1. Calculus on Gaussian Spaces

Malliavin calculus is about calculus on Gaussian Sobolev Spaces. It is instructive to start with Gaussian measures on n-dimensional Euclidean space. We will quickly encounter the important operators of Malliavin calculus, the gradient, the divergence and the Ornstein-Uhlenbeck operator. Let $C_p^\infty(\mathbb{R}^n)$, be the set of all infinitely differentiable functions with polynomial growth on $\mathbb{R}^n$. We let ν^n denote the standard Gaussian measure on $\mathbb{R}^n$, i.e.,

$$d\nu^n(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} e^{-\langle \mathbf{x},\mathbf{x}\rangle/2} d\mathbf{x}.$$

Since polynomials are integrable with respect to Gaussian measures, $C_p^\infty(\mathbb{R}^n)$ is a suitable class of functions. In probabilistic terms this is equivalent to the fact that Gaussian random variables have finite moments of any order. Let $f \in C_p^\infty(\mathbb{R}^n)$ and $g = (g_1, \ldots, g_n)$ be a smooth vector field with $g_i \in C_p^\infty(\mathbb{R}^n)$, $i = 1, \ldots, n$. We use the notation ∂_i for partial derivatives, $\nabla = (\partial_1, \ldots, \partial_n)$ for the gradient and $\Delta = \sum_{i=1}^n \partial_i^2$ for the Laplace operator. The divergence is defined by $\nabla \cdot g = \sum_{i=1}^n \partial_i g_i$. What is the adjoint operator of ∇, acting on the vector field g under a Gaussian measure? Using integration by parts in one variable at a time we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^n} \nabla f(\mathbf{x}) \cdot g(\mathbf{x}) \, \mathrm{d}\nu^n(\mathbf{x}) \\
&\quad = \sum_{i=1}^n \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \partial_i f(x_1, \ldots, x_i, \ldots, x_n) g_i(x_1, \ldots, x_i, \ldots, x_n) e^{-x_i^2/2} \, \mathrm{d}x_i \\
&\qquad \times e^{-\sum_{k\neq i} x_k^2/2} \frac{dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n}{(2\pi)^{n/2}} \\
&\quad = \int_{\mathbb{R}^n} f(\mathbf{x}) \sum_{i=1}^n [x_i g_i(\mathbf{x}) - \partial_i g_i(\mathbf{x})] \, \mathrm{d}\nu^n(\mathbf{x}) \\
&\quad = \int_{\mathbb{R}^n} f(\mathbf{x}) [\mathbf{x} \cdot g(\mathbf{x}) - \nabla \cdot g(\mathbf{x})] \, \mathrm{d}\nu^n(\mathbf{x}) \\
&\quad := \int_{\mathbb{R}^n} f(\mathbf{x}) \delta^n g(\mathbf{x}) \, \mathrm{d}\nu^n(\mathbf{x}).
\end{aligned} \tag{1.1}$$

The operator

$$\delta^n g(\mathbf{x}) := \mathbf{x} \cdot g(\mathbf{x}) - \nabla \cdot g(\mathbf{x}). \tag{1.2}$$

is called the divergence and acts as the adjoint of ∇ under a standard Gaussian measure. Notice that in calculus w.r.t. Lebesgue measure λ^n we have that

$$\int_{\mathbb{R}^n} \nabla f(\mathbf{x}) \cdot g(\mathbf{x}) \, \mathrm{d}\lambda^n(\mathbf{x}) = - \int_{\mathbb{R}^n} f(\mathbf{x}) \nabla \cdot g(\mathbf{x}) \, \mathrm{d}\lambda^n(\mathbf{x}), \tag{1.3}$$

which explains the terminology. The minus is hidden in δ^n as is seen by comparison of (1.2) and (1.3).

If g is the gradient of some scalar field h, i.e., $g = \nabla h$, then we have

$$\begin{aligned}
&\int_{\mathbb{R}^n} \nabla f(\mathbf{x}) \cdot \nabla h(\mathbf{x}) \, \mathrm{d}\nu^n(\mathbf{x}) \\
&\quad = \int_{\mathbb{R}^n} f(\mathbf{x}) \sum_{i=1}^{n} [\partial_i h(\mathbf{x}) x_i - \partial_i^2 h(\mathbf{x})] \, \mathrm{d}\nu^n(\mathbf{x}) \\
&\quad = \int_{\mathbb{R}^n} f(\mathbf{x}) [\nabla h(\mathbf{x}) \cdot \mathbf{x} - \Delta h(\mathbf{x})] \, \mathrm{d}\nu^n(\mathbf{x}) \\
&\quad := - \int_{\mathbb{R}^n} f(\mathbf{x}) L^n h(\mathbf{x}) \, \mathrm{d}\nu^n(\mathbf{x}).
\end{aligned}$$

The operator

$$L^n h(\mathbf{x}) := \Delta h(\mathbf{x}) - \nabla h(\mathbf{x}) \cdot \mathbf{x}$$

is called the Ornstein-Uhlenbeck operator. Putting the pieces together we see that

$$L^n = -\delta^n \nabla.$$

The Ornstein-Uhlenbeck operator plays the same role on $L_2(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \nu^n)$ as the Laplace operator on $L_2(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \lambda^n)$, where λ^n is the n-dimensional Lebesgue measure. It is self adjoint, negative and densely defined. It generates an important semi-group called the Ornstein-Uhlenbeck semi-group.

