta(x) \mathbf{\Gamma}(dx) = \alpha_j \big(H_\alpha, H_\beta \big)_{L^2(\mathbf{\Gamma})}.$$

This implies that

$$\|DH_\alpha\|_{L^2}^2 = |\alpha| \|H_\alpha\|_{L^2}.$$

In particular, we deduce that

$$\mathscr{X}^{:n:} \subset \mathbb{D}^{1,2}(\mathscr{X}), \ \ \|F\|_{\mathbb{D}^{1,2}}^2 = (1+n)\|F\|_{L^2}^2, \ \ \forall F \in \mathscr{X}^{:n:}.$$

The proposition is now an immediate consequence of the above fact. $\square$

Example 7.10. For any $n \in \mathbb{N}$, and any $p \in \mathbb{N}_0$, the n-th chaos $\mathscr{X}^{:n:}$ is contained in $\mathbb{D}^{p,2}(\mathscr{X})$. $\square$

Since $\mathcal{S}(\mathscr{X})$ is dense in in $\mathbb{D}^{1,q}(\mathscr{X})$ we obtain the following useful result.

Proposition 7.11 (Chain Rule). *Suppose that $\varphi : \mathbb{R}^m \to \mathbb{R}$ is a C^1-function with bounded derivatives. Then for any $F_1, \dots, F_n \in \mathbb{D}^{1,q}$ we have $\varphi(F_1, \dots, F_n) \in \mathbb{D}^{1,p}$ and*

$$D\varphi(F_1, \dots, F_m) = \sum_{j=1}^m \frac{\partial \varphi}{\partial x_j}(F_1, \dots, F_m) DF_m. \tag{7.6}$$

$\square$

The Chain Rule holds in the more general case when φ is a Lipschitz function, [28, Prop. 1.2.4].

Proposition 7.12 (Extended Chain Rule). *Suppose that $\varphi : \mathbb{R}^m \to \mathbb{R}$ is a Lipschitz function, then for any $F_1, \dots, F_n \in \mathbb{D}^{1,q}$ such that the probability distribution of*

$$\vec{F} = (F_1, \dots, F_m) : \Omega \to \mathbb{R}^m,$$

is absolutely continuous[10] with respect to the Lebesgue measure on $\mathbb{R}^m$, then $\varphi(\vec{F}) \in \mathbb{D}^{1,q}$ and (7.6) continues to hold with $\frac{\partial \varphi}{\partial x_i}$ defined a.e. $\square$

[10] This assumption is needed to give a precise meaning to $\frac{\partial \varphi}{\partial x_i}(\vec{F})$ since $\frac{\partial \varphi}{\partial x_i}(x)$ is well definite only for x outside a Lebesgue negligible subset $\mathcal{N} \subset \mathbb{R}^m$.

Remark 7.13. The Gaussian Hilbert spaces $\mathbb{D}^{k,p}$ can be given an alternate definition; see [12, Def. 15.59]. The fact that our definition agrees with the definition in [12] requires some work and it follows from [12, Thm. 15.104].

The definition in [12, Def. 15.59] has certain technical advantages. In particular, it leads naturally to the following important result.

> *The probability distribution of any non-constant random variable $F \in \mathbb{D}^{1,p}(\mathscr{X})$, $p \in [1, \infty)$, is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

For a proof we refer to [12, Thm. 15.50]. Yet other approaches to this absolute continuity theorem can be found in [21, Thm. III.7.1] or [33]. □

8. The divergence operator

The *divergence operator* δ is the adjoint of the Malliavin gradient viewed as a closed *unbounded* operator

$$D : \mathbb{D}^{1,2}(\mathscr{X}) \subset L^2_{\mathscr{X}}(\Omega) \to L^2_{\mathscr{X}}(\Omega, \mathscr{X}).$$

Similarly, for $p \in \mathbb{N}$, the operator δ^p is the adjoint of the closed unbounded operator

$$D^p : \mathbb{D}^{p,2}(\mathscr{X}) \subset L^2_{\mathscr{X}}(\Omega) \to L^2_{\mathscr{X}}(\Omega, \mathscr{X}).$$

The domain of δ^p is the space

$$\operatorname{Dom}(\delta^p) := \Big\{ u \in L^2(\Omega, \mathscr{X}^{\otimes p}); \exists C > 0 \;\; \big| \mathbb{E}\big[(D^p F, u)_{\mathscr{X}^{\hat{\otimes} p}} \big] \big| \leq C \sqrt{\mathbb{E}\big[\, F^2 \,\big]}, \;\; \forall F \in \mathcal{S}(\mathscr{X}) \Big\}.$$

If $u \in \operatorname{Dom}(\delta^p)$, then $\delta^p u$ is the unique element in $L^2_{\mathscr{X}}(\Omega) = \widehat{\mathscr{X}}$ such that

$$\mathbb{E}\big[\, F \delta^p u \,\big] = \mathbb{E}\big[\, (D^p F, u)_{\mathscr{X}^{\otimes p}} \,\big], \;\; \forall F \in \mathcal{S}(\mathscr{X}). \tag{8.1}$$

Example 8.1. (a) Suppose that $\dim \mathscr{X} = n < \infty$. Fix an orthonormal basis $\{X_1, \dots, X_n\}$ of $\mathscr{X}$. Let

$$u = (u_1, \dots, u_n) \in L^2_{\mathscr{X}}(\Omega, \mathbb{R}^n).$$

Then each u_j is a measurable function of $(X_1, \dots, X_n)$. For any admissible function $f \in C^\infty(\mathbb{R}^n)$ we have

$$\mathbb{E}\big[\, f(X_1, \dots, X_n) \delta u \,\big] = \mathbb{E}\left[\sum_{j=1}^n f'_{x_j}(X_1, \dots, X_n) u_j(X_1, \dots, X_n) \right]$$

$$= (2\pi)^{-n/2} \int_{\mathbb{R}^n} \left(\sum_{j=1}^n f'_{x_j}(x) u_j(x) e^{-\frac{x_j^2}{2}} \right) dx = \int_{\mathbb{R}^n} f(x) \sum_{j=1}^n \big(-\partial_{x_j} u_j(x) + x_j u_j \big) \boldsymbol{\Gamma}_1(dx).$$

Thus

$$\delta(u_1, \dots, u_n) = \sum_{j=1}^n \big(-\partial_{x_j} u_j(X_1, \dots, X_n) + X_j u_j(X_1, \dots, X_n) \big).$$

Observe that in the case $n = 1$ the divergence operator coincides with the creation operator (2.2).

(b) Suppose that $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a separable Gaussian Hilbert space and $X \in \mathscr{X}$. It is not hard to verify that $DX \in \operatorname{Dom}(\delta)$. We want to compute δDX.

For $F \in L^2_{\mathscr{X}}(\Omega)$ we have

$$\mathbb{E}[F \delta DX] = \mathbb{E}\big[\, (Df, DX)_{\mathscr{X}} \,\big], \;\; \forall \mathcal{S}(\mathscr{X}).$$

We can assume that $\|X\|_{L^2} = 1$ and that $F = f(X_1, \dots, X_n)$, where $\{X_1, \dots, X_n\}$ is an orthonormal system, $X = X_1$ and f is an admissible function. We have

$$\mathbb{E}\big[\, (Df, DX)_{\mathscr{X}} \,\big] = \int_{\mathbb{R}^n} f'_{x_1}(x)\boldsymbol{\Gamma}_1(dx) = \int_{\mathbb{R}^n} f(x)x_1\boldsymbol{\Gamma}(dx) = \mathbb{E}[FX].$$

Hence

$$\delta(DX) = X, \;\; \forall X \in \mathscr{X}.$$

(c) Suppose that $F \in \mathcal{S}(\mathscr{X})$ and $X \in \mathscr{X}$. Then $D_X F = (DF, DX)_{\mathscr{X}} \in \mathcal{S}(X)$. Indeed, we can assume that $F = f(X_1, \dots, X_n)$, f admissible, $\{X_1, \dots, X_n\}$ orthonormal system, $X_1 = X$. Then

$$(DF, DX)_{\mathscr{X}} = f'_{x_1}(X_1, \dots, X_n) \in \mathcal{S}(X).$$

Observe that

$$D(D_X F) = D(DF, DX)_{\mathscr{X}} = \sum_{j=1}^n f''_{x_1 x_j}(X_1, \dots, X_n)DX_j = \frac{1}{2!} i_{X_1} D^2 F,$$

where for any $X \in \mathscr{X}$ we denoted by i_X the contraction

$$i_X : \mathscr{X}^{\otimes k} \to \mathscr{X}^{\otimes (k-1)}, \;\; k \in \mathbb{N}$$

which is the $\otimes$-derivation uniquely determined by the condition

$$i_X Y = (X, Y)_{\mathscr{X}}, \;\; \forall Y \in \mathscr{X}. \qquad \square$$

The next result follows immediately from the definition of δ. We refer to [26, Prop. 2.5.4] for details.

Proposition 8.2. *Let $F \in \mathbb{D}^{1,2}(\mathscr{X})$ and $u \in \operatorname{Dom}(\delta)$ such that*

$$\mathbb{E}[F^2\|u\|^2_{\mathscr{X}}] + \mathbb{E}[F^2\delta(u)^2] + \mathbb{E}[(DF, u)^2_{\mathscr{X}}] < \infty.$$

Then

$$\delta(Fu) = F\delta u - (DF, u)_{\mathscr{X}}. \qquad \square$$

Example 8.3. Suppose that $F \in \mathcal{S}(\mathscr{X})$, $X \in \mathscr{X}$ so that

$$u = FDX \in \mathcal{S}(\mathscr{X}, \mathscr{X}).$$

Then $u \in \operatorname{Dom}(\delta)$ and we deduce from Proposition 8.2 that

$$\delta u = FX - (DF, DX)_{\mathscr{X}} = FX - D_X F.$$

This shows that $\delta u \in \mathbb{D}^{1,2}$ and for any $Y \in \mathscr{X}$ we have

$$D_Y(\delta u) = (D_Y F)X + FD_Y X - D_Y D_X F = (D_Y F)X + F(X, Y)_{\mathscr{X}} - D_Y D_X F.$$

On the other hand

$$D_Y u = D_Y F \otimes DX,$$
$$\delta(D_Y F u) = (D_Y F)X - D_X D_Y F.$$

Hence

$$\begin{aligned} D_Y(\delta u) - \delta(D_Y u) &= (D_Y F)X + F(X, Y)_{\mathscr{X}} - D_Y D_X F - (D_Y F)X + D_X D_Y F \\ &= F(X, Y) + [D_X, D_Y]F = F(X, Y) = (u, Y)_{\mathscr{X}}. \end{aligned}$$

We have thus proved the *Heisenberg identity*

$$\forall u \in \operatorname{Dom}(\delta), \;\; Y \in \mathscr{X} \;\; [D_Y, \delta]u = (u, Y)_{\mathscr{X}}. \tag{8.2}$$

$\square$

The operator δ^p is closely related to the multiple Ito integrals. We have the following result.

Proposition 8.4. *Let $X_1, \dots, X_p \in \mathscr{X}$. Then (compare with (6.5))*

$$\delta^p(DX_1 \otimes \cdots \otimes DX_p) = \delta^p\big(\,\mathbf{Sym}[DX_1 \otimes \cdots \otimes DX_p]\,\big) = :X_1 \cdots X_p: . \tag{8.3}$$

Proof. Fix an orthonormal basis $\{Y_n\}_{n\in\mathbb{N}}$ of $\mathscr{X}$. Clearly it sufficers to prove the result in the special case when

$$DX_1 \odot \cdots \odot DX_p = (DY)^{\odot\alpha}, \;\; \alpha \in \mathbb{N}_0^{\mathbb{N}}, \;\; |\alpha| = p.$$

Suppose that $f = f(y_1, \dots, y_n)$ is an admissible function. Then

$$\mathbb{E}\big[\, f(Y_1, \dots, Y_n)\delta(DY^{\odot\alpha})\,\big] = \mathbb{E}\big[\, (D^p f(Y_1, \dots, Y_n), (DY)^{\odot\alpha})_{\mathscr{X}^{\odot p}}\,\big]$$

From Remark 7.4 we deduce

$$\mathbb{E}\big[\, (D^p f(Y_1, \dots, Y_n), (DY)^{\odot\alpha})_{\mathscr{X}^{\odot p}}\,\big] = \sqrt{p!}\sum_{|\beta|=p}\frac{1}{\beta!}\mathbb{E}\big[\,\partial_y^\beta f(Y_1, \dots, Y_n)(DY^{\odot\beta}, DY^{\odot\alpha})_{\mathscr{X}^{\odot p}}\,\big]$$

$$= \sqrt{p!}\mathbb{E}\big[\,\partial_y^\alpha f(Y_1, \dots, Y_n)\,\big] = \sqrt{p!}\int_{\mathbb{R}^n}\partial_y^\alpha f(y_1, \dots, y_n)\mathbf{\Gamma}(dy)$$

($\delta_{y_k} = -\partial_{y_k} + y_k\cdot$, $\delta_{y_k}^{\alpha_k}1 = H_{\alpha_k}(y_k)$)

$$= \sqrt{p!}\int_{\mathbb{R}^n} f(y)H_\alpha(y)\mathbf{\Gamma}(dy).$$

Hence

$$\delta^p((DY)^{\odot\alpha}) = \sqrt{|\alpha|!}H_\alpha(Y),$$

i.e.,

$$\delta^p\big(\,\mathbf{Sym}[(DY)^{\otimes\alpha}]\,\big) = \frac{1}{\sqrt{|\alpha|!}}\delta^p((DY)^{\odot\alpha}) = H_\alpha(Y) =: Y_1^{\alpha_1}\cdots Y_n^{\alpha_n}: .$$

This proves the second equality of (8.3). The first one is proved in a similar fashion. □

Remark 8.5. Using the equalities (5.1) and (5.4) we deduce that

$$\delta^p(u) = \sqrt{p!}\Theta_p(u), \;\; \forall p \in \mathbb{N}, \;\; \forall u \in \mathscr{X}^{\odot p}.$$

If we are given a Hilbert space isomorphism

$$Z : L^2(M, \mathcal{M}, \mu) \to \mathscr{X},$$

then the resulting map

$$L^2(M^p, \mu^{\otimes p}) \to \mathscr{X}^{\otimes p} \xrightarrow{\delta^p} \widehat{\mathscr{X}}$$

coincides with the multiple Wiener-Ito integral $\boldsymbol{I}_n$; see (6.5). For this reason we set

$$\boxed{\boldsymbol{I}_p[F] := \delta^p F = \delta^p(\mathbf{Sym}[F]) = \sqrt{p!}\Theta_p(\mathbf{Sym}[F]), \;\; \forall F \in \mathscr{X}^{\otimes p}}. \tag{8.4}$$

Using the isometry relation (5.5) we deduce that

$$\boxed{\mathbb{E}\big[\,\boldsymbol{I}_p[F]^2\,\big] = \|\,\boldsymbol{I}_p[F]\,\|^2 = p!\|F\,\|^2, \;\; \forall F \in \mathscr{X}^{\odot p}}. \tag{8.5}$$

□

Remark 8.6. For any Hilbert space H and any $k \in \mathbb{N}$ we have a Malliavin derivative

$$D_H^k : \mathcal{S}(\mathscr{X}, H) \to \mathcal{S}(\mathscr{X}, H \otimes \mathscr{X}^{\otimes k})$$

with adjoint δ_H^k defined by the equality

$$\mathbb{E}\big[\, F\big(\, h, \delta(Gh' \otimes u)\,\big)_H\,\big] = \mathbb{E}\big[\,\big(\, D^k F \otimes h, Gh' \otimes u\big)_{H \otimes \mathscr{X}^{\otimes k}}\,\big],$$

$\forall F, G \in \mathcal{S}(\mathscr{X})$, $h, h' \in H$, $u \in \mathscr{X}^{\otimes k}$. For any $p \in \mathbb{N}$ we have

$$D^{p+1} = D^p_{\mathscr{X}} \circ D, \quad \delta^{p+1} = \delta^p_{\mathscr{X}} \circ \delta.$$

Arguing as in the proof of Proposition 8.4 one can show

$$p\delta^{p-1}(u) = D\delta^p(u), \quad \forall u \in \mathscr{X} \otimes \mathscr{X}^{\odot(p-1)}. \tag{8.6}$$

The above equality generalizes (2.4a). In fact, (8.6) follows from (2.4a). If as in the previous remark we set

$$\boldsymbol{I}_{p-1}[u] = \delta^{p-1}(u), \quad \forall u \in \mathscr{X} \otimes \mathscr{X}^{\odot(p-1)}.$$

We can rewrite (8.6) as

$$p\boldsymbol{I}_{p-1}[u] = D\boldsymbol{I}_p[u], \quad \forall u \in \mathscr{X} \otimes \mathscr{X}^{\odot(p-1)}. \tag{8.7}$$

□

9. The Ornstein-Uhlenbeck semigroup

Let $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ be a separable Gaussian Hilbert space.