It is possible to extend the operators to be well defined on a larger space of function, just as the case with λ^n, to Sobolev spaces. We omit that, and instead pass to the probabilistic case. There, in the situation of possibly countably many Gaussian random variables on some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we will develop the Sobolev theory.

2. Simplified Malliavin Calculus

The calculus above will here move into a probabilistic setting. Let the probability space $(\Omega, \mathcal{F}, \mathbb{P}) = (\mathbb{R}^n, \mathcal{B}^n, \nu^n)$, i.e, a Gaussian probability space admitting at most n independent random variables. We denote a point of Ω by ω. For every $h \in \mathbb{R}^n$ we define a random variable $I(h)(\omega) := \langle \omega, h \rangle_{\mathbb{R}^n}$. It is easy to check that $\mathbb{E}[I(h)] = 0$, for all $h \in \mathbb{R}^n$ and $\mathbb{E}[I(h)I(g)] = \langle h, g \rangle_{\mathbb{R}^n}$, for all $h, g \in \mathbb{R}^n$. Now, let $\{h_1, \ldots, h_n\} \subset \mathbb{R}^n$, be an orthonormal system. It is then clear that $I(h_1), \ldots, I(h_n)$ are independent Gaussian random variables. Define, for $f \in C_p^\infty(\mathbb{R}^n)$, the random variable $F = f(I(h_1), \ldots, I(h_n))$. We want to differentiate F with respect to the chance parameter ω. We denote the corresponding gradient D. Clearly

$$DI(h)(\omega) = D\langle \omega, h \rangle_{\mathbb{R}^n} = h,$$

and thus,

$$
\begin{aligned}
DF &= Df(I(h_1),\dots,I(h_n)) = \nabla f(I(h_1),\dots,I(h_n))[h_1,\dots,h_n]^* \\
&= \sum_{i=1}^n \partial_i f(I(h_1),\dots,I(h_n))h_i.
\end{aligned}
$$

Here $*$ denotes matrix transpose. The directional derivative of F in direction $h \in H$ is denoted $D^h F$. It is given by

$$
D^h F = \langle DF, h\rangle_{\mathbb{R}^n}.
$$

For $F = f(I(h_1),\dots,I(h_n))$ we have that

$$
\begin{aligned}
&\frac{\partial}{\partial\epsilon} f\big(\langle \omega - \epsilon h, h_1\rangle_{\mathbb{R}^n},\dots,\langle \omega - \epsilon h, h_n\rangle_{\mathbb{R}^n}\big)\Big|_{\epsilon=0} \\
&\quad = \frac{\partial}{\partial\epsilon} f\big(I(h_1) - \epsilon\langle h_1, h\rangle_{\mathbb{R}^n},\dots,I(h_n) - \epsilon\langle h_n, h\rangle_{\mathbb{R}^n}\big)\Big|_{\epsilon=0} \\
&\quad = \sum_{i=1}^n \partial_i f(I(h_1),\dots,I(h_n))\langle h_i, h\rangle_{\mathbb{R}^n} = D^h F.
\end{aligned}
$$

So, D^h seems to be reasonably defined.

Taking the second derivative yields

$$
D^2 F = \sum_{i,j=1}^n \partial_i\partial_j f(I(h_1),\dots,I(h_n)) h_i \otimes h_j,
$$

where the tensor product $h_i \otimes h_j$ is nothing but the outer product $h_i h_j^*$, resulting in an n by n matrix. It acts on a vector $\mathbf{x}$ by $(h_i \otimes h_j)\mathbf{x} = h_i\langle h_j, \mathbf{x}\rangle_{\mathbb{R}^n}$.

The differential calculus in the previous section moves over directly to this setting. Let $f \in C_p^\infty(\mathbb{R}_n)$, $h_1,\dots,h_n \in \mathbb{R}^n$ orthonormal and $F = f(I(h_1),\dots,I(h_n))$. Further, define the smooth random vector field $u = G_1 h_1 + \cdots + G_n h_n$ with $G_i = g_i(I(h_1),\dots,I(h_n))$ and $g_i \in C_p^\infty(\mathbb{R}^n)$. Here we have chosen to express u in terms of the basis $\{h_1,\dots,h_n\}$. Any other basis would do fine. We would then make a change of basis and get different random coefficients $G_1,\dots,G_n$. Working with two different bases would make the final expression ugly.

$$
\begin{aligned}
\mathbb{E}[DF \cdot u] &= \sum_{i,j=1}^{n} \mathbb{E}\Big[\partial_i f(I(h_1), \ldots, I(h_n)) g_j(I(h_1), \ldots, I(h_n))\Big] \langle h_i, h_j \rangle_{\mathbb{R}^n} \\
&= \int_{\mathbb{R}^n} \sum_{i=1}^{n} \partial_i f(\mathbf{x}) g_i(\mathbf{x}) \,\mathrm{d}\nu^n(\mathbf{x}) \\
&= \int_{\mathbb{R}^n} f(\mathbf{x}) \sum_{i=1}^{n} [g_i(\mathbf{x}) x_i - \partial_i g_i(\mathbf{x})] \,\mathrm{d}\nu^n(\mathbf{x}) \\
&= \mathbb{E}\Big[F \sum_{i=1}^{n} [G_i I(h_i) - D^{h_i} G_i]\Big] \\
&:= \mathbb{E}[F \delta u].
\end{aligned}
$$

Here we used that, by orthonormality,

$$
\partial_i g_i(\mathbf{x}) = \partial_i g_i(\mathbf{x}) \langle h_i, h_i \rangle_{\mathbb{R}^n} = \Big\langle \sum_{j=1}^{n} \partial_j g_i(\mathbf{x}) h_j, h_i \Big\rangle_{\mathbb{R}^n} = \langle DG_i, h_i \rangle_{\mathbb{R}^n} = D^{h_i} G_i.
$$

So, by the above calculation the divergence δ for $u = G_1 h_1 + \cdots + G_n h_n$ is given by

$$
\delta u = \sum_{i=1}^{n} [G_i I(h_i) - D^{h_i} G].
$$

Now to the Ornstein-Uhlenbeck operator. Let $f, g \in C_p^\infty(\mathbb{R}^n)$, $\{h_1, \cdots, h_n\} \subset \mathbb{R}^n$ orthonormal, $F = f(I(h_1), \ldots, I(h_n))$ and $G = g(I(h_1), \ldots, I(h_h))$.

$$
\begin{aligned}
\mathbb{E}[\langle DF, DG \rangle_{\mathbb{R}^n}] &= \sum_{i,j=1}^{n} \mathbb{E}\Big[\partial_i f(I(h_1), \ldots, I(h_n)) \partial_j g(I(h_1), \ldots, I(h_n))\Big] \langle h_i, h_j \rangle_{\mathbb{R}^n} \\
&= \int_{\mathbb{R}^n} \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x}) \,\mathrm{d}\nu^n(\mathbf{x}) \\
&= \int_{\mathbb{R}^n} f(\mathbf{x}) \sum_{i=1}^{n} [x_i \partial_i g(\mathbf{x}) - \partial_i^2 g(\mathbf{x})] \,\mathrm{d}\nu^2(\mathbf{x}) \\
&= \mathbb{E}\Big[F \sum_{i=1}^{n} [I(h_i) D^{h_i} G - \langle D^2 G h_i, h_i \rangle_{\mathbb{R}^n}]\Big] \\
&= \mathbb{E}\Big[F \Big[\sum_{i=1}^{n} I(h_i) D^{h_i} G - \mathrm{Tr}(D^2 G)\Big]\Big] \\
&:= \mathbb{E}[FLG]
\end{aligned}
$$

Here we used that, by orthogonality and the definition of the outer product,

$$\begin{aligned}
\partial_i^2 g(\mathbf{x}) &= \sum_{j,k=1}^{n} \partial_j \partial_k g(\mathbf{x}) \langle h_j, h_i \rangle_{\mathbb{R}^n} \langle h_i, h_k \rangle_{\mathbb{R}^n} \\
&= \Big\langle \sum_{j,k=1}^{n} \partial_j \partial_k g(\mathbf{x}) (h_j \otimes h_k) h_i), h_i \Big\rangle_{\mathbb{R}^n} \\
&= \langle D^2 G h_j, h_j \rangle_{\mathbb{R}^n}.
\end{aligned}$$

The Ornstein-Uhlenbeck operator is hence given by

$$LF = \sum_{i=1}^{n} I(h_i) D^{h_i} F - \mathrm{Tr}(D^2 F)$$

Those explicit expressions of δ and L will not be of big importance. We will take the closure of the above used random variables, in an appropriate norm. Naturally, the closure contains more general random variables. The calculations above gives us a motivation for the abstract framework to come. We will define families of Gaussian random variables $\{I(h) : h \in H\}$ for an arbitrary separable Hilbert space H.