Definition 9.1. The *Ornstein-Uhlenbeck semigroup* is the semigroup of contractions $P_t : \widehat{\mathscr{X}} \to \widehat{\mathscr{X}}$, $t \geq 0$, defined by

$$T_t F = \sum_{n \geq 0} e^{-nt} \operatorname{Proj}_n F, \quad \forall F \in \widehat{\mathscr{X}}, \quad \forall t \geq 0,$$

where we recall that $\operatorname{Proj}_n : \widehat{\mathscr{X}} \to \mathscr{X}^{:n:}$ denotes the orthogonal projection onto the n-th chaos. □

The above definition shows that T_t is indeed a semigroup of selfadjoint L^2-contractions. We want to present an equivalent, coordinate dependent description of this semigroup.

Fix an *orthonormal* basis of $\mathscr{X}$,

$$\underline{X} = (X_1, X_2, \dots, X_n, \dots).$$

Observe that the semigroup T_t is uniquely determined by its action on $\mathcal{P}(\mathscr{X})$.

Proposition 9.2 (Mehler's formula)**.** *Let $P : \mathbb{R}^m \to \mathbb{R}$ be a polynomial in m real variables. Set*

$$\vec{X} := (X_1, \dots, X_m).$$

Then

$$T_t[P(\vec{X})](\omega) = \int_{\mathbb{R}^m} P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1 - e^{-2t}}y\Big)\boldsymbol{\Gamma}_{\mathbb{1}}(dy), \tag{9.1}$$

where $\boldsymbol{\Gamma}_{\mathbb{1}}$ denotes the canonical Gaussian measure on the Euclidean space $\mathbb{R}^m$.

Proof. It suffices to prove the result in the special case when $P(\vec{X}) = H_\alpha(\vec{X})$, $\alpha \in \mathbb{N}_0^m$. In this case, the left-hand side of (9.1) is equal to

$$T_t[P(\vec{X})](\omega) = e^{-|\alpha|t} H_\alpha(\vec{X}(\omega)) = \prod_{j=1}^m e^{-\alpha_j t} H_{\alpha_j}(X_j(t)).$$

The Fubini theorem shows that the right-hand side of (9.1) is equal in this case to

$$\prod_{j=1}^{m}\int_{\mathbb{R}} H_{\alpha_j}\Big(e^{-t}X_j(\omega)+\sqrt{1-e^{-2t}}\,y\Big)\boldsymbol{\gamma}_1(dy)$$

Thus, to prove (9.1) it suffices to prove that

$$\int_{\mathbb{R}} H_n\Big(e^{-t}x++\sqrt{1-e^{-2t}}\,y\Big)\boldsymbol{\gamma}(dy)=e^{-nt}H_n(x),\;\;\forall n\in\mathbb{N}_0,\;\;t\geq 0,\;\;\forall x\in\mathbb{R}. \tag{9.2}$$

We follow closely the presentation in the proof of [20, Prop. V.1.5.4]. We have the following useful identities.

Lemma 9.3. *Define the linear operator*

$$T_t:\mathbb{R}[x]\to\mathbb{R}[x],\;\;T_tP(x)=\int_{\mathbb{R}} P\Big(e^{-t}x++\sqrt{1-e^{-2t}}\,y\Big)\boldsymbol{\gamma}_1(dy),$$

Then the following hold.

(i) *The operator T_t is symmetric with respect to the $L^2(\boldsymbol{\gamma})$-inner product on $\mathbb{R}[x]$.*
(ii) $\partial_x T_t=e^{-t}T_t\partial_x$.
(iii) $T_t\delta_x=e^{-t}\delta_x T_t$.

Proof of Lemma 9.3. To prove (i) observe that

$$(T_tP,Q)=\int_{\mathbb{R}}\int_{\mathbb{R}} P\Big(e^{-t}x++\sqrt{1-e^{-2t}}\,y\Big)Q(x)\boldsymbol{\gamma}_1(dy)\boldsymbol{\gamma}_1(dx). \tag{9.3}$$

Set $a=e^{-t}$, $b=\sqrt{1-e^{-2t}}$ so that $a^2+b^2=1$. We have

$$(T_tP,Q)=\int_{\mathbb{R}^2} P(ax+by)Q(x)\boldsymbol{\Gamma}_{\mathbb{1}}(dxdy).$$

Now consider the *orthogonal* change in variables

$$\left[\begin{array}{c} x\\ y\end{array}\right]=\left[\begin{array}{cc} b & a\\ -a & b\end{array}\right]\cdot\left[\begin{array}{c} u\\ v\end{array}\right].$$

Since $\boldsymbol{\Gamma}_{\mathbb{1}}$ is invariant under orthogonal transformations we deduce

$$\int_{\mathbb{R}^2} P(ax+by)Q(x)\boldsymbol{\Gamma}_{\mathbb{1}}(dxdy)=\int_{\mathbb{R}^2} P(v)Q(av+bu)\boldsymbol{\Gamma}_{\mathbb{1}}(dudv)=(P,T_tQ).$$

This proves (i). The equality (ii) follows by differentiating the definition (9.3) of $T_t[P]$. The equality (iii) is obtained from (ii) by passing to adjoints, and using the symmetry of T_t proved in (i). □

Clearly, $T_t1=1$. From Lemma 9.3(iii) we deduce that

$$T_tH_n=T_t\delta_x^n 1=e^{-nt}\delta_x^n T_t1=e^{-nt}H_n.$$

□

Definition 9.4. The *Ornstein-Uhlenbeck operator*, denoted by L, is the infinitesimal generator of the Ornstein-Uhlenbeck semigroup. □

Recall that $F\in\operatorname{Dom}(L)\subset\widehat{\mathscr{X}}$ if and only if the limit

$$\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF-F\Big)$$

exists in $L^2(\Omega,\widehat{\mathfrak{F}},\mathbb{P})$. We denote by LF the above limit.

Proposition 9.5.

$$\operatorname{Dom}(L) = \Big\{ F \in \mathscr{X};\ \sum_{n\geq 0} n^2 \|\operatorname{Proj}_n\|_{L^2}^2 < \infty \Big\} = W^{2,2}(\mathscr{X}).$$

$$\forall n \in \mathbb{N},\ \ \forall F \in \mathscr{X}^{:n:}:\ \ LF = -nF = -\delta DF.$$

Proof. Let $F \in \widehat{\mathscr{X}}$. We set $F_n = \operatorname{Proj}_n F$. Then

$$\frac{1}{t}\Big(T_tF - F\Big) = \sum_{n\geq 0} \frac{e^{-nt}-1}{t} F_n.$$

Now observe that

$$\Big|\frac{e^{-nt}-1}{t}\Big| \leq n,\ \ \forall t > 0,\ \ N \in \mathbb{N}_0.$$

so that

$$\Big\|\frac{1}{t}\Big(T_tF - F\Big)\Big\|_{L^2}^2 \leq \sum_{n\geq 0} n^2 \|F_n\|_{L^2}^2$$

This proves that if

$$\sum_{n\geq 0} n^2 \|F_n\|_{L^2}^2 < \infty,$$

then

$$\lim_{t\searrow 0} \frac{1}{t}\Big(T_tF - F\Big)$$

exists in L^2 and it is equal to

$$\sum_{n\geq 0} \frac{d}{dt}\Big|_{t=0} e^{-nt}F_n = -\sum_{n\geq 0} nF_n.$$

Conversely, if the above limit exists in L^2, then

$$\operatorname{Proj}_n\left(\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF - F\Big)\right) = \lim_{t\searrow 0}\operatorname{Proj}_n\left(\frac{1}{t}\Big(T_tF - F\Big)\right) = -nF_n.$$

Thus

$$\lim_{t\searrow 0}\frac{1}{t}\Big(T_tF - F\Big) = -\sum_{n\geq 0} nF_n \in L^2 \Rightarrow \sum_{n\geq 0} n^2\|F_n\|_{L^2}^2 < \infty.$$

The equality $LF = -nF$, $f \in \mathscr{X}^{:n:}$ follows from the above discussion. To prove the equality $\delta DF = n$, $F \in \mathscr{X}^{:n:}$ it suffices to consider only the special case when $F = H_\alpha(X_1, \dots, X_k)$ where (X_j) is an orthonormal system and α is a multi-index such that $|\alpha| = n$. In this case the equality follows from (2.4b). □

Remark 9.6. Using Proposition 7.9 we deduce that $\operatorname{Dom}(L) = \mathbb{D}^{2,2}(\mathscr{X})$. Note also that L is *non-positive*. □

Example 9.7. (a) Suppose that $\dim \mathscr{X} = n$. By fixing an orthonormal basis $X_1, \dots, X_n$ of $\mathscr{X}$ we can identify $\widehat{\mathscr{X}}$ with $L^2(\mathbb{R}^n, \boldsymbol{\Gamma}_{\mathbb{1}})$. Then

$$Lf = \sum_{j=1}^n \partial_{x_j}^2 f - \sum_{j=1}^n x_j\partial_{x_j} f = (-\Delta - x\nabla)f,$$

for any function $f \in C^2(\mathbb{R}^n)$ with bounded 2nd order derivatives. Above, Δ is the Euclidean *geometers'* Laplacian. □

Definition 9.8. We define L^{-1} to be the bounded operator $L^{-1}:\widehat{\mathscr{X}}\to\widehat{\mathscr{X}}$ given by

$$L^{-1}F=-\sum_{n\geq 1}\frac{1}{n}\operatorname{Proj}_n F. \qquad \square$$

Note that L^{-1} is a pseudo-inverse of L. More precisely, if $F\in\mathbb{D}^{2,2}(\mathscr{X})$ is such that $\mathbb{E}[F]=0$, i.e., $\operatorname{Proj}_0 F=0$, then

$$L^{-1}LF=LL^{-1}F=F.$$

Proposition 9.9. *Let $F\in\mathbb{D}^{1,2}(\mathscr{X})$. Then for any $X\in\mathscr{X}$, $\|X\|_{L^2}=1$, we have*

$$D_XL^{-1}F=-\int_0^\infty e^{-t}T_tD_XFdt=(L-\mathbb{1})^{-1}D_XF. \tag{9.4}$$

Proof. It suffices to prove the result in the special case

$$F=H_\alpha(X_1,\dots,X_m),$$

where $\{X_1,\dots,X_m\}\subset\mathscr{X}$ is an orthonormal system, $X=X_1$, $|\alpha|=n>0$. Note that

$$D_XF=\alpha_1H_\beta(X_1,\dots,X_m),\;\;\beta=(\alpha_1-1,\alpha_2,\dots,\alpha_m).$$

Using the identity

$$\frac{1}{n}=\int_0^\infty e^{-nt}dt$$

we deduce

$$L^{-1}F=-\frac{1}{n}F=-\int_0^\infty T_tFdt\Rightarrow D_XL^{-1}F=-\int_0^\infty D_XT_tFdt=-\int_0^\infty e^{-t}T_tD_XFdt.$$

On the other hand

$$(L-\mathbb{1})^{-1}D_XF=(L-\mathbb{1})^{-1}[\alpha_1H_\beta]=-\frac{1}{|\beta|+1}\alpha_1H_\beta=-\frac{1}{n}D_XF=D_XL^{-1}F.$$

$\square$

Proposition 9.10 (Key integration by parts formula)**.** *Suppose that $F,G\in\mathbb{D}^{1,2}(\mathscr{X})$ are non-constant and $g:\mathbb{R}\to\mathbb{R}$ is a Lipschitz function. Then*

$$\boxed{\mathbb{E}\big[\,Fg(G)\,\big]=\mathbb{E}[F]\cdot\mathbb{E}\big[\,g(G)\,\big]+\mathbb{E}\big[\,g'(G)\cdot(DG,-DL^{-1}F)_{\mathscr{X}}\,\big]}. \tag{9.5}$$

Proof. Suppose first that $g\in C^1(\mathbb{R})$. Let $F_\perp=F-\mathbb{E}[F]$. Then

$$\mathbb{E}\big[\,Fg(G)\,\big]=\mathbb{E}[F]\cdot\mathbb{E}\big[\,g(G)\,\big]+\mathbb{E}\big[\,F_\perp g(G)\,\big].$$

Since $F_\perp=LL^{-1}F$ we deduce

$$\begin{gathered}\mathbb{E}\big[\,F_\perp g(G)\,\big]=\mathbb{E}\big[\,LL^{-1}Fg(G)\,\big]=-\mathbb{E}\big[\,\delta DL^{-1}Fg(G)\,\big]\\=-\mathbb{E}\big[\,(\,DL^{-1}F,Dg(G)\,)_{\mathscr{X}}\,\big]=-\mathbb{E}\big[\,(\,DL^{-1}F,g'(G)DG\,)_{\mathscr{X}}\,\big]\\=\mathbb{E}\big[\,g'(G)(\,DG,-DL^{-1}F)_{\mathscr{X}}\,\big].\end{gathered}$$

To prove the general case when g is only Lipschitz, we approximate g by C^1 Lipschitz functions g_ε using a nonnegative mollifying family $(\rho_\varepsilon)_\varepsilon>0$,

$$g_\varepsilon=\rho_\varepsilon*g.$$

Then $\partial_xg_\varepsilon=\rho_\varepsilon*(\partial_xg)$ so that

$$\|\partial_xg_\varepsilon\|_\infty\leq\operatorname{Lip}(g)$$

where $\mathrm{Lip}(g)$ is the (best) Lipschitz constant of g. Since F, G are nonconstant, we deduce from Remark 7.13 that the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$. Thus, although g' is only defined a.e., the composition $g'(F)$ is a.s. well defined. To obtain (9.5) in the general case, use (9.5) for the functions g_ε and then let $\varepsilon \to 0$. □