3. Gaussian Hilbert Spaces

Let H be a separable Hilbert space and Ω its algebraic dual, i.e., the space of all linear but not necessarily bounded functionals $h \mapsto \omega(h) \in \mathbb{R}$ for $h \in H$. Define a family of functions $\{I(h) : h \in H\}$ on Ω by $\omega \mapsto I(h)(\omega) = \omega(h)$. Let $\mathcal{F}$ be the σ-algebra generated by the functions $\{I(h) : h \in H\}$. Itô has proven that there exists a unique probability measure $\mathbb{P}$ that makes $\{I(h) : h \in H\}$ satisfy the properties of the following definition.

Definition 3.1. *A collection $\{I(h)\}_{h \in H}$ of real random variables is called an isonormal Gaussian process on $(\Omega, \mathcal{F}, \mathbb{P})$ if*

- *$I(h)$ is centered Gaussian $\forall h \in H$,*
- *$\mathbb{E}[I(h)I(g)] = \langle h, g \rangle_H, \forall h, g \in H$.*

Such a family is called an isonormal Gaussian process. It offers us a set of linear functionals on $(\Omega, \mathcal{F}, \mathbb{P})$ suited for differential calculus. By definition rather than by calculation we will have $DI(h) = h$. The isometry enables us to easily create independent random variables via orthogonality in H. Notice that the family $\{I(h) : h \in H\} \subset L_2(\Omega)$ is a closed subspace and hence a Hilbert space. This follows from the fact that L_2-limits and hence limits in probability of Gaussian random variables are Gaussian together with the above properties. Such spaces are refereed to as Gaussian Hilbert spaces.

The Itô isometry offers a concrete example of an isonormal Gaussian process. Let $H = L_2([0,T])$, $\{W(t)\}_{t \in [0,T]}$ be a standard Wiener process. Define $I(h) = \int_0^T h(s)\,\mathrm{d}W(s)$, i.e. the Itô integral of a deterministic square integrable integrand.

It is well known that $I(h)$ is centered Gaussian. Moreover by the Itô isometry for $h, g \in H$

$$\mathbb{E}[I(h)I(g)] = \mathbb{E}\Big[\int_0^T h(t)\,\mathrm{d}W(t)\int_0^T g(t)\,\mathrm{d}W(t)\Big] = \int_0^T h(t)g(t)\,\mathrm{d}t = \langle h, g\rangle_H.$$

This is the Wiener integral defined on the measure space $([0,T], \mathcal{B}_{[0,T]}, \lambda_{[0,T]})$ where $\lambda_{[0,T]}$ is the Lebesgue measure on $[0,T]$. We will now construct the Wiener integral on a general measure space. This is not central for what to come and can be skipped, but it offers a wide range of examples of isonormal Gaussian processes.

Let $(X, \mathcal{G}, \nu)$ be a positive measure space. Define the function $C : \mathcal{G} \times \mathcal{G} \to [0,\infty], (\mathcal{A}, \mathcal{B}) \mapsto \nu(\mathcal{A} \cap \mathcal{B})$. It is clear that C is symmetric and it can be shown that it is positive definite. It is known that given any mean function $\mu : T \to \mathbb{R}$ and covariance function $Q : T \times T \to \mathbb{R}$, where T is some set, there exist a Gaussian process $\{G(t)\}_{t\in T}$ with $\mathbb{E}[G(t)] = \mu(t)$, $\forall t \in T$, $Cov(G(s), G(t)) = Q(s,t), \forall s, t \in T$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let now the index set be $T = \mathcal{G}$, the mean be $\mu = 0$ and the covariance be given by $Q = C$ defined before. By the discussion above there exists a Gaussian process $\{\dot{W}(\mathcal{A})\}_{\mathcal{A}\in\mathcal{G}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying

- $\mathbb{E}[\dot{W}(\mathcal{A})] = 0, \forall \mathcal{A} \in \mathcal{G}$,
- $\mathbb{E}[\dot{W}(\mathcal{A})\dot{W}(\mathcal{B})] = \nu(\mathcal{A} \cap \mathcal{B}), \forall \mathcal{A}, \mathcal{B} \in \mathcal{G}$.

Let h be a simple real valued function on X, i.e. $h(x) = \sum_{i=1}^n a_i \chi_{A_i}$ where $a_1, \ldots, a_n \in \mathbb{R}$ and $A_1, \ldots, A_n \in \mathcal{G}$ are disjoint. We define the Wiener integral $I(h)$ of h by

$$I(h) = \sum_{i=1}^n a_i \dot{W}(A_i).$$

Clearly $\mathbb{E}[I(h)] = 0$ and moreover

$$\|I(h)\|^2_{L_2(\Omega,\mathcal{F},\mathbb{P})} = \mathbb{E}[I(h)^2] = \sum_{i=1}^n a_i^2 \nu(A_i) = \|h\|^2_{L_2(X,\mathcal{G},\nu)}.$$

This is the so called Wiener isometry. Let $\{h_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence of simple functions in $L_2(X, \mathcal{G}, \nu)$ converging to $h \in L_2(X, \mathcal{G}, \nu)$. Then by the Wiener isometry $\{I(h_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $L_2(\Omega, \mathcal{F}, \mathbb{P})$. Since this is a complete space it converges. We denote the limit by $I(h)$. If we let $H = L_2(X, \mathcal{G}, \nu)$ then clearly we have that $\{I(h) : h \in H\}$ is an isonormal Gaussian process.