10. THE HYPER-CONTRACTIVITY OF THE ORNSTEIN-UHLENBECK SEMIGROUP

We know that T_t defines a continuous semigroup of contractions $L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P}) \to L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P})$. It is not hard to see that T_t defines continuous semigroup of contractions $L^p(\Omega, \widehat{\mathcal{F}}, \mathbb{P}) \to L^p(\Omega, \widehat{\mathcal{F}}, \mathbb{P})$ for any $p \in (1, \infty)$. We limit ourself to proving that

$$\|T_t P\|_{L^p} \leq \|P\|_{L^p}, \;\; \forall P \in \mathcal{P}(\mathscr{X}).$$

To see this assume $P = P(\vec{X})$, where $\vec{X} = (X_1, \dots, X_m)$ is an orthonormal system in $\mathscr{X}$. Using Mehler's formula (9.1) we deduce

$$T_t[P(\vec{X})](\omega) = \int_{\mathbb{R}^m} P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1 - e^{-2t}}y\Big)\boldsymbol{\Gamma}_{\mathbb{1}}(dy)$$

Since the function $f(x) = x^p$, $x > 0$, is convex for $p > 1$ we deduce from Jensen's inequality that

$$\big|\, T_t[P(\vec{X})](\omega)\,\big|^p \leq \int_{\mathbb{R}^m} \Big|\, P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1 - e^{-2t}}\, y\Big)\,\Big|^p \boldsymbol{\Gamma}_{\mathbb{1}}(dy)$$

Invoking Jensen's inequality once again we conclude that

$$\mathbb{E}\big[\, |T_t P|^p\,\big] \leq \int_{\mathbb{R}^m} \mathbb{E}\Big[\Big|\, P\Big(e^{-t}\vec{X}(\omega) + \sqrt{1 - e^{-2t}}\, y\Big)\,\Big|^p\Big]\boldsymbol{\Gamma}_{\mathbb{1}}(dy)$$

$$= \int_{\mathbb{R}^m \times \mathbb{R}^m} \Big|\, P(e^{-t}x + \sqrt{1 - e^{-2t}}y)\Big|^p \boldsymbol{\Gamma}_{\mathbb{1}}(dx)\boldsymbol{\Gamma}_{\mathbb{1}}(dy) = \int_{\mathbb{R}^m} \big|\, P(\,x\,)\,\big|^p \boldsymbol{\Gamma}_{\mathbb{1}}(dx)$$

where at the last step we used the fact that if X, Y are independent standard normal random variables and $a^2 + b^2 = 1$, then $aX + bY$ is also a standard normal random variable.

The semigroup T_t satisfies a hypercontractivity property, namely, for any $p_0 \in (1, \infty)$ there exists an increasing, unbounded function $p : [0, \infty) \to (0, \infty)$ such that $p_0 = p(0)$ and, $\forall t \geq 0$, the operator T_t induces a bounded linear map $T_t : L^{p_0} \to L^{p(t)}$. We will spend the remainder of this subsection proving this fact.

Theorem 10.1 (The log-Sobolev inequality)**.** *For any $n \in \mathbb{N}$, and any $f \in W^{1,2}(\mathbb{R}^n, \boldsymbol{\Gamma})$ we have*

$$\int_{\mathbb{R}^n} f^2(x) \log f^2(x) \boldsymbol{\Gamma}(dx) \leq 2 \int_{\mathbb{R}^n} |\nabla f(x)|^2\, \boldsymbol{\Gamma}(dx) + \int_{\mathbb{R}^n} f^2(x)\boldsymbol{\Gamma}(dx) \log\left(\int_{\mathbb{R}^n} f^2(x)\, \boldsymbol{\Gamma}(dx)\right), \tag{10.1}$$

where $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{\mathbb{1}}$ is the canonical Gaussian measure on $\mathbb{R}^n$ and $0 \cdot \log 0 := 0$.

Proof. We follow the presentation in [3, §1.6]Assume first that $f \in C_b^\infty(\mathbb{R}^n)$, i.e., f and all its derivatives are bounded. We distinguish three cases.

A. $\exists c > 0$ such that $f(x) > c$, $\forall x \in \mathbb{R}^n$. Set $\varphi = f^2$ so that

$$\nabla f = \frac{1}{2\sqrt{\varphi}} \nabla\varphi$$

and (10.1) is equivalent to

$$\int_{\mathbb{R}^n} \varphi \log \varphi \, d\boldsymbol{\Gamma} - \int_{\mathbb{R}^n} \varphi d\boldsymbol{\Gamma} \log\left(\int_{\mathbb{R}^n} \varphi \, d\boldsymbol{\Gamma}\right) \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{1}{\varphi} |\nabla\varphi|^2\, d\boldsymbol{\Gamma}. \tag{10.2}$$

Consider the Ornstein-Uhlenbeck semigroup

$$T_t : L^2(\mathbb{R}^n, \mathbf{\Gamma}) \to L^2(\mathbb{R}^n, \mathbf{\Gamma}).$$

Using the equality (9.1) we deduce that

$$T_t[\varphi](x) \geq c, \;\; \forall x \in \mathbb{R}^n, \;\; t \geq 0.$$

Since

$$\lim_{t\to\infty} T_t[\varphi] \log T_t[\varphi] = \int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma} \log\left(\int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma} \right)$$

we see that the left-hand side of (10.2) is equal to

$$-\int_0^\infty \frac{d}{dt} \int_{\mathbb{R}^n} T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}.$$

Taking into account the fact that

$$\frac{d}{dt} T_t[g] = LT_t[g], \;\; \forall g \in C_b^\infty(\mathbb{R}^n)$$

we deduce

$$-\int_0^\infty \frac{d}{dt} \int_{\mathbb{R}^n} T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}$$
$$= -\int_0^\infty \int_{\mathbb{R}^n} LT_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma} - \int_0^\infty \int_{\mathbb{R}^n} T_t[\varphi] \frac{1}{T_t[\varphi]} \frac{d}{dt} T_t[\varphi] \, d\mathbf{\Gamma}$$
$$= -\int_0^\infty \int_{\mathbb{R}^n} LT_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma} - \int_0^\infty \int_{\mathbb{R}^n} LT_t[\varphi] \, d\mathbf{\Gamma}.$$

Since L is symmetric and $L1 = 0$ we deduce

$$\int_{\mathbb{R}^n} LT_t[\varphi] \, d\mathbf{\Gamma} = 0.$$

Hence

$$\int_{\mathbb{R}^n} \varphi \log \varphi \, d\mathbf{\Gamma} - \int_{\mathbb{R}^n} \varphi d\mathbf{\Gamma} \log\left(\int_{\mathbb{R}^n} \varphi \, d\mathbf{\Gamma} \right) = -\int_0^\infty \int_{\mathbb{R}^n} LT_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma}$$
$$= \int_0^\infty \int_{\mathbb{R}^n} \delta D T_t[\varphi] \log T_t[\varphi] \, d\mathbf{\Gamma} = \int_0^\infty \int_{\mathbb{R}^n} \big(\nabla T_t[\varphi], \nabla \log T_t[\varphi] \big) \, d\mathbf{\Gamma}$$
$$= \int_0^\infty \underbrace{\int_{\mathbb{R}^n} \frac{1}{T_t[\varphi]} \big| \nabla T_t[\varphi] \big|^2 d\mathbf{\Gamma}}_{F(t)}.$$

Using Lemma 9.3 (ii) we deduce

$$\partial_{x_i} T_t[\varphi] = e^{-t} T_t[\partial_{x_i} \varphi], \;\; \forall i = 1, \dots, n,$$

so that

$$F(t) = e^{-2t} \int_{\mathbb{R}^n} \frac{1}{T_t[\varphi]} \sum_{i=1}^n \big(T_t[\partial_{x_i} \varphi] \big)^2 d\mathbf{\Gamma}.$$

The equality (9.1) implies that for any $g, h \in C_b^\infty(\mathbb{R}^n)$ we have

$$T_t[g] \leq T_t[\,|g|\,] \leq \|g\|_{L^\infty}, \;\; \big(T_t[gh] \big)^2 \leq T_t[g^2] T_t[h]^2.$$

Hence

$$\frac{1}{T_t[\varphi]} \big(T_t[\partial_{x_i} \varphi] \big)^2 = \frac{1}{T_t[\varphi]} \left(T_t \left[\sqrt{\varphi} \cdot \frac{\partial_{x_i} \varphi}{\sqrt{\varphi}} \right] \right)^2 \leq T_t \left[\frac{(\partial_{x_i} \varphi)^2}{\varphi} \right] \leq \frac{(\partial_{x_i} \varphi)^2}{\varphi}.$$

Thus

$$F(t) \leq e^{-2t} \int_{\mathbb{R}^n} \frac{|\nabla \varphi|^2}{\varphi} d\mathbf{\Gamma}.$$

The inequality (10.2) follows by integrating the above inequality.

B. $f \in W^{1,2}(\mathbb{R}^n, \mathbf{\Gamma})$, $f \geq 0$ a.s. . This case follows from case **A** by choosing a family of functions $f_\varepsilon \in C_b^\infty(\mathbb{R}^n)$, $f_\varepsilon \geq \varepsilon$, $f_\varepsilon \to f$ in $W^{1,2}$ and then letting $\varepsilon \searrow 0$.

The general case, $f \in W^{1,2}(\mathbb{R}^n, \mathbf{\Gamma})$, follows from case **B** applied to $|f|$. □

Remark 10.2. If $(\Omega, \mathcal{O}, \mu)$ is a probability space and $f : \Omega \to [0, \infty)$ is measurable function, then its entropy with respect to μ is

$$\operatorname{Ent}_\mu(f) = \begin{cases} \mathbb{E}_\mu[f \log f] - \mathbb{E}_\mu[f] \log \mathbb{E}_\mu[f], & \mathbb{E}_\mu[\log(1+f)] < \infty, \\ \\ +\infty, & \mathbb{E}_\mu[\log(1+f)] = \infty. \end{cases}$$

where $0 \log 0 := 0$. Observe that $\operatorname{Ent}_\mu(f)$ is nonnegative and positively homogeneous of degree 1. The log-Sobolev inequality (10.1) can be rewritten as

$$\operatorname{Ent}_{\mathbf{\Gamma}}[f^2] \leq 2 \int_{\mathbb{R}^n} |\nabla f(x)|^2 \mathbf{\Gamma}(dx).$$

As explained in [18, Sec.5.1], the log-Sobolev inequality leads to rather sharp concentration of measure inequalities. □

Theorem 10.3 (Hypercontractivity)**.** *Let* $p \in (1, \infty)$. *Define*

$$q(t) := 1 + e^{2t}(p-1), \ \ \forall t \geq 0.$$

Then

$$\|T_t f\|_{L^{q(t)}} \leq \|f\|_{L^p}, \ \ \forall f \in L^p(\mathbb{R}^n, \mathbf{\Gamma}), \ \ t \geq 0. \tag{10.3}$$

Proof. We follow closely the arguments in [3, Thm. 1.6.2]. It suffices to prove the inequality for smooth functions $f \in C_b^\infty(\mathbb{R}^n)$ such that

$$c := \inf_{x \in \mathbb{R}^n} f(x) > 0.$$

Under this assumption the function $[0, \infty) \ni t \mapsto G(t) = \|f\|_{L^{q(t)}}$ is differentiable. The inequality (10.3) reads $G(t) \leq G(0)$ so it it suffices to prove that $G'(t) \leq 0$, $\forall t \geq 0$.

Applying the log-Sobolev inequality to the function $f^{r/2}$, $r > 0$, we deduce

$$\int_{\mathbb{R}^n} f^r \log f d\mathbf{\Gamma} - \frac{1}{r} \int_{\mathbb{R}^n} f^r d\mathbf{\Gamma} \left(\log \int_{\mathbb{R}^n} f^r d\mathbf{\Gamma} \right) \leq \frac{r}{2} \int_{\mathbb{R}^n} (f^{r-2} \nabla f, \nabla f) d\mathbf{\Gamma}$$
$$= \frac{r}{2(r-1)} \int_{\mathbb{R}^n} (\nabla f^{r-1}, \nabla f) d\mathbf{\Gamma} = -\frac{r}{2(r-1)} \int_{\mathbb{R}^n} f^{r-1} L f d\mathbf{\Gamma}.$$

Hence

$$\int_{\mathbb{R}^n} f^r \log f d\mathbf{\Gamma} - \frac{1}{r} \int_{\mathbb{R}^n} f^r d\mathbf{\Gamma} \left(\log \int_{\mathbb{R}^n} f^r d\mathbf{\Gamma} \right) \leq -\frac{r}{2(r-1)} \int_{\mathbb{R}^n} f^{r-1} L f d\mathbf{\Gamma}, \ \ \forall r > 0. \tag{10.4}$$

We set

$$F(t) := \int_{\mathbb{R}^n} T_t[f] \, d\mathbf{\Gamma}.$$

Then $G(t) = F(t)^{1/q(t)}$ and we have

$$G'(t) = G(t)\left(-\frac{q'(t)}{q(t)^2}\log F(t) + \frac{F'(t)}{q(t)F(t)}\right).$$

Since $q'(t) = 2q(t) - 2 > 0$ it suffices to show that

$$-\frac{1}{q(t)}F(t)\log F(t) + \frac{F'(t)}{q(t)} \leq 0. \tag{10.5}$$

Observing that

$$F'(t) = \int_{\mathbb{R}^n} (T_t[f])^{q(t)}\left(q'(t)\log T_t[f] + q(t)\frac{LT_t[f]}{T_t[f]}\right)d\mathbf{\Gamma}$$

we conclude that (10.4) is equivalent to

$$-\frac{F(t)\log F(t)}{q(t)} + \int_{\mathbb{R}^n}\big(T_t[f]\big)^{q(t)}\log T_t[f]d\mathbf{\Gamma} + \frac{q(t)}{q'(t)}\int_{\mathbb{R}^n}\big(T_t[f]\big)^{q(t)-1}LT_t[f]d\mathbf{\Gamma} \leq 0.$$

This is precisely the inequality (10.4) with $r = q(t)$. □

We conclude by mentioning, without proof, the Kree-Meyer inequality.

Theorem 10.4 (Kree-Meyer). *For any $p \in (1,\infty)$, and any $k, \ell \in \mathbb{N}_0$, there exist positive constants $c_p(k,\ell) < C_p(k\ell$ such that*

$$c_p\|F\|_{\mathbb{D}^{k+\ell,p}} \leq \|(\mathbb{1} - L)^{\frac{\ell}{2}}F\|_{\mathbb{D}^{k,p}} \leq C_p\|F\|_{\mathbb{D}^{k+\ell,p}}, \quad \forall F \in \mathcal{S}(\mathscr{X}). \tag{10.6}$$

□

For a proof we refer to [3, Sec. 5.6], [20, Chap 2] or [28, Sec. 1.5]

PART 3. LIMIT THEOREMS

11. THE STEIN METHOD

11.1. **Metrics on spaces of probability measures.** Let us recall several concepts of of pseudo-distances on the spaces of Borel probability measures on $\mathbb{R}^d$.

Definition 11.1. Let $\mathcal{H}$ be a set of Borel measurable functions $\mathbb{R}^d \to \mathbb{R}$. We denote by $\mathcal{P}(\mathbb{R}^d)$ the space of Borel probability measures on $\mathbb{R}^d$.