Example 3.2. For $(X, \mathcal{G}, \nu) = (\mathbb{R}^+, \mathcal{B}_{\mathbb{R}+}, \lambda)$ W is the Wiener (Itô) integral of square integrable functions on $[0,\infty)$. Moreover, we have that $I(\chi_{[0,t]}) = \beta(t)$, i.e. a standard Brownian motion. The Gaussian independent increments are immediate from the construction. To prove the almost sure continuity Kolmogorovs continuity criterion can easily be used.

4. The Malliavin Derivative

Here, we will leave all special cases behind us and make no assumptions on the topology of Ω. What we have is an isonormal Gaussian process $\{I(h) : h \in H\}$ indexed by a separable Hilbert space H and defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The σ-algebra $\mathcal{F}$ is assumed to be generated by $\{I(h) : h \in H\}$. The theory starts with the sort of smooth random variables we have seen already, namely the type $F = f(I(h_1), \ldots, I(h_n))$, where $n \in \mathbb{N}$, $f \in C_p^\infty(\mathbb{R}^n, \mathbb{R})$ and $\{h_1, \ldots, h_n\} \subset H$. We denote this class of smooth random variables $\mathcal{S}$. The fact that the probability space admits a countable number of independent r.v. does not affect the calculus. The vector $(I(h_1), \ldots, I(h_n))$ still has the standard normal distribution on $\mathbb{R}^n$. Further, define $\mathcal{S}_b \subset \mathcal{S}$ and $\mathcal{S}_0 \subset \mathcal{S}$ to be smooth random variables with $f \in C_b^\infty(\mathbb{R}^n, \mathbb{R})$ (bounded) and $f \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ (compact support) respectively. Clearly $\mathcal{S}_0 \subset \mathcal{S}_b \subset \mathcal{S}$ and $\mathcal{S}_0$ is dense in $L_2(\Omega)$.

Definition 4.1. *For $F = f(I(h_1), \ldots, I(h_n)) \in \mathcal{S}$ we define the Malliavin derivative of F to be the H-valued random variable*

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

Example 4.2. Let $\{\beta(t)\}_{t \in [0,T]}$ be a Brownian motion. We have that

$$D\beta(t) = D \int_0^T \chi_{[0,t]}(s) \, \mathrm{d}\beta(s) = DW(\chi_{[0,t]}) = \chi_{[0,t]}, t \in [0,T].$$

The directional derivative $\langle DF, h \rangle_H$ will be denoted $D^h F$. The next result is an integration by parts formula for smooth random variables.

Lemma 4.3. *Let $F \in \mathcal{S}$ and $h \in H$. Then*

$$\mathbb{E}[\langle DF, h \rangle_H] = \mathbb{E}[F I(h)].$$

Proof. We can without loss of generality assume for $n \in \mathbb{N}$, $f \in C_p^\infty(\mathbb{R}^n, \mathbb{R})$, $F = f(I(h_1), \ldots, I(h_n))$ that $h_1, \ldots, h_n \in H$ are orthonormal. Otherwise we can redefine f since W is linear. We have that $I(h_1), \ldots, I(h_n)$ are independent standard normal random variables. Now, for $h \in Span(h_1, \ldots, h_n) \subset H$, we integrate by parts as in (1.1).

$$
\begin{aligned}
\mathbb{E}[\langle DF, h\rangle_H] &= \sum_{i=1}^n \mathbb{E}[\partial_i f(I(h_1),\dots,I(h_n))]\langle h_i, h\rangle_H \\
&= \sum_{i=1}^n \int_{\mathbb{R}^n} \partial_i f(\mathbf{x}) e^{-\mathbf{x}\cdot\mathbf{x}/2} \frac{\mathrm{d}\mathbf{x}}{(2\pi)^{n/2}} \langle h_i, h\rangle_H \\
&= \sum_{i=1}^n \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \partial_i f(x_1,\dots,x_i,\dots,x_n) e^{-x_i^2/2}\,\mathrm{d}x_i \\
&\qquad\qquad \times e^{-\sum_{k\neq i} x_k^2/2} \frac{\mathrm{d}x_1\cdots\mathrm{d}x_{i-1}\,\mathrm{d}x_{i+1}\cdots\mathrm{d}x_n}{(2\pi)^{n/2}} \langle h_i, h\rangle_H \\
&= \sum_{i=1}^n \int_{\mathbb{R}^n} f(\mathbf{x}) x_i e^{-\mathbf{x}\cdot\mathbf{x}/2} \frac{\mathrm{d}\mathbf{x}}{(2\pi)^{n/2}} \langle h_i, h\rangle_H \\
&= \sum_{i=1}^n \mathbb{E}[FW(h_i)]\langle h_i, h\rangle_H \\
&= \mathbb{E}[FW(\sum_{i=1}^n \langle h_i, h\rangle_H h_i)] \\
&= \mathbb{E}[FI(h)]
\end{aligned}
$$