(i) We set

$$\mathcal{P}(\mathbb{R},\mathcal{H}) := \big\{\,\mu \in \mathcal{P}(\mathbb{R}^d);\;\; \mathcal{H} \subset L^1(\mathbb{R}^d,\mu)\,\big\}.$$

(ii) We say that $\mathcal{H}$ is called *separating* if for any $\mu,\nu \in \mathcal{P}(\mathbb{R}^d)$

$$\mu = \nu \Longleftrightarrow \mathbb{E}_\mu[h] = \mathbb{E}_\nu[h], \;\; \forall h \in \mathcal{H} \cap L^1(\mathbb{R}^d,\,\mu) \cap L^1(\mathbb{R},\,\nu).$$

(iii) If $\mathcal{F}$ is separating and $\mu,\nu \in \mathcal{P}(\mathbb{R}^d,\mathcal{H})$, we set

$$\operatorname{dist}_{\mathcal{H}}(\mu,\nu) := \sup_{h\in\mathcal{H}}\big|\,\mathbb{E}_\mu[h] - \mathbb{E}_\nu[h]\,\big|.$$

(iv) If $(\Omega,\mathcal{O},\mathbb{P})$ is a probability space and $F,G : \Omega \to \mathbb{R}^d$ are random variables whose probability distributions belong to $\mathcal{P}(\mathbb{R}^d,\mathcal{H})$, then we set

$$\operatorname{dist}_{\mathcal{H}}(F,G) := \sup_{h\in\mathcal{H}}\big|\,\mathbb{E}[h(F)] - \mathbb{E}[h(G)]\,\big|.$$

□

It is easy to check that if $\mathcal{H}$ is separating, then $\mathrm{dist}_{\mathcal{H}}$ is indeed a metric on $\mathcal{P}(\mathbb{R}^d, \mathcal{H})$.

Example 11.2. (a) If $\mathcal{H}$ is the class of functions

$$\boldsymbol{1}_{(-\infty,c_1]\times\cdots\times(-\infty,c_d]},\quad c_1,\dots,c_d\in\mathbb{R},$$

then the resulting metric $\mathrm{dist}_{\mathcal{H}}$ on $\mathcal{P}(\mathbb{R}^d)$ is called the *Kolmogorov distance* and it is denoted by dist_{Kol}.

(b) If $\mathcal{H}$ is the class of functions Borel measurable functions $h:\mathbb{R}^d\to[0,1]$, then $\mathcal{H}$ is separating then the resulting metric on $\mathcal{P}(\mathbb{R}^d)$ is called the *total variation* metric and it is denoted by dist_{TV}.

(c) If $\mathcal{H}$ is the class of Lipschitz continuous functions $\mathbb{R}^d\to\mathbb{R}$ satisfying $\mathrm{Lip}(h)\leq 1$, where $\mathrm{Lip}(h)$ is the (best) Lipschitz constant of h, then $\mathcal{H}$ is separating, the resulting metric is called the *Wasserstein* metric and it is denoted by dist_W.

(d) If $\mathcal{H}$ denotes the class of Lipschitz continuous functions $h:\mathbb{R}^d\to\mathbb{R}$ such that

$$\|h\|_{L^\infty}+\mathrm{Lip}(h)\leq 1,$$

then $\mathcal{H}$ is separating, the resulting distance is called the *Fortet-Mourier* metric and it is denoted by dist_{FM}.

(e) If $\mathcal{H}\subset C_b^2(\mathbb{R}^d)$ denotes the class of C^2-functions $f:\mathbb{R}^d\to\mathbb{R}$ satisfying

$$\|f\|_{C^2}\leq 1,$$

then $\mathcal{F}$ is separating. We denote by dist_{C^2} the resulting metric. □

Remark 11.3. (a) Clearly

$$\mathrm{dist}_{Kol}\leq\mathrm{dist}_{TV},\quad \mathrm{dist}_W\geq\mathrm{dist}_{FM}\geq\mathrm{dist}_{C^2}\,.$$

Thus

$$\lim_{n\to\infty}\mathrm{dist}_{TV}(F_n,F)=0\Rightarrow\lim_{n\to\infty}\mathrm{dist}_{Kol}(F_n,F)=0.$$

Moreover, if

$$\lim_{n\to\infty}\mathrm{dist}_{Kol}(F_n,F)=0,$$

then $F_n\to F$ in law.

(b) Also, one can prove (see [9, Thm.11.3.3] that $F_n\to F$ in distribution of and only if $F_n\to F$ in the Fortet-Mornier metric. It is not hard to see that dist_{C^2} induces on $\mathcal{P}(\mathbb{R}^d)$ the same topology as dist_{FM}, the topology of convergence in law. Moreover, (see [6, Thm.3.3]), if $N\sim\mathcal{N}(0,1)$, then

$$\mathrm{dist}_{Kol}(F,N)\leq 2\sqrt{\mathrm{dist}_W(F,N)}.$$ □

The Stein method provides a way of estimating the distance between a random variable and a normal random variable. We present the bare-bones minimum referring to [5, 6, 32] for more details and many more applications. We follow the presentation in [26, Chap.3,4]. It all starts with the following simple observation.

11.2. **The one-dimensional Stein method.** Suppose that $N\sim\mathcal{N}(0,1)$ and $g\in\mathbb{D}^{1,2}(\mathbb{R})$, i.e., $g(N),g'(N)\in L^2$. Then

$$\int_{\mathbb{R}}\big(-g'(x)+xg(x)\big)\boldsymbol{\gamma}_1(dx)=\int_{\mathbb{R}}\delta_x g(x)\cdot 1\,\boldsymbol{\gamma}_1(dx)=\int_{\mathbb{R}}g(x)\cdot(\partial_x 1)\boldsymbol{\gamma}_1(dx)=0,$$

so that

$$\mathbb{E}[\,Ng(N)\,]=\mathbb{E}[\,g'(N)\,],\quad\forall g\in\mathbb{D}^{1,2}(\mathbb{R}).\tag{11.1}$$

It turns out the the converse is also true.

Lemma 11.4 (Stein's Lemma)**.** *A random variable N is a standard normal random variable if and only if for all $g \in \mathbb{D}^{1,2}(\mathbb{R})$ we have $g(N), g'(N) \in L^1$ and*

$$\mathbb{E}[\, Ng(N)\,] = \mathbb{E}[\, g'(N)\,]. \tag{11.2}$$

Proof. We have to prove only the *if* part. Applying (11.2) with $g(x) = x^k$, $k = 0, 1, \dots,$ we deduce

$$\mathbb{E}\big[\, N^{k+1}\,\big] = k\mathbb{E}\big[\, N^{k-1}\,\big],\ \ \forall k = 0, 1, 2, \dots .$$

This proves that

$$\mathbb{E}\big[\, N^k\,\big] = \int_{\mathbb{R}} x^k \boldsymbol{\gamma}_1(dx),\ \ \forall k = 0, 1, 2 \dots,$$

so that

$$\mathbb{E}\big[\, e^{\boldsymbol{i}tN}\,\big] = \int_{\mathbb{R}} e^{\boldsymbol{i}tx} \boldsymbol{\gamma}_1(dx),\ \ \forall t \in \mathbb{R} \Rightarrow N \sim \mathcal{N}(0,1).$$

□

Stein's lemma suggests that for a random variable X the quantity $\mathbb{E}[Xf(X) - f'(X)\,\big]$ should give an indication of how far away is the distribution of X from the normal distribution.

Definition 11.5. Let $N \sim \mathcal{N}(0,1)$ and $h \in L^1(\mathbb{R}, \boldsymbol{\gamma}_1)$. The *Stein's equation* associated to h is the o.d.e.

$$\underbrace{g'(x) - xg(x)}_{=-\delta_x g(x)} = h(x) - \int_{\mathbb{R}} h(x)\boldsymbol{\gamma}_1(dx) = h(x) - \mathbb{E}[h(N)]. \tag{11.3}$$

We set $h_\perp(x) := h(x) - \mathbb{E}[h(N)]$ so that

$$\mathbb{E}[\, h_\perp(N)\,] = 0.$$

□

Observe that Stein's equation can be rewritten as

$$e^{\frac{x^2}{2}} \partial_x\big(\, e^{-\frac{x^2}{2}} g(x)\,\big) = h_\perp(x),$$

This implies immediately the following result.

Proposition 11.6. *The general solution of (11.3) has the form*

$$g(x) = g_{h,c}(x) = ce^{\frac{x^2}{2}} + e^{\frac{x^2}{2}} \int_{-\infty}^x h_\perp(y) e^{-\frac{y^2}{2}}\, dy,\ \ x \in \mathbb{R}, \tag{11.4}$$

where $c \in \mathbb{R}$ is an arbitrary real constant. Moreover the solution

$$g_h(x) := g_{h,c=0} = e^{\frac{x^2}{2}} \int_{-\infty}^x h_\perp(y) e^{-\frac{y^2}{2}}\, dy \tag{11.5}$$

is the unique solution $g(x)$ of (11.3) such that

$$\lim_{x \to \pm\infty} e^{-\frac{x^2}{2}} g(x) = 0. \tag{11.6}$$

□

If now F is a random variable, then taking the expectation of the equality

$$g_h'(x) - xg_h(x) = h(x) - \mathbb{E}[\, h(N)\,]$$

with respect to the probability distribution of F we deduce

$$\mathbb{E}[\, h(F)\,] - \mathbb{E}[\, h(N)\,] = \mathbb{E}\big[\, g_h'(F) - Fg_h(F)\,\big]. \tag{11.7}$$

Thus, if $\mathcal{H}$ is a separating collection of Borel measurable functions $h : \mathbb{R} \to \mathbb{R}$ we deduce

$$\operatorname{dist}_{\mathcal{H}}(F, N) = \sup_{h\in\mathcal{H}} \Big| \mathbb{E}\big[g_h'(F) - F g_h(F) \big] \Big|. \tag{11.8}$$

We want to use the above equality to produce estimates on the Wasserstein distance between two Borel probability measures on $\mathbb{R}$.

Proposition 11.7. *Let $h : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function. Set $K := \operatorname{Lip}(h)$. Then the function g_h given by (11.5) admits the representation*

$$g_h(x) = -\int_0^\infty \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \mathbb{E}\big[h\big(e^{-t}x + \sqrt{1-e^{-2t}}N \big) N \big] dt. \tag{11.9}$$

Moreover, g_h is a C^1 function and

$$\|g_h'\|_\infty \leq \sqrt{\frac{2}{\pi}} K. \tag{11.10}$$

Almost a proof. We know that $-\delta_x g_h = h$ and we conclude that

$$-\partial_x \delta_x g_h = \partial_x h \ \text{ a.e. on } \mathbb{R}.$$

Using the Heisenberg identity $[\partial_x, \delta_x] = \mathbb{1}$ we deduce $-\partial_x\delta_x = -\mathbb{1} - \delta_x\partial_x = (L - \mathbb{1})$. Thus

$$(L - \mathbb{1}) g_h = \partial_x h.$$

From the condition (11.6) we deduce[11] $g_h \in L^2(\mathbb{R}, \boldsymbol{\gamma}_1)$. Clearly $\partial_x h \in L^2(\mathbb{R}, \boldsymbol{\gamma}_1)$. Hence

$$g_h = (L-1)^{-1}\partial_x h \overset{(9.4)}{=} -\int_0^\infty e^{-t} T_t[\partial_x h] dt.$$

Using Mehler's formula (9.1) we deduce

$$T_t[\partial_x h](x) = \int_{\mathbb{R}} h'(e^{-t}x + \sqrt{1-e^{-2t}}y) \boldsymbol{\gamma}(dy).$$

We set $u_x := e^{-t}x + \sqrt{1-e^{-2t}}y$ and we observe that for fixed x we have

$$\frac{d}{dy} h(u_x) = h'(u_x) \frac{du_x}{dy} = \sqrt{1-e^{-2t}}\, h'(u_x) \Rightarrow h'(u_x) = \frac{1}{\sqrt{1-e^{-2t}}} \frac{d}{dy} h(u_x).$$

Hence

$$\begin{aligned} T_t[\partial_x h](x) &= \frac{1}{\sqrt{1-e^{-2t}}} \int_{\mathbb{R}} \frac{d}{dy} h(u_x) \boldsymbol{\gamma}(dy) = \frac{1}{\sqrt{1-e^{-2t}}} \int_{\mathbb{R}} h(u_x) y\, \boldsymbol{\gamma}(dy) \\ &= \frac{1}{\sqrt{1-e^{-2t}}} \mathbb{E}\big[h\big(e^{-t}x + \sqrt{1-e^{-2t}}N \big) N \big]. \end{aligned}$$

This proves (11.9).

Clearly g_h is a C^1-function. To prove the estimate (11.10), we derivate (11.9) we respect to x and we deduce

$$g_h'(x) = -\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}} \mathbb{E}\big[h'\big(e^{-t}x + \sqrt{1-e^{-2t}}N \big) N \big] dt.$$

Since $|h'| \leq K$ we deduce

$$\big| g_h'(x) \big| \leq K\mathbb{E}[\,|N|\,] \int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}} dt = K\sqrt{\frac{2}{\pi}} \int_0^1 \frac{dv}{2\sqrt{1-v}} = K\sqrt{\frac{2}{\pi}}.$$

□

[11] This needs a proof, but we skip it since it is neither hard, nor particularly revealing.

From the above proposition and the equality (11.8) we obtain immediately the following useful result.

Corollary 11.8. *Let $N \sim \mathcal{N}(0,1)$. Then for any square integrable random variable F we have*

$$\operatorname{dist}_{FM}(F,N) \leq \operatorname{dist}_W(F,N) \leq \sup_{g\in\mathcal{F}_W} \Big| \mathbb{E}\big[\, g'(F) - Fg(F)\,\big] \Big|, \tag{11.11}$$

where

$$\mathcal{F}_W := \Big\{ g \in C^1(\mathbb{R});\;\; \|g'\|_\infty \leq \sqrt{\frac{2}{\pi}} \Big\}. \tag{11.12}$$

□

11.3. **The multidimensional Stein method.** The Stein method has a multidimensional counterpart. To describe it we need to introduce some notation. Denote by $\mathcal{L}(\mathbb{R}^n)$ the space of bounded linear operators $\mathbb{R}^n \to \mathbb{R}^n$. We define the *Hilbert-Schmidt* inner product on $L(\mathbb{R}^n)$ to be

$$(A,B)_{HS} := \operatorname{tr} AB^* = \sum_{i,j} A_{ij}B_{ij}, \;\; \forall A,B \in \mathcal{L}(\mathbb{R}^n).$$

The next result generalizes the one-dimensional Stein lemma

Lemma 11.9 (Multidimensional Stein lemma)**.** *Let $d \in \mathbb{N}$ and $C \in L(\mathbb{R}^d)$ be a symmetric operator such that $C \geq 0$. Let $\boldsymbol{N} = (N_1,\dots,N_d)$ be a random d-dimensional vector. Then the following statements are equivalent.*

(i) $\boldsymbol{N} \sim \mathcal{N}(0,C)$

(ii) *For any C^2 function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded first and second order derivatives we have*

$$\mathbb{E}\big[\, \big(\boldsymbol{N}, \nabla f(\boldsymbol{N})\,\big)\,\big] = \mathbb{E}\big[\, \big(\, C, \operatorname{Hess} f(\boldsymbol{N})\,\big)_{HS}\,\big]. \tag{11.13}$$

Proof. (i) ⇒ (ii). If $C > 0$, then the implication follows from an immediate integration by parts and the equality

$$\boldsymbol{\Gamma}_C(dx) = \frac{1}{\sqrt{\det(2\pi C)}} e^{-\frac{1}{2}(C\boldsymbol{x},\boldsymbol{x})} d\boldsymbol{x}.$$

The general case follows from the general case applied to the nondegenerate matrices $C_\varepsilon = C + \varepsilon\mathbb{1}$ and then (carefully) letting $\varepsilon \to 0$.