If $h \perp Span(h_1,\dots,h_n)$ then the left hand side is clearly zero. For the right hand side, F and $I(h)$ becomes independent, so $\mathbb{E}[FI(h)] = \mathbb{E}[F]\mathbb{E}[I(h)] = 0$. □

Remark 4.4. Notice that Lemma 4.3, written in non probabilistic terms, is equivalent to the following. For $F \in \mathcal{S}$ and $h \in H$ we have

$$\langle DF, h\rangle_{L_2(\Omega,H)} = \langle F, I(h)\rangle_{L_2(\Omega)}$$

It is clear that W is the adjoint of D in this smooth "semi-deterministic" setting. In section 10 we will be able to extend this to random h in a subset of $L_2(\Omega, H)$ and by the next proposition F in a larger class than $\mathcal{S}$.

Example 4.5. Let $H = L_2([0,T])$ and $\{\beta(t)\}_{t\in[0,T]}$ be a Brownian motion. Then Lemma 4.3 reads

$$\mathbb{E}\int_0^T D_s F h(s)\,\mathrm{d}s = \mathbb{E}F \int_0^T h(s)\,\mathrm{d}\beta(s).$$

We will now extend the derivative operator outside the class $\mathcal{S}$. Let $p \geq 1$. The Watanabe graph norm of $D : L^p(\Omega,\mathcal{F},\mathbb{P}) \to L^p(\Omega,\mathcal{F},\mathbb{P}; H)$ is defined by

$$\|F\|_{1,p} = \Big(\mathbb{E}[|F|^p] + \mathbb{E}[\|DF\|_H^p]\Big)^{1/p}. \tag{4.1}$$

We will make use of the following observation.

$$\mathbb{E}[G\langle DF, h\rangle_H] = \mathbb{E}[FGI(h)] - \mathbb{E}[F\langle DG, h\rangle_H]. \tag{4.2}$$

It follows by application of Lemma 4.3 to $\mathbb{E}[FGI(h)]$ for $F, G \in \mathcal{S}$, $h \in H$ and by using the product rule. The latter is an obvious consequence of the usual product rule.

PROPOSITION 4.6. *Let $p \geq 1$. The closure of $\mathcal{S}$ under the norm (4.1) is a Banach space and a dense subspace of $L_p(\Omega)$. It is called the Watanabe-Sobolev space and we denote it $\mathbb{D}^{1,p}$. The space $\mathbb{D}^{1,2}$ is a Hilbert space with inner product*

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

PROOF. We need to prove that D is a closable operator from $L_p(\Omega)$ to $L_p(\Omega, H)$. For this it suffices to prove that for a sequence $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{S}$ such that $F_n \to 0 \in L_p(\Omega)$ and $DF_n \to \eta \in L_p(\Omega)$ then $\eta = 0$ [1].

Fix $h \in H$ and let $G \in \mathcal{S}$ such that G and $GI(h)$ are bounded. Such random variables are dense in $\mathcal{S}_b$ and hence in $L_2(\Omega)$. To see this let $\tilde{G} \in \mathcal{S}_b$, $\tilde{G}_n = \tilde{G}e^{-\epsilon_n I(h)^2}$ for $\epsilon_n > 0$, $n = 1, 2, \ldots$ and $\epsilon_n \to 0$. Then $\tilde{G}_n I(h)$ are bounded $n = 1, 2, \ldots$ and $\tilde{G}_n \to \tilde{G}$ in $L_2(\Omega)$. Let now $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{S}$ such that $F_n \to 0 \in L_p(\Omega)$ and $DF_n \to \eta \in L_p(\Omega)$. Then, using bounded convergence and (4.2) we get,

$$\begin{aligned}\mathbb{E}[G\langle \eta, h\rangle_H] &= \lim_{n\to\infty} \mathbb{E}[G\langle DF_n, h\rangle_H] \\ &= \lim_{n\to\infty} \mathbb{E}[F_n GI(h)] - \mathbb{E}[F_n\langle DG, h\rangle_H] = 0\end{aligned}$$

since $F_n \to 0$ in $L_p(\Omega) \subset L_1(\Omega)$ and $GI(h)$ and $\langle DG, h\rangle_H$ are bounded. This implies that $\langle \eta, h\rangle_H = 0$ with probability one. The same result can be obtained for any $h \in H$ and hence $\eta = 0$ almost surely. ☐

The following proposition is the Malliavin chain rule.