(ii) ⇒ (i). Fix $\boldsymbol{G} \sim \mathcal{N}(0,C)$ independent of $\boldsymbol{N}$ and a C^2 function $f : \mathbb{R}^d \to \mathbb{R}$ as in (ii). We set

$$\varphi(t) := \mathbb{E}\big[\, f\big(\sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{G}\,\big)\,\big].$$

Then

$$\varphi(1) = \mathbb{E}[\, f(\boldsymbol{N})\,], \;\; \varphi(0) = \mathbb{E}[\, f(\boldsymbol{G})\,]$$

and thus

$$\mathbb{E}[\, f(\boldsymbol{N})\,] - \mathbb{E}[\, f(\boldsymbol{G})\,] = \int_0^1 \varphi'(t)dt$$

$$= \int_0^1 \mathbb{E}\big[\, \big(\, \nabla f(\,\sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{G}\,), \boldsymbol{N}\,\big)\,\big] \frac{dt}{2\sqrt{t}} - \int_0^1 \mathbb{E}\big[\, \big(\, \nabla f(\,\sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{G}\,), \boldsymbol{G}\,\big)\,\big] \frac{dt}{2\sqrt{1-t}}$$

Using (11.13) we deduce by conditioning on $\boldsymbol{G}$ that, for any $x \in \mathbb{R}^d$, we have

$$\mathbb{E}\big[\, \big(\, \nabla f(\,\sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{x}\,), \boldsymbol{N}\,\big)\,\big] = \underbrace{\sqrt{t}\mathbb{E}\big[\, \big(\, C, \operatorname{Hess} f(\sqrt{t}\boldsymbol{N} + \sqrt{1-t}\boldsymbol{x}\,\big)_{HS}\,\big]}_{=:h_1(\boldsymbol{x},t)}.$$

Since $\boldsymbol{G} \sim \mathcal{N}(0, C)$ it satisfies (11.13) and, conditioning on $\boldsymbol{N}$, we deduce that for any $x \in \mathbb{R}^d$, we have

$$\mathbb{E}\big[\,\big(\nabla f(\sqrt{t}\boldsymbol{x} + \sqrt{1-t}\boldsymbol{G}\,), \boldsymbol{G}\,\big)\,\big] = \underbrace{\mathbb{E}\big[\,\big(\,C, \mathrm{Hess}\, f(\sqrt{t}\boldsymbol{x} + \sqrt{1-t}\boldsymbol{G})\,\big)_{HS}\,\big]}_{=:h_2(\boldsymbol{x},t)}.$$

Integrating $h_1(\boldsymbol{x}, t)$ and $h_2(\boldsymbol{x}, t)$ respectively with respect to the law of $\boldsymbol{G}$ and the law of $\boldsymbol{N}$, and then integrating with respect to t we deduce that

$$\boldsymbol{E}\big[\,f(\boldsymbol{N})\,\big] = \boldsymbol{E}\big[\,f(\boldsymbol{G})\,\big],$$

for any C^2-function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded first and second order derivatives. Since the class of such functions is separating we deduce that $\boldsymbol{N} \sim \boldsymbol{G} \sim \mathcal{N}(0, C)$. □

Definition 11.10. Let $\boldsymbol{N} \sim \mathcal{N}(0, \mathbb{1}_d)$ and $h : \mathbb{R}^n \to \mathbb{R}$ a measurable function such that $\boldsymbol{e}[\,|h(\boldsymbol{N})|\,] < \infty$. The *Stein's equation* associated to h and $\boldsymbol{N}$ is the p.d.e.

$$Lf(\boldsymbol{x}) = -\Delta f(\boldsymbol{x}) - \boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) = h(\boldsymbol{x}) - \mathbb{E}[\,h(\boldsymbol{N})\,], \;\; \Delta = -\sum_{j=1}^{d} \partial_{x_j}^2. \tag{11.14}$$

Observe that if $h : \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz continuous function, then the function

$$h_\perp(\boldsymbol{x}) := h(\boldsymbol{x}) - \mathbb{E}[\,h(\boldsymbol{N})\,] \in L^2(\mathbb{R}^d, \boldsymbol{\Gamma}) \;\text{ and }\; \int_{\mathbb{R}^d} h_\perp(\boldsymbol{x})\boldsymbol{\Gamma}(d\boldsymbol{x}) = 0.$$

Thus, $h_\perp$ lies in the range of the Ornstein-Uhlenbeck operator $L : \mathbb{D}^{2,2}(\mathbb{R}^d) \to \mathbb{D}^{0,2}(\mathbb{R}^d)$ so there exists a unique function $f_h \in \mathbb{D}^{2,2}(\mathbb{R}^d)$ such that

$$Lf_h(\boldsymbol{x}) = h_\perp(\boldsymbol{x}) \;\text{ and }\; \int_{\mathbb{R}^d} f_h(\boldsymbol{x})\boldsymbol{\Gamma}(d\boldsymbol{x}) = 0.$$

More precisely, $f_h = L^{-1}h = L^{-1}h_\perp$. We can now state the multidimensional counterpart of Proposition 11.7.

Proposition 11.11. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous function. Then the function*

$$f_h = L^{-1}h = L^{-1}h_\perp$$

is well defined, C^2 and admits the representation

$$f_h(\boldsymbol{x}) = -\int_0^\infty T_t[h_\perp]dt = \int_0^\infty \mathbb{E}\big[\,h(\boldsymbol{N}) - h\big(\,e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big)\,\big]dt. \tag{11.15}$$

Moreover, if $\mathrm{Lip}(h) \leq K$ *then,*

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d} \|\, \mathrm{Hess}\, f_h(\boldsymbol{x})\,\|_{HS} \leq K\sqrt{d}. \tag{11.16}$$

Proof. Let $h_n \in L^2(\mathbb{R}^d, \boldsymbol{\Gamma})$ be the n-th chaos component of $h(x)$. Then, in L^2, we have the following equalities

$$h(x) = \sum_{n\geq 0} h_n(x), \;\; h_\perp(x) = \sum_{n\geq 1} h_n(x),$$

$$L^{-1}h_\perp(x) = -\sum_{n\geq 1} \frac{1}{n} h_n(x) = -\sum_{n\geq 1}\int_0^\infty e^{-nt}h_n(x) = -\int_0^\infty T_t[h_\perp]dt.$$

This proves the first part of (11.15). The second part of this equality follows from Mehler's formula. The C^2-regularity of f_h is a consequence of basic elliptic regularity results.

To prove (11.16) we observe that

$$\partial^2_{x_ix_j} f_h(\boldsymbol{x}) = -\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbb{E}\Big[\,\partial_{x_j} h\big(\,e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big) N_i\,\Big] dt$$

Thus, if $B \in \mathcal{L}(\mathbb{R}^d)$, we have

$$\big|\,\big(\,B, \operatorname{Hess} f_h(\boldsymbol{x})\,\big)_{HS}\,\big| = \left|\sum_{i,j} \partial^2_{x_ix_j} f_h(\boldsymbol{x})\right|$$

$$= \left|\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbb{E}\Big[\,\Big(\,B\boldsymbol{N}, \nabla h\big(\,e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big)\,\Big)\,\Big] dt\,\right| dt$$

$$\leq \|\nabla h\|_\infty \mathbb{E}\Big[\,\big|B\boldsymbol{N}\,\big|_{\mathbb{R}^d}\,\Big]\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}} dt \leq K\sqrt{d}\,\sqrt{\mathbb{E}\Big[\,\big|B\boldsymbol{N}\,\big|^2_{\mathbb{R}^d}\,\Big]},$$

because $\|\nabla h\|_\infty \leq K\sqrt{d}$ and

$$\int_0^\infty \frac{e^{-2t}}{\sqrt{1-e^{-2t}}} dt = 1.$$

A simple computation shows that

$$\mathbb{E}\Big[\,\big|B\boldsymbol{N}\,\big|^2_{\mathbb{R}^d}\,\Big] = \|B\|^2_{HS}.$$

This completes the proof of (11.16). □

Proposition 11.11 admits the following immediate generalization.

Proposition 11.12. *Fix a symmetric positive definite operator $C \in \mathcal{L}(\mathbb{R}^d)$. Denote by $\lambda_{\min}(C)$ and respectively $\lambda_{\max}(C)$ the smallest and the largest eigenvalue of C. Fix a random vector $\boldsymbol{N} \sim \mathcal{N}(0, C)$ and a Lipschitz continuous function $h : \mathbb{R}^d \to \mathbb{R}$. Set $K := \operatorname{Lip}(h)$. Then the function*

$$f_h(\boldsymbol{x}) = \int_0^\infty \mathbb{E}\big[\,h(\boldsymbol{N}) - h\big(\,e^{-t}\boldsymbol{x} + \sqrt{1-e^{-2t}}\boldsymbol{N}\,\big)\,\big] dt \tag{11.17}$$

is well defined, it is C^2 and satisfies the Stein's equation

$$\Big(\,C, \operatorname{Hess} f(\boldsymbol{x})\,\Big)_{HS} - \big(\boldsymbol{x}, \nabla f(\boldsymbol{x})\,\big) = h(\boldsymbol{x}) - h(\boldsymbol{N}). \tag{11.18}$$

Moreover

$$\sup_{x\in\mathbb{R}^d}\big\|\operatorname{Hess} f_h(\boldsymbol{x})\,\big\|_{HS} \leq K\frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)}. \tag{11.19}$$

Main Idea. The above proposition can be obtained from Proposition 11.11 by choosing an orthonormal basis $\boldsymbol{f}_1, \dots, \boldsymbol{f}_d$ of $\mathbb{R}^d$ that diagonalizes C,

$$C\boldsymbol{f}_k = \lambda_k \boldsymbol{f}_k,\;\; k = 1, \dots, d,\;\; 0 < \lambda_1 \leq \cdots \leq \lambda_d.$$

□

The last result implies the following multi-dimensional counterpart of Corollary 11.8.

Corollary 11.13. *Fix a symmetric positive definite operator $C \in \mathcal{L}(\mathbb{R}^d)$ and a random vector $\boldsymbol{N} \sim \mathcal{N}(0, C)$. If F is a square integrable $\mathbb{R}^d$-valued random variable, then*

$$\operatorname{dist}_{FM}(F, \boldsymbol{N}) \leq \operatorname{dist}_W(F, \boldsymbol{N}) \leq \sup_{f\in\mathcal{F}_d}\left|\,\mathbb{E}\Big[\,\big(\,C, \operatorname{Hess} f(F)\,\big)_{HS} - \big(F, \nabla f(F)\,\big)\,\Big]\,\right|, \tag{11.20}$$

where $\mathcal{F}_d$ consists of the C^2-functions $f : \mathbb{R}^d \to \mathbb{R}$ satisfying (11.19) with $K = 1$. □

12. LIMIT THEOREMS

12.1. **An abstract limit theorem.** Fix a separable Gaussian Hilbert space $\mathscr{X} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$. As usual, we set $\widehat{\mathscr{X}} = L^2(\Omega, \widehat{\mathcal{F}}, \mathbb{P})$ and we denote by Proj_n the orthogonal projection onto the n-th chaos $\mathscr{X}^{:n:}$. For any number $N \in \mathbb{N}_0$ we set

$$\operatorname{Proj}_{\leq N} = \bigoplus_{0\leq n\leq N} \operatorname{Proj}_n, \;\; \operatorname{Proj}_{>N} = \mathbb{1} - \operatorname{Proj}_{\leq N}.$$

For $F \in \widehat{\mathscr{X}}$ and $n \in \mathbb{N}_0$ we denote by $\operatorname{Var}_n(F)$ the variance of $\operatorname{Proj}_n(F)$.

We have

$$\operatorname{Var}(F) = \sum_{n\geq 1} \operatorname{Var}_n(F),$$

and we set

$$\operatorname{Var}_{\leq N} := \sum_{n=1}^{N} \operatorname{Var}_n(F), \;\; \operatorname{Var}_{>N}(F) = \sum_{n>N} \operatorname{Var}_n(F) = \operatorname{Var}(F) - \operatorname{Var}_{\leq N}(F).$$

We begin by describing a simple sufficient condition guaranteeing the convergence in law to a normal random variable of a sequence of random variables in $\widehat{\mathscr{X}}$.

Proposition 12.1. *Consider a sequence of random variables $(F_\nu)_{\nu\geq 1}$ in $\widehat{\mathscr{X}}$ such that*

$$\mathbb{E}[F_\nu] = 0, \forall \nu,$$

i.e., $\operatorname{Proj}_0(F_\nu) = 0$, $\forall \nu$. Suppose that the following hold.

- (C_1) *For any $n \in \mathbb{N}$, the sequence of variances $\operatorname{Var}_n(F_\nu)$ converges as $\nu \to \infty$ to a nonnegative number v_n.*
- (C_2) *The sequence*
$$V_N := \sup_{\nu\geq 1} \operatorname{Var}_{>N}(F_\nu)$$
converges to 0 as $N \to \infty$. In other words, as $N \to \infty$, the "tails" $\operatorname{Proj}_{>N} F_\nu$ converge to 0 in L^2, uniformly *with respect to ν.*
- (C_3) *For any $N > 0$ the sequence of random variables $\operatorname{Proj}_{\leq N}(F_\nu)$ converges in law to a normal random variable.*

Then the following hold.

- (i) *The series $\sum_{n\geq 1} v_\nu$ is convergent. We denote by v its sum.*
- (ii)
$$\lim_{\nu\to\infty} \operatorname{Var}(F_\nu) = v.$$
- (iii) *As $\nu \to \infty$, the random variable F_ν converges in law to a random variable $F_\infty \sim \mathcal{N}(0, v)$.*

Proof. (i) Fix $\varepsilon > 0$. We can find $N(\varepsilon) > 0$ such that for any $N > N(\varepsilon)$ we have $V_N < \varepsilon$. For all $n > m > N(\varepsilon)$ we have

$$\sum_{k=m}^{n} \operatorname{Var}_k(F_\nu) \leq \sum_{k>N} \operatorname{Var}_k(F_\nu) \leq V_N < \varepsilon$$

which shows that

$$\forall n > M > N(\varepsilon) : \;\; \sum_{k=m}^{n} v_k = \lim_{\nu\to\infty} \sum_{k=m}^{n} \operatorname{Var}_k(F_\nu) \leq \varepsilon.$$

To prove (ii) observe that for any $N>0$ we have

$$
\begin{aligned}
|\operatorname{Var}(F_\nu)-v| &\leq \sum_{n\leq N}\Big|\operatorname{Var}_n(F_\nu)-v_n\Big| + \sum_{n>N}\operatorname{Var}_n(F_\nu)+\sum_{n>N}v_n\\
&\leq \sum_{n\leq N}\Big|\operatorname{Var}_n(F_\nu)-v_n\Big| + V_N+\sum_{n>N}v_n
\end{aligned}
$$

This proves that

$$\limsup_{\nu\to\infty}|\operatorname{Var}(F_\nu)-v|\leq V_N+\sum_{n>N}v_n,\ \ \forall N>0.$$

The conclusion (ii) is obtained by letting $N\to\infty$ in the above inequality.