PROPOSITION 4.7. *Let $\varphi \in C_b^1(\mathbb{R}^n, \mathbb{R})$ and $F = (F^1, \ldots, F^n)$, where $F^i \in \mathbb{D}^{1,p}$, $i = 1, \ldots, n$. Then $\varphi(F) \in \mathbb{D}^{1,p}$ and*

$$D\varphi(F) = \sum_{i=1}^n \partial_i \varphi(F) DF^i.$$

PROOF. Let $\psi \in C_0^\infty(\mathbb{R}^n, \mathbb{R})$ be positive with support in the unit ball and with total mass one. Define the so called mollifiers $\phi_k = k^n \phi(kx)$. They have the same properties as ψ only that the support vanishes as $k \to \infty$. The regularized functions

$$\varphi_k(x) = \varphi * \phi_k(x) = \int_{\mathbb{R}^n} \varphi(y)\phi_k(x-y)\,\mathrm{d}y \tag{4.3}$$

[1]Let (F, u) and (F, v) be in the closure $\overline{\mathcal{G}(D)}$ of the graph $\mathcal{G}(D)$. Then there exist sequences $\{F_n^1\}_{n\in\mathbb{N}} \subset \mathcal{S}$ and $\{F_n^2\}_{n\in\mathbb{N}} \subset \mathcal{S}$ such that $(F_n^1, DF_n^1) \to (F, u)$ and $(F_n^2, DF_n^2) \to (F, v)$. We have that $\mathcal{S} \ni F_n^1 - F_n^2 \to 0$ and $D(F_n^1 - F_n^2) \to u - v$. But then $u - v = 0$ (if η above is zero). So we can extend D for $(F, u) \in \overline{\mathcal{G}(D)}$ by $DF = u$.

is $C_b^\infty(\mathbb{R}^n,\mathbb{R})$, $k \in \mathbb{N}$. It is well known that $\varphi_k \to \varphi$ uniformly and since $\varphi \in C_b^1(\mathbb{R}^n,\mathbb{R})$ also $\partial_i\varphi_k \to \partial_i\varphi$ uniformly, $i = 1,\dots,n$. Let $\{F_k^j\}_{k\in\mathbb{N}}$, $j = 1,\dots,n$, be a sequences of smooth random variables converging to F^j respectively in $\mathbb{D}^{1,p}$. Since $\varphi_k(F_l) \in \mathcal{S}$ we have that

$$D\varphi_k(F_l) = \sum_{i=1}^n \partial_i\varphi_k(F_l)DF_l^i.$$

Now, we check convergence.

$$\begin{aligned}
&\mathbb{E}\Big[\Big\|\sum_{i=1}^n \partial_i\varphi_k(F_l)DF_l^i - \sum_{i=1}^n \partial_i\varphi(F)DF^i\Big\|^p\Big] \\
&\le C\Big\{\sum_{i=1}^n \mathbb{E}[\|\partial_i\varphi_k(F_l)DF_l^i - \partial_i\varphi_k(F_l)DF^i\|_H^p] \\
&\quad + \sum_{i=1}^n \mathbb{E}[\|\partial_i\varphi_k(F_l)DF^i - \partial_i\varphi_k(F)DF^i\|_H^p] \\
&\quad + \sum_{i=1}^n \mathbb{E}[\|\partial_i\varphi_k(F)DF^i - \partial_i\varphi(F)DF^i\|_H^p]\Big\} \\
&\to 0, \qquad l,k \to \infty.
\end{aligned}$$

The terms of the first sum converges to zero by dominated convergence since $F_l^i \to F^i$ in $\mathbb{D}^{1,p}$ and the derivatives $\partial_i\varphi_k$ are uniformly bounded in k. The second sum converges to zero by bounded convergence since $F^i \in \mathbb{D}^{1,p}$ and the continuity of $\partial_i\varphi$ and $\partial_i\varphi_k$, $k \in \mathbb{N}$. For the third sum bounded convergence together with the uniform convergence of $\partial\varphi_k \to \partial_i\varphi$ yields the result. □

A similar result will be proved in section 9, for the case when the function φ is globally Lipschitz continuous.

5. Hilbert Space Tensor Products

When defining higher derivatives or the Malliavin derivative of Hilbert space valued random variables we will need tensor products. These are characterized by the following definition but can be realized in several different but isomorphic ways.

Definition 5.1. *Let H and U be two separable Hilbert spaces. The tensor product of H and U, denoted $H \otimes U$, is a Hilbert space together with a bilinear map $H\times U \ni (h,u) \mapsto h\otimes u \in H\otimes U$ with dense range satisfying for all $h_1, h_2 \in H$ and $u_1, u_2 \in U$*

$$\langle h_1 \otimes u_1, h_2 \otimes u_2\rangle_{H\otimes U} = \langle h_1, h_2\rangle_H\langle u_1, u_2\rangle_U. \tag{5.1}$$

We will realize $H \otimes U$ as the space of Hilbert-Schmidt operators from U to H. Let $L(U, H)$ denote the Banach space of all bounded linear operators $L : U \to H$.