(iii) Let $X\in\mathscr{X}$, $\|X\|=1$, so that $X\in\mathcal{N}(0,1)$, $\sqrt{v}X\in\mathcal{N}(0,v)$. We will show that for any bounded Lipschitz function $h:\mathbb{R}\to\mathbb{R}$ we have

$$\lim_{\nu\to\infty}\mathbb{E}\big[\,h(F_\nu)\,\big]=\mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]. \tag{12.1}$$

Observe that if $v=0$, we deduce from (ii) that $F_\nu\to 0$ in L^2 so F_ν converges in law to the degenerate normal random variable with variance 0. Assume $v>0$. Without loss of generality we can assume $v=1$.

Fix a bounded Lipschitz function $h:\mathbb{R}\to\mathbb{R}$ and set

$$K:=\|h\|_\infty+\operatorname{Lip}(h).$$

For $N>0$ we set

$$G_{\nu,N}=\operatorname{Proj}_{\leq N}(F_\nu),\ \ H_{\nu,N}=F_\nu-G_{\nu,N}=\operatorname{Proj}_{>N}(F_\nu)$$

$$v_N=\sum_{n\leq N}v_n,\ \ \sigma_N=\sqrt{v_N}$$

so that, as $\nu\to\infty$ $G_{\nu,N}$ converges in law to σ_NX and $H_{\nu,N}$ converges in L^2 to 0.

$$
\begin{aligned}
\Big|\mathbb{E}\big[\,h(F_\nu)\,\big]-\mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]\,\Big| &\leq \Big|\mathbb{E}\big[\,h(F_\nu)\,\big]-\mathbb{E}\big[\,h(G_{\nu,N})\,\big]\,\Big|\\
+\Big|\mathbb{E}\big[\,h(G_{\nu,N})\,\big]-\mathbb{E}\big[\,h(\sigma_NX)\,\big]\,\Big| &+\Big|\mathbb{E}\big[\,h(\sigma_NX)\,\big]-\mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big].\,\Big|
\end{aligned}
$$

Now observe that

$$\Big|\mathbb{E}\big[\,\big|\,h(G_{\nu,N}+H_{\nu,N})\ -\ h(G_{\nu,N})\,\big|\,\big]\,\Big|\leq K\mathbb{E}\big[\,\big|H_{\nu,N}\,\big|\,\big]\leq K\big\|H_{\nu,N}\|_{L^2},$$

$$\lim_{\nu\infty}\big\|H_{\nu,N}\|_{L^2}=0,\ \ \lim_{\nu\to\infty}\Big|\mathbb{E}\big[\,h(G_{\nu,N})\,\big]-\mathbb{E}\big[\,h(\sigma_NX)\,\big]\,\Big|=0,$$

so that

$$\limsup_{\nu\to\infty}\Big|\mathbb{E}\big[\,h(F_\nu)\,\big]-\mathbb{E}\big[\,h(X)\,\big]\,\Big|\leq\Big|\mathbb{E}\big[\,h(\sigma_NX)\,\big]-\mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]\,\Big|.$$

Letting $N\to\infty$ we deduce

$$\lim_{\nu\to\infty}\Big|\mathbb{E}\big[\,h(F_\nu)\,\big]-\mathbb{E}\big[\,h(\sqrt{v}\,X)\,\big]\,\Big|=0,$$

for any bounded Lipschitz function h. This proves (iii).

$\square$

In the remainder of this section we will explain how to combine the Stein method with the Malliavin calculus to prove central limit results of the type described in Proposition 12.1, with condition C_3 replaced by one that is easier to verify in concrete situations. These techniques were pioneered by D. Nualart and G. Peccati in [29] and have since generated a lot of follow-up investigations[12]; see e.g. [25, 27] and the references therein. We follow the presentation in the recent, award winning monograph of I. Nourdin and G. Peccati, [26].

12.2. **Central limit theorem: single chaos.** The following proposition is the key result in the implementation of the Stein method in the Wiener chaos context.

Proposition 12.2 (Key abstract estimate). *Let $F \in \mathbb{D}^{1,2}(\mathscr{X})$ such that $\mathbb{E}[F] = 0$, $\mathbb{E}[F^2] = 1$. If $g : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function and $K = \mathrm{Lip}(f)$, then*

$$\boxed{\left|\,\mathbb{E}\big[\,g'(F)\,\big] - \mathbb{E}\big[\,Fg(F)\,\big]\,\right| \leq K \cdot \left|\,\mathbb{E}\Big[\,\Big(1 - \big(\,DF, -DL^{-1}F\big)_{\mathscr{X}}\,\Big)\Big]\,\right|}. \tag{12.2}$$

Proof. Note first that g' is defined only a.e.. However, the law of F has a density, so the random variable $g'(F)$ is a.s. well defined. Using the integration-by-parts formula (9.5) with $F = G$ we deduce

$$\begin{aligned}\left|\,\mathbb{E}\big[\,g'(F)\,\big] - \mathbb{E}\big[\,Fg(F)\,\big]\,\right| &= \left|\,\mathbb{E}\Big[\,g'(F)\Big(1 - \big(\,DF, -DL^{-1}F\big)_{\mathscr{X}}\,\Big)\Big]\,\right| \\ &\leq K \cdot \left|\,\mathbb{E}\Big[\,\Big(1 - \big(\,DF, -DL^{-1}F\big)_{\mathscr{X}}\,\Big)\Big]\,\right|.\end{aligned}$$

□

Corollary 12.3. *Let $F \in \mathbb{D}^{1,2}(\mathscr{X})$ with $\mathbb{E}[F] = 0$, $\mathbb{E}[F^2] = \sigma^2 > 0$. If $N \sim \mathcal{N}(0, \sigma^2)$, then*

$$\mathrm{dist}_W(F, N) \leq \sqrt{\frac{2}{\pi\sigma^2}}\;\mathbb{E}\big[\,\big|\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\big|\,\big]. \tag{12.3}$$

If, in addition, $F \in \mathbb{D}^{1,4}(\mathscr{X})$, then

$$\mathbb{E}\Big[\,\Big|\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\Big|\,\Big] \leq \sqrt{\mathrm{Var}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\Big]}. \tag{12.4}$$

Proof. The case $\sigma = 1$ follows from Corollary 11.8 and the inequality (12.2). The general case of (12.3) follows from the case $\sigma = 1$ applied to the new random variable $\sigma^{-1}F$.

To prove (12.4) we observe that

$$\mathbb{E}\Big[\,\big|\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\big|\,\Big] \leq \sqrt{\mathbb{E}\big[\,\big(\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\big)^2\,\big]}.$$

From the integration by parts formula (9.5) we deduce that

$$\mathbb{E}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\Big] = \sigma^2,$$

so that,

$$\mathbb{E}\Big[\,\big(\,\sigma^2 - \big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\big)^2\,\Big] = \mathrm{Var}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\Big].$$

To show that the above variance is finite observe that

$$\mathbb{E}\Big[\,\big(DF, -DL^{-1}F\,\big)^2_{\mathscr{X}}\,\Big] \leq \sqrt{\mathbb{E}\Big[\,\|DF\|^4_{\mathscr{X}}\,\Big]} \cdot \sqrt{\mathbb{E}\Big[\,\|DL^{-1}F\|^4_{\mathscr{X}}\,\Big]}.$$

The Kree-Meyer inequalities (10.6) imply that the quantities in the right-hand-side above are finite. □

[12]Ivan Nourdin maintains a site dedicated to this novel way of approaching limit theorems https://sites.google.com/site/malliavinstein/home.

Remark 12.4. The method of proof of Proposition 12.2 and the statement of Corollary 12.3 rely on the assumption $\sigma > 0$ which may not be easy to verify in some concrete situations. □

Proposition 12.5. *Let $F \in \mathbb{D}^{1,2}(\mathscr{X})$ such that $\mathbb{E}[F] = 0$, $\mathbb{E}[F^2] = \sigma^2$. If $h \in C^2_b(\mathbb{R})$ and $N \sim \mathcal{N}(0, \sigma^2)$, then*

$$\big|\, \mathbb{E}\big[\, h(F)\,\big] - \mathbb{E}\big[\, h(N)\,\big]\,\big| \leq \frac{1}{2}\|h''\|_\infty \cdot \mathbb{E}\Big[\,\Big|\, \big(\, DF, -DL^{-1}F\big)_{\mathscr{X}}\,\big) - \sigma^2\,\Big|\,\Big]. \tag{12.5}$$

In particular, if $F \in \mathbb{D}^{1,4}$, then

$$\mathrm{dist}_{C^2}(F, N) \leq \frac{1}{2}\mathbb{E}\Big[\,\Big|\, \big(\, DF, -DL^{-1}F\big)_{\mathscr{X}}\,\big) - \sigma^2\,\Big|\,\Big] \leq \frac{1}{2}\sqrt{\mathrm{Var}\Big[\,\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\Big]}. \tag{12.6}$$

Proof. The results is obviously true if $\sigma^2 = 0$ so we can assume that $\sigma^2 > 0$. We set

$$\varphi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathbb{E}\big[\, h\big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\big)\,\big]dx.$$

Note that

$$\varphi(\infty) = \mathbb{E}\big[\, h(F)\,\big],\ \ \varphi(0) = \mathbb{E}\big[\, h(N)\,\big],$$

so that

$$\mathbb{E}\big[\, h(F)\,\big] - \mathbb{E}\big[\, h(N)\,\big] = \int_0^\infty \varphi'(t)dt.$$

We have

$$\varphi'(t) = \frac{e^{-t}\sigma}{\sqrt{2\pi}}\mathbb{E}\left[\,\int_{-\infty}^{\infty} h'\big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\big) x e^{-\frac{x^2}{2}}dx\,\right]$$

$$+ \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \times \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F\mathbb{E}\Big[\, h'\big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\big)\,\Big] e^{-\frac{x^2}{2}}dx.$$

Performing an usual integration by parts in the first integral and using the Malliavin integration by parts formula (9.5) in the second integrand we deduce

$$\varphi'(t) = -\frac{e^{-2t}\sigma^2}{\sqrt{2\pi}}\mathbb{E}\left[\,\int_{-\infty}^{\infty} h''\big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\big) e^{-\frac{x^2}{2}}dx\,\right]$$

$$+ \frac{e^{-2t}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathbb{E}\Big[\, h''\big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\big)\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}}\,\Big] e^{-\frac{x^2}{2}}dx$$

$$= \frac{e^{-2t}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathbb{E}\Big[\, h''\Big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\Big) \cdot \Big(\,\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}} - \sigma^2\,\Big)\,\Big] e^{-\frac{x^2}{2}}dx.$$

We deduce

$$\mathbb{E}\big[\, h(F)\,\big] - \mathbb{E}\big[\, h(N)\,\big]$$

$$= \int_{-\infty}^{\infty} \frac{e^{-2t}dt}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathbb{E}\Big[\, h''\Big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\Big) \cdot \Big(\,\big(DF, -DL^{-1}F\,\big)_{\mathscr{X}} - \sigma^2\,\Big)\,\Big] e^{-\frac{x^2}{2}}dx.$$

We reach the desired conclusion by observing that

$$\mathbb{E}\Big[\,\Big|\, h''\Big(\, e^{-t}\sigma x + \sqrt{1 - e^{-2t}}F\,\Big)\,\Big|\,\Big] \leq \|h''\|_\infty,\ \ \forall x.$$

□

Observe that when $F \in \mathscr{X}^{:q:}$, $q > 0$, then $F \in \mathbb{D}^{1,4}$ and

$$\big(DF, -DL^{-1}F\big)_{\mathscr{X}} = \frac{1}{q}\|DF\|^2_{\mathscr{X}}.$$

In this case we can provide more detailed information. This will require a bit of Ito calculus and a bit more terminology.

Given $p, q \in \mathbb{N}$ and $r \in \mathbb{N}_0$ such that $r \leq \min\{p, q\}$ we define the map

$$\otimes_r : \mathscr{X}^{\otimes p} \times \mathscr{X}^{\otimes q} \to \mathscr{X}^{\otimes(p+q-2r)}$$

to be the unique continuous bilinear map such that

$$(X_1 \otimes \cdots \otimes X_p) \otimes_r (Y_1 \otimes \cdots \otimes Y_q) = \left(\prod_{j=1}^{r} \mathbb{E}[X_j Y_j]\right) X_{r+1} \otimes \cdots \otimes X_p \otimes Y_{t+1} \otimes \cdots \otimes Y_q.$$

We define

$$\tilde{\otimes}_r : \mathscr{X}^{\odot p} \times \mathscr{X}^{\odot q} \to \mathscr{X}^{\odot(p+q-2r)}$$

to be

$$u \tilde{\otimes}_r v := \mathbf{Sym}\big[\, u \otimes_r v \,\big], \;\; \forall u \in \mathscr{X}^{\odot p}, \;\; v \in \mathscr{X}^{\odot q}.$$

Remark 12.6. If $W : L^2(M, \mathcal{M}, \mu) \to \mathscr{X}$ is a white noise isomorphism then we can *isometrically* identify $\mathscr{X}^{\otimes p}$ with the space $L^2(M^p, \mathcal{M}^{\otimes p}, \mu^{\otimes p})$. Thus we can view $f \in \mathscr{X}^{\otimes p}$ and $g \in \mathscr{X}^{\otimes q}$ as L^2-functions

$$f : M^p \to \mathbb{R}, \quad g : M^q \to \mathbb{R}.$$

Then $f \otimes_r g$ can be identified with the function $M^{p-r} \times M^{q-r} \to \mathbb{R}$ given by

$$f \otimes_r g(x_{r+1}, \dots, x_p, y_{r+1}, \dots, y_q)$$

$$= \int_{M^r} f(t_1, \dots, t_r, x_{r+1}, \dots, x_p) g(t_1, \dots, t_r, y_{r+1}, \dots y_q) \mu^{\otimes r}(dt_1 \cdots dt_r).$$

Lemma 12.7. *Let $q \in \mathbb{N}$, $q \geq 2$ and $f \in \mathscr{X}^{\odot q}$, Set $F = \boldsymbol{I}_p[f]$. Then the following hold.*

$$\frac{1}{q}\|\, DF\|^2_{\mathscr{X}} = \mathbb{E}[F^2] + q \sum_{r=1}^{q-1} (r-1)! \binom{q-1}{r-1}^2 \boldsymbol{I}_{2q-2r}[f \tilde{\otimes}_r f], \tag{12.7a}$$

$$\operatorname{Var}\left(\frac{1}{q}\|\, DF\|^2_{\mathscr{X}}\right) = \frac{1}{q^2} \sum_{r=1}^{q-1} r^2 (r!)^2 \binom{q}{r}^2 (2q-2r)! \|f \tilde{\otimes}_r f\|^2_{\mathscr{X}^{\odot(2q-2r)}}, \tag{12.7b}$$

$$\begin{aligned} \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2 &= \frac{3}{q} \sum_{r=1}^{q-1} r^2 (r!)^2 \binom{q}{r}^2 (2q-2r)! \,\|\, f \,\tilde{\otimes}_r f \,\|^2_{\mathscr{X}^{\otimes(2q-2r)}} \\ = \sum_{r=1}^{q-1} (q!)^2 \binom{q}{r}^2 &\left(\|\, f \otimes_r f\|^2_{\mathscr{X}^{\otimes(2q-2r)}} + \binom{2q-2r}{q-r} \|\, f \,\tilde{\otimes}_r f \,\|^2_{\mathscr{X}^{\otimes(2q-2r)}} \right), \end{aligned} \tag{12.7c}$$

$$\boxed{\operatorname{Var}\left(\frac{1}{q}\|\, DF\|^2_{\mathscr{X}}\right) \leq \frac{q-1}{3q}\Big(\mathbb{E}\big[F^4\big] - 3\mathbb{E}\big[F^2\big]^2\Big) \leq (q-1)\operatorname{Var}\left(\frac{1}{q}\|\, DF\|^2_{\mathscr{X}}\right)}. \tag{12.7d}$$

About the proof. Let us point out that (12.7b) follows immediately from (12.7a) via the isometry (5.4). The inequality (12.7d) follows immediately from (12.7b, 12.7c). Thus it suffices to prove only (12.7a) and (12.7c).

To prove (12.7a) it is convenient to consider a more general problem, that of finding the chaos decomposition of

$$\big(DF, DG\big)_{\mathscr{X}},\;\; F, G \in \mathscr{X}^{:q:}$$

We write $F = \boldsymbol{I}_p[f]$, $G = \boldsymbol{I}_p[g]$, $f, g \in \mathscr{X}^{\odot q}$. Using the polarization trick we can reduce the problem to the special case

$$f = X^{\otimes q},\;\; g = Y^{\otimes q},\;\; X, Y \in \mathscr{X},\;\; \mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1.$$

Thus

$$F = H_q(X),\;\; DF = qH_{q-1}(X)DX,$$

$$G = H_q(Y),\;\; DG = qH_{q-1}(Y)DY,$$

$$\big(DF, DG\big)_{\mathscr{X}} = q^2 H_{q-1}(X)H_{q-1}(Y)\mathbb{E}[XY].$$

The equality (12.7a) now follows by invoking (4.5), (4.6) and the isometry equality (8.5).

The proof of (12.7c) requires a bit more work. The hardest part is the 2nd half of this equality. It is based on the (non-obvious) elementary identity

$$\boxed{(2q)!\|f\tilde{\otimes}f\|^2_{\mathscr{X}^{\otimes(2q)}} = 2(q!)^2\|f\|^4_{\mathscr{X}^{\otimes q}} + (q!)^2\sum_{r=1}^{q-1}\binom{q}{r}^2\|f\otimes_r f\|^2_{\mathscr{X}^{\otimes(2q-2r)}},\;\; f \in \mathscr{X}^{\odot q}}. \tag{12.8}$$

A convenient way to prove this is to use a white noise isomorphism as in Remark 12.6. We refer to [26, Lemma 5.2.4] for details. □

Corollary 12.8 (The fourth moment theorem, [29])**.** *Suppose that $F \in \mathscr{X}^{:q:}$, $q \geq 2$ and $\mathbb{E}[F^2] = \sigma^2 > 0$. Then for $N \in \mathcal{N}(0, \sigma)$ we have*

$$\operatorname{dist}_W(F, N) \leq \frac{1}{\sigma}\sqrt{\operatorname{Var}\Big(\frac{2}{q\pi}\|DF\|^2_{\mathscr{X}}\Big)} \leq \frac{1}{\sigma}\sqrt{\frac{(2q-2)\big(\mathbb{E}[F^4] - 3\sigma^4\big)}{3\pi q}}. \tag{12.9}$$

Thus, given a sequence $(F_n)_{n\geq 0}$ in $\mathscr{X}^{:q:}$, $q \geq 2$ and $N \sim \mathcal{N}(0, \sigma)$ the following statements are equivalent.

(i) *The sequence $(F_n)_{n\geq 0}$ converges in probability to N.*
(ii) *As $n \to \infty$, $\mathbb{E}[F_n^2] \to \mathbb{E}[N^2] = \sigma^2$and $\mathbb{E}[F_n^4] \to \mathbb{E}[N^4] = 3\sigma^4$.*
(iii) *If $F_n = \boldsymbol{I}_q[f_n]$, $f_n \in \mathscr{X}^{\odot q}$, then*

$$\lim_{n\to\infty}\|f_n\tilde{\otimes}_r f_n\|_{\mathscr{X}^{\odot(2q-2r)}} = 0,\;\; \forall r = 1\dots, q-1.$$

(iv) $\operatorname{Var}\big(\|DF_n\|^2\big) \to 0$ *as* $n \to \infty$.

Proof. In this case we have

$$\big(DF, -DL^{-1}F\big)_{\mathscr{X}} = \frac{1}{q}\|\,DF\,\|^2_{\mathscr{X}}.$$

The desired conclusions follow from Corollary 12.3, (12.7b) and (12.7d). □

12.3. **Central limit theorem: multiple chaoses.** The results proved in the previous subsection have a multidimensional counterpart. The next result, is the multi-dimensional counterpart of Proposition 12.2 and Corollary 12.3

Proposition 12.9. *Fix $d \geq 2$ and let $\boldsymbol{F} = (F_1, \dots, F_d)$ be a random vector such that $F_1, \dots, F_d \in \mathbb{D}^{1,4}(\mathscr{X})$ with $\mathbb{E}[F_i] = 0$, i. Let $C \in \mathcal{L}(\mathbb{R}^d)$ be a symmetric* positive *definite operator and let $\boldsymbol{N} \sim \mathcal{N}(0, C)$. Then*

$$\operatorname{dist}_W(\boldsymbol{F}, \boldsymbol{N}) \leq \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)} \sqrt{\sum_{i,j=1}^{d} \mathbb{E}\Big[\big(C_{ij} - \big(DF_i, -DL^{-1}F_j \big)_{\mathscr{X}} \big)^2 \Big]} \tag{12.10}$$

Proof. Let M be the random operator $M : \Omega \to \mathcal{L}(\mathbb{R}^d)$ with the (i, j)-th entry given by

$$M_{ij} := \big(DF_j, -DL^{-1}F_i \big)_{\mathscr{X}}.$$

Arguing as in the proof of Corollary 12.3 we deduce that $M_{ij} \in L^2$ since $F_i, F_j \in \mathbb{D}^{1,4}(\mathscr{X})$. For $g \in C^2(\mathbb{R}^d)$ such that

$$\sup_{x \in \mathbb{R}^d} \big\| \operatorname{Hess} g(x) \big\|_{HS} \leq \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)}$$

we have

$$\Big| \mathbb{E}\big[(C, \operatorname{Hess} g)_{HS}(\boldsymbol{F}) - (\boldsymbol{F}, \nabla g(\boldsymbol{F}))_{\mathbb{R}^d} \big] \Big| = \left| \sum_{i,j=1}^{d} C_{ij} \mathbb{E}\big[\partial^2_{x_i x_j} g(\boldsymbol{F}) \big] - \sum_{i=1}^{d} \mathbb{E}\big[F_i \partial_{x_i} g(\boldsymbol{F}) \big] \right|$$

(use the integration by parts formula (9.5))

$$= \left| \sum_{i,j=1}^{d} C_{ij} \mathbb{E}\big[\partial^2_{x_i x_j} g(\boldsymbol{F}) \big] - \sum_{i,j=1}^{d} \mathbb{E}\big[\partial^2_{x_i x_j} g(\boldsymbol{F}) (DF_j, -DL^{-1}F_i)_{\mathscr{X}} \big] \right|$$

$$= \left| \sum_{i,j=1}^{d} C_{ij} \mathbb{E}\big[\partial^2_{x_i x_j} g(F) \big(C_{ij} - (DF_j, -DL^{-1}F_i)_{\mathscr{X}} \big] \right|$$

$$= \Big| \mathbb{E}\big[(\operatorname{Hess} g(\boldsymbol{F}), C - M)_{HS} \big] \Big| \leq \sqrt{\mathbb{E}\big[\| \operatorname{Hess} g(\boldsymbol{F}) \|^2_{HS}} \cdot \sqrt{\mathbb{E}\big[\| C - M \|^2_{HS}}$$

$$\leq \frac{\sqrt{d\lambda_{\max}(C)}}{\lambda_{\min}(C)} \sqrt{\mathbb{E}\big[\| C - M \|^2_{HS}}.$$

We conclude by invoking Corollary 11.13. □

The next result, is the multidimensional counterpart of Proposition 12.5 and explains what to do when the covariance matrix C is possible degenerate.

Proposition 12.10. *Fix $d \geq 2$ and let $\boldsymbol{F} = (F_1, \dots, F_d)$ be a random vector such that $F_1, \dots, F_d \in \mathbb{D}^{1,4}(\mathscr{X})$ with $\mathbb{E}[F_i] = 0$, i. Let $C \in \mathcal{L}(\mathbb{R}^d)$ be a symmetric, nonnegative definite operator and let $\boldsymbol{N} \sim \mathcal{N}(0, C)$. Then for every $h \in C^2(\mathbb{R}^d)$ such that $\|h''\|_\infty < \infty$ we have*

$$\big| \mathbb{E}\big[h(\boldsymbol{F}) \big] - \mathbb{E}\big[h(\boldsymbol{N}) \big] \big| \leq \frac{1}{2} \|h''\|_\infty \sqrt{\sum_{i,j=1}^{d} \mathbb{E}\Big[\big(C_{ij} - \big(DF_j, -DL^{-1}F_i \big)_{\mathscr{X}} \big)^2 \Big]} \tag{12.11}$$

Proof. Without any loss of generality we can assume $\boldsymbol{N}$ is independent of the Gaussian space $\mathscr{X}$. Let h as in the statement of the proposition. For $t \in [0,1]$ we set

$$\Psi(t) := \mathbb{E}\big[\, h(\sqrt{1-t}\,\boldsymbol{F} + \sqrt{t}\,\boldsymbol{N})\,\big].$$

Then

$$\boldsymbol{E}\big[\, h(\boldsymbol{N})\,\big] - \boldsymbol{E}\big[\, h(\boldsymbol{F})\,\big] = \Psi(1) - \Psi(0) = \int_0^1 \Psi'(t)dt.$$

We have

$$\Psi'(t) = \sum_{i=1}^d \mathbb{E}\left[\, \partial_{x_i} h(\sqrt{1-t}\,\boldsymbol{F} + \sqrt{t}\,\boldsymbol{N})\, \left(\frac{1}{2\sqrt{t}}N_i - \frac{1}{2\sqrt{1-t}}F_i\right)\right].$$

At this point we want to use the following elementary but useful identity.

Lemma 12.11. *If $f = f(y_1,\dots,y_d) : \mathbb{R}^d \to \mathbb{R}$ is C^1 with bounded derivatives, $\widehat{N} \sim \mathcal{N}(0, \mathbb{1}_d)$ and $T, S \in \mathcal{L}(\mathbb{R}^d)$, then*

$$\mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big)\big(\, T\hat{\boldsymbol{N}}\,\big)_i\,\big] = \sum_{k=1}^d \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big)(TS^*)_{ik}\,\big], \tag{12.12}$$

where $\big(\, T\hat{\boldsymbol{N}}\,\big)_i$ denotes the i-th component of the random vector $T\hat{\boldsymbol{N}}$ and $(TS^*)_{ik}$ denote the (i,k)-entry of the matrix TS^*

Proof of the lemma. We have

$$\mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big)\big(\, T\hat{\boldsymbol{N}}\,\big)_i\,\big] = \sum_{j=1}^d \mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big)T_{ij}N_j\big]$$

$(\delta_j = -\partial_{N_j} + N_j)$

$$= \sum_{j=1}^d \mathbb{E}\big[\, f\big(\, S\hat{\boldsymbol{N}}\,\big)T_{ij}\delta_j(1)\big]$$

(integrate by parts using the equalities $\partial_{N_j} = \sum_k \partial_{y_k}\partial_{N_j} y_k$, $y_k = \sum_j S_{kj}N_j$)

$$= \sum_{j=1}^d\sum_{k=1}^d \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big)S_{kj}T_{ij}\big] = \sum_{k=1}^d \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big)(TS^*)_{ik}\,\big].$$

□

Now observe that if $f : \mathbb{R}^d \to \mathbb{R}$ is a C^1-function with bounded derivatives, and $\hat{\boldsymbol{N}} \sim \mathcal{N}(0, \mathbb{1}_d)$ is such that, $\boldsymbol{N} = \sqrt{C}\hat{\boldsymbol{N}}$, then (12.12) shows that

$$\mathbb{E}\big[\, f(\boldsymbol{N})N_i\,\big] = \mathbb{E}\big[\, f(\,\sqrt{C}\hat{\boldsymbol{N}}\,)\big(\sqrt{C}\hat{\boldsymbol{N}}\big)_i\,\big] = \sum_{k=1}^d \mathbb{E}\big[\, \partial_{y_k} f\big(\, S\hat{\boldsymbol{N}}\,\big)C_{ik}\,\big]. \tag{12.13}$$

We have

$$\mathbb{E}\Big[\, \partial_{x_i} h(\sqrt{1-t}\,\boldsymbol{F} + \sqrt{t}\,\boldsymbol{N})N_i\Big] = \mathbb{E}_{\boldsymbol{x}}\Big[\, \mathbb{E}\Big[\, \partial_{x_i} h(\sqrt{1-t}\,\boldsymbol{x} + \sqrt{t}\,\boldsymbol{N})N_i\Big|\boldsymbol{F} = \boldsymbol{x}\Big]\,\Big]$$

$$\stackrel{(12.13)}{=} \sqrt{t}\sum_{j=1}^d \mathbb{E}_{\boldsymbol{x}}\Big[\, C_{ij}\mathbb{E}\Big[\, \partial^2_{x_i x_j} h(\sqrt{1-t}\,\boldsymbol{x} + \sqrt{t}\,\boldsymbol{N})\Big|\boldsymbol{F} = \boldsymbol{x}\Big]\,\Big]$$

$$= \sqrt{t}\sum_j C_{ij}\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})\,\Big].$$

Using the integration by parts formula (9.5) we deduce

$$\mathbb{E}\Big[\,\partial_{x_i}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})F_i\Big] = \mathbb{E}_{\boldsymbol{x}}\Big[\,\mathbb{E}\Big[\,\partial_{x_i}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{x})F_i\Big|\boldsymbol{N}=\boldsymbol{x}\Big]\,\Big]$$

$$= \sqrt{1-t}\sum_{j=1}^d\mathbb{E}_{\boldsymbol{x}}\Big[\,\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{x})\big(DF_j,-DL^{-1}F_i\big)_{\mathscr{X}}\Big|\boldsymbol{N}=\boldsymbol{x}\Big]\,\Big]$$

$$= \sqrt{1-t}\sum_{j=1}^d\mathbb{E}\Big[\,\partial^2_{x_ix_j}h(\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N})\big(DF_j,-DL^{-1}F_i\big)_{\mathscr{X}}\Big]$$

Hence

$$\boldsymbol{E}\big[\,h(\boldsymbol{N})\,\big]-\boldsymbol{E}\big[\,h(\boldsymbol{F})\,\big]=\Psi(1)-\Psi(0)=\int_0^1\Psi'(t)dt$$
$$=\frac{1}{2}\sum_{i,j=1}^d\int_0^1\mathbb{E}\Big[\,\partial^2_{x_ix_j}h\big(\,\sqrt{1-t}\,\boldsymbol{F}+\sqrt{t}\,\boldsymbol{N}\,\big)\Big(\,C_{ij}-\big(\,DF_j,-DL^{-1}F_i\,\big)_{\mathscr{X}}\,\Big)\,\Big]. \tag{12.14}$$

□

We now have (almost) all the information we need to prove the following remarkable result.

Theorem 12.12. *Let $d\geq 2$ and $q_1,\dots,q_d\in\mathbb{N}$. Consider the d-dimensional random vector*

$$\boldsymbol{F}=(F_1,\dots,F_d),\quad F_i\in\mathscr{X}^{:q_i:},\quad i=1,\dots,d.$$

Let $f_i\in\mathscr{X}^{\odot q_i}$ such that $\boldsymbol{I}_{q_i}[f_i]=F_i$. Denote by C the covariance matrix of the random vector $\boldsymbol{F}$, $C_{ij}=\mathbb{E}[F_iF_j]$, and let $\boldsymbol{N}\sim\mathcal{N}(0,C)$. Consider the continuous function

$$\psi:(\mathbb{R}\times\mathbb{R}_{>0})^d\to\mathbb{R}>0$$

given by

$$\Psi(x_1,y_1,\dots,x_d,y_d)=\sum_{i,j=1}^d\delta_{q_iq_j}\left(\sqrt{\sum_{r=1}^{q_i-1}\binom{2r}{r}}\right)|x_i|^{\frac{1}{2}}$$
$$+\sum_{i,j=1}^d(1-\delta_{q_iq_j})\left(\big(\,2|y_j|\,\big)^{\frac{1}{2}}\,|x_i|^{\frac{1}{4}}++\sum_{r=1}^{\min(q_i,q_j)-1}\sqrt{(2(q_i+q_j-2r)!}\binom{q_j}{r}|x_i|^{\frac{1}{2}}\right),$$

and set

$$m(\boldsymbol{F})=\psi\big(\,m_4(F_1)-3m_2(F_1)^2,m_2(F_1),\dots,m_4(F_d)-3m_2(F_d)^2,m_2(F_d)\,\big),$$

where we recall that $m_k(X)$ denotes the k-th moment of a random variable X. Note that

$$\psi\big(\,x_1,y_1,\dots,x_d,y_d\,\big)_{x_1=\cdots=x_d=0}=0.$$

If $h:\mathbb{R}^d\to\mathbb{R}$ is a C^2 function with bounded second derivatives, then

$$\big|\,\mathbb{E}[\boldsymbol{F}]-\mathbb{E}[\boldsymbol{N}]\,\big|\leq\frac{1}{2}\|h''\|_\infty m(\boldsymbol{F}).$$

The main ideas. We plan to use Proposition 12.10 so we need to estimate from above the quantities

$$\mathbb{E}\Big[\,\Big(\,C_{ij}-\big(\,DF_j,-DL^{-1}F_i\big)_{\mathscr{X}}\,\Big)^2\,\Big]=\mathbb{E}\Big[\,\Big(\,\mathbb{E}[F_iF_j]-\frac{1}{q_i}\big(\,DF_i,DF_j\big)_{\mathscr{X}}\,\Big)^2\,\Big].$$

Note that $C_{ij}=0$ if $q_i\neq q_j$. Thus, we need to produce suitable upper estimates for quantities of the form

$$\mathbb{E}\Big[\,\alpha-\frac{1}{p}(DF,DG)_{\mathscr{X}}\,\Big],\;\;F\in\mathscr{X}^{:p:},\;\;G\in\mathscr{X}^{:q:},\;\;\alpha\in\mathbb{R}.$$

This is what the next lemma accomplishes.

Lemma 12.13. *Let $F=\boldsymbol{I}_p[f]$, $f\in\mathscr{X}^{\odot p}$ and $G=\boldsymbol{I}_q[g]$, $g\in\mathscr{X}^{\odot q}$, $p,q\geq 1$. Suppose that α is a real constant.*

(i) *If $p=q$, then*

$$\mathbb{E}\Big[\,\Big(\,\alpha-\frac{1}{p}(DF,DG)_{\mathscr{X}}\,\Big)^2\,\Big]\leq\big(\,\alpha-\mathbb{E}[FG]\,\big)^2$$
$$+\frac{p^2}{2}\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^4(2p-2r)!\big(\,\|f\otimes_{p-r}f\|^2_{\mathscr{X}^{\otimes 2r}}+\|g\otimes_{p-r}g\|^2_{\mathscr{X}^{\otimes 2r}}\,\big)\tag{12.15}$$

(ii) *If $p<q$, then*

$$\mathbb{E}\Big[\,\Big(\,\frac{1}{q}(DF,DG)_{\mathscr{X}}\,\Big)^2\,\Big]\leq(p!)^2\binom{q-1}{p-1}^2(q-p)!\|f\|^2_{\mathscr{X}^{\otimes p}}\|\,g\otimes_{q-p}g\,\|_{\mathscr{X}^{\otimes 2p}}$$
$$+\frac{p^2}{2}\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^2\binom{q-1}{r-1}^2(p+q-2r)!\big(\,\|f\otimes_{p-r}f\|^2_{\mathscr{X}^{\otimes 2r}}+\|g\otimes_{q-r}g\|^2_{\mathscr{X}^{\otimes 2r}}\,\big).\tag{12.16}$$

Main idea of the proof. The lemma follows from the identity

$$\big(DF,DG\,\big)_{\mathscr{X}}=pq\sum_{r=1}^{\min(p,q)}(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1}\boldsymbol{I}_{p+q-2r}[f\otimes_r g],$$

which can be reduced to the equality (4.5). □

Using (12.7c) we deduce that for any $q\geq 2$ and any $f\in\mathscr{X}^{\odot q}$ we have

$$\|f\otimes_r f\|^2_{\mathscr{X}^{\otimes(2q-2r)}}\leq\frac{\big(\,r!(q-r)!\,\big)^2}{(q!)^4}\Big(\,\mathbb{E}\big[\,\boldsymbol{I}_q[\,f\,]^4\,\big]-3\mathbb{E}\big[\,\boldsymbol{I}_q[\,f\,]^2\,\big]^2\,\Big).$$

Theorem 12.12 now follows from the above lemma after some simple algebraic manipulations □

Theorem 12.12 implies the following remarkable result.

Theorem 12.14 (Peccati-Tudor, [30]). *Let $d\geq 1$ and $q_1,\dots,q_d\in\mathbb{N}$. Consider the sequence of d-dimensional random vectors*

$$\boldsymbol{F}_n=(F_{1,n},\dots,F_{d,n}),\;\;F_{j,n}\in\mathscr{X}^{:q_j:},\;\;j=1,\dots,d,\;\;n\in\mathbb{N}.$$

Suppose that $C\in\mathcal{L}(\mathbb{R}^d)$ is symmetric and nonnegative definite and

$$\lim_{n\to\infty}\mathbb{E}\big[\,F_{i,n}F_{j,n}\,\big]=C_{ij},\;\;\forall i,j=1,\dots,d.$$

Then the following statements are equivalent.

(i) *The random vector $\boldsymbol{F}_n$ converges in probability to a Gaussian vector $\boldsymbol{N}\sim\mathcal{N}(0,C)$.*

(ii) *For each* $j = 1, \dots, d$ *the sequence of random variables* $\big(F_{i,n} \big)_{n\in\mathbb{N}}$ *converges in probability to a Gaussian r.v.* $N_i \sim \mathcal{N}(0, C_{ii})$.

□

The above result leads to the following substantial strengthening of Proposition 12.1

Theorem 12.15. *Consider a sequence of random variables* $(F_\nu)_{\nu\geq 1}$ *in* $\widehat{\mathscr{X}}$ *such that* $\mathbb{E}[F_\nu] = 0$, $\forall\nu$, *i.e.,* $\mathrm{Proj}_0(F_\nu) = 0$, $\forall\nu$. *Suppose that the following hold.*

(C_1) *For any* $k \in \mathbb{N}$, $\exists v_k \geq 0$ *such that*

$$\lim_{\nu\to\infty} \mathbb{E}\Big[\Big(\mathrm{Proj}_k F_\nu \Big)^2 \Big] = v_k.$$

(C_2) *The sequence*

$$V_N := \sup_{\nu\geq 1} \sum_{k>N} \mathbb{E}\Big[\Big(\mathrm{Proj}_k F_\nu \Big)^2 \Big]$$

converges to 0 *as* $N \to \infty$.

(C_3') *For any* $k \in \mathbb{N}$

$$\lim_{\nu\to\infty} \mathbb{E}\Big[\Big(\mathrm{Proj}_k F_\nu \Big)^4 \Big] = 3v_k^2.$$

Then the following hold.

(i) *The series* $\sum_{n\geq 1} v_\nu$ *is convergent. We denote by* v *its sum.*

(ii)

$$\lim_{\nu\to\infty} \mathrm{Var}(F_\nu) = v.$$

(iii) *As* $\nu \to \infty$, *the random variable* F_ν *converges in law to a random variable* $F_\infty \sim \mathcal{N}(0, v)$.

Remark 12.16. (a) The fourth moment theorem (Corollary 12.8) shows that the conditions $C_1 + C_3'$ are equivalent with the requirement that, $\forall k \in \mathbb{N}$, as $\nu \to \infty$ the random variables $\mathrm{Proj}_k[F_\nu]$ converge in probability as $\nu \to \infty$ to a normal random variable $N_k \in \mathcal{N}(0, v_k)$.

(b) If we write

$$\mathrm{Proj}_k[F_\nu] = \boldsymbol{I}_k[f_{\nu,k}], \;\; f_{\nu,k} \in \mathscr{X}^{:k:},$$

the Corollary 12.8 shows that the condition C_3'' is equivalent to

$$\lim_{\nu\to\infty} \| f_{\nu,k} \tilde{\otimes}_r f_{\nu,k} \|_{\mathscr{X}^{\odot(2q-2r)}} = 0, \;\; \forall k \geq 1, \;\; \forall r = 1 \dots, k-1$$

□

References

[1] R. J Adler, J. E. Taylor: *Random Fields and Geometry*, Springer Verlag, 2007.
[2] S. Bochner: *Harmonic Analysis and the Theory of Probability*, Dover, 2005.
[3] V. I. Bogachev: *Gaussian Measures*, Math. Surveys and Monographs, vol. 62., Amer. Math. Soc., 1998.
[4] P. Breuer, P. Major: *Central limit theorems for non-linear functionals of Gaussian fields*, J. of Multivariate Anal., **13**(1983), 425-441.
[5] S. Chatterjee: *A short survey of Stein's method*, 2014 ICM talk, arXiv: 1404.1392.
[6] L.H.Y. Chen, L. Goldstein, Q,-M, Shao: *Normal Approximations by Stein's Method* Springer Verlag, 2011.
[7] E.Çinlar: *Probability and Stochastics*, Grad. Texts in Math., vol. 261 Springer Verlag, 2011.
[8] G. Da Prato: *An Introduction to Infinite Dimensional Analysis*, Springer Verlag, 2006.
[9] R.M. Dudley: *Real Analysis and Probability*, Cambridge University Press, 2004.
[10] A. Estrade, J. R. León: *A central limit theorem for the Euler characteristic of a Gaussian excursion set*, Ann. of Probability, to appear. MAP5 2014-05. 2015. hal-00943054v3.
[11] I.M. Gelfand, N.Ya. Vilenkin: *Generalized Functions*, vol. 4, Academic Press, New York, 1964.
[12] S. Janson: *Gaussian Hilbert Spaces*, Cambridge Tracts in Math., vol. 129, Cambridge University Press, 1997.

[13] I. Karatzas, S.E.Shreve: *Brownian Motion and Stochastic Calculus*, 2nd Edition, Grad. Texts in Math., Springer Verlag, 1991.
[14] A. Klenke: *Probability Theory. A Comprehensive Course*, Universitext, Springer Verlag, 2008.
[15] M. Kratz, J. R. León: *Hermite polynomial expansion for non-smooth functionals of stationary Gaussian processes: crossings and extremes*, Stoch. Proc. Appl. **77**(1997), 237-252.
[16] M. Kratz, J. R. León: *Central limit theorems for level functionals of stationary Gaussian processes and fields*, J. Theor. Probab. **14**(2001), 639-672.
[17] H.-H. Kuo: *Introduction to Stochastic Integration*, Universitext, Springer Verlag, 2006.
[18] M. Ledoux: *The Concentration of Measure Phenomenon*, Math. Surveys and Monographs, vol. 89, Amer. Math. Soc., 2001.
[19] P. Major: *Multiple Wiener-Ito integrals*, Lect. Notes in Math., vol. 849, Springer Verlag, 1981.
[20] P. Malliavin: *Integration and Probability*, Grad. Texts. in Math., vol. 157, Springer Verlag, 1995.
[21] P. Malliavin: *Stochastic Analysis*, Grundlehren. der math. Wissenschaften, vol. 313, Springer Verlag, 1997.
[22] F. J. Murray, J. v. Neumann: *Rings of operators*, Ann. Math., **37**(1936), 116-229.
[23] L.I. Nicolaescu: *A CLT concerning critical points of random functions on a Euclidean space*, arXiv: 1509.06200.
[24] L.I. Nicolaescu: *Critical points of multidimensional random Fourier series: central limits*, arXiv: 1511.04965.
[25] I. Nourdin, G. Peccati: *Stein's method on Wiener chaos*, Prob. Theor. Rel. Fields, **145**(2009), 75-118.
[26] I. Nourdin, G. Peccati: *Normal Approximations with Malliavin Calculus. From Stein's Method to Universality*, Cambridge Tracts in Math., vol.192, Cambridge University Press, 2012.
[27] I. Nourdin, G. Peccati, M. Podolskij: *Quantitative Breuer-Major theorems*, Stoch. Proc. and their Appl., **121**(2011), 793-812.
[28] D. Nualart: *The Malliavin Calculus and Related Topics*, 2nd Edition, Springer Verlag, 2006.
[29] D. Nualart, G. Peccati: *Central limit theorems for sequences of multiple stochastic integrals*, Ann. Probab. **33**(2005), 177-193.
[30] G. Peccati, C.A. Tudor: *Gaussian limits for vector-valued multiple stochastic integrals*, *Séminaire de Probabilités XXXVIII*, p. 247-262, Lect. Notes in Math. vol. 1857, Springer Verlag, 2005.
[31] M. Reed, B. Simon: *Methods of Modern Mathematical Physics I. Functional Analysis*, Academic Press, 1980.
[32] N. Ross: *Fundamentals of Stein's method*, arXiv: 1109.1880.
[33] I. Shigekawa: *Derivatives of Wiener functionals and absolute continuity of induced measures*, J.Math. Kyoto Univ., **20**(1980), 263-289.
[34] D. W. Stroock: *Probability Theory. An Analytic View*, 2nd Edition, Cambridge University Press, 2011.
[35] D. W. Stroock: *Gaussian Measures in Finite and Infinite Dimensions*, Springer Verlag, 2023.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NOTRE DAME, NOTRE DAME, IN 46556-4618.
Email address: nicolaescu.1@nd.edu
URL: http://www.nd.edu/~lnicolae/