DEFINITION 5.2. *An operator $T \in L(U,H)$ is Hilbert-Schmidt if for any ON-basis $\{e_n\}_{n\in\mathbb{N}} \subset U$*

$$\sum_{k=1}^{\infty} \|Te_k\|_H^2 < \infty.$$

PROPOSITION 5.3. *The space of Hilbert-Schmidt operators, denoted $L_2(U,H)$, is a separable Hilbert space with inner product and norm*

$$\langle S,T\rangle_{L_2(U,H)} = \sum_{k=1}^{\infty}\langle Se_k, Te_k\rangle_H, \quad \|S\|_{L_2(U,H)} = \Big(\sum_{k=1}^{\infty}\|Se_k\|_H^2)\Big)^{\frac{1}{2}},$$

where $\{e_k\}_{k\in\mathbb{N}} \subset U$ is an ON-basis.

Together with the map $H \times U \ni (h,u) \mapsto h\otimes u$ with $(h\otimes u)v = h\langle u,v\rangle_U$ for $v \in U$, we have that $L_2(U,H) = H \otimes U$.

PROOF. It is straight forward to show that the inner product is well defined, i.e. does not depend on the choice of orthonormal basis. Moreover, property (5.1) is satisfied since

$$\begin{aligned}\langle h_1\otimes u_1, h_2\otimes u_2\rangle_{L_2(U,H)} &= \sum_{k=1}^{\infty}\langle (h_1\otimes u_1)e_k, (h_2\otimes u_2)e_k\rangle_H \\ &= \langle h_1,h_2\rangle_H \sum_{k=1}^{\infty}\langle u_1,e_k\rangle_U\langle u_2,e_k\rangle_U \\ &= \langle h_1,h_2\rangle_H\langle u_1,u_2\rangle_U.\end{aligned}$$

Let $\{e_k\}_{k\in\mathbb{N}} \subset H$ and $\{f_k\}_{k\in\mathbb{N}} \subset U$ be ON-bases. It is immediate that $\{e_k\otimes f_l\}_{k,l\in\mathbb{N}}$ is an ON-system for $L_2(U,H)$. To prove that it is complete let $\langle T, e_k\otimes f_l\rangle_{L_2(U,H)} = 0$, for all $k,l\in\mathbb{N}$. Then, for all $k,l\in\mathbb{N}$,

$$\begin{aligned}\langle T, e_k\otimes f_l\rangle_{L_2} &= \sum_{n=1}^{\infty}\langle Tf_n, (e_k\otimes f_l)f_n\rangle_H \\ &= \sum_{n=1}^{\infty}\langle Tf_n, e_k\rangle_H\langle f_n,f_l\rangle_U = \langle Tf_l, e_k\rangle_H = 0.\end{aligned}$$

This implies that $Tf_l = 0$ for all $l \in \mathbb{N}$ and hence $T = 0$. We have proved separability.

Now to completeness. Let $\{T_n\}_{n\in\mathbb{N}} \subset L_2(U,H)$ be a Cauchy sequence. Then, since $\|T\|_{L(U,H)} \le \|T\|_{L_2(U,H)}$, for all $T \in L_2(U,H)$[2], $\{T_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $L(U,H)$. Now, $L(U,H)$ is complete so there exists $T \in L(U,H)$ such that

[2]For any $\epsilon > 0$, $\exists\{e_k\}_{k\in\mathbb{N}} \subset U$ ONB: $\|T\|_{L(U,H)}^2 \le \|Te_1\|_H^2 + \epsilon \le \sum_{k=1}^{\infty}\|Te_k\|_H^2 + \epsilon = \|T\|_{L_2(U,H)}^2 + \epsilon$.

$\|T_n - T\|_{L(U,H)} \to 0$. Let $\epsilon > 0$. For large enough $n \in \mathbb{N}$ it follows by Fatou's Lemma that

$$\begin{aligned}
\|T_n - T\|^2_{L_2(U,H)} &= \sum_{k=1}^{\infty} \|(T_n - T)f_k\|^2_H = \sum_{k=1}^{\infty} \lim_{m\to\infty} \|(T_n - T_m)f_k\|^2_H \\
&\leq \liminf_{m\to\infty} \sum_{k=1}^{\infty} \|(T_n - T_m)f_k\|^2_H = \liminf_{m\to\infty} \|T_n - T_m\|_{L_2(U,H)} \\
&< \epsilon.
\end{aligned}$$

Since ϵ is arbitrary $\|T_n - T\|_{L_2(U,H)} \to 0$ as $n \to \infty$ and we have proved the completeness. This finishes the proof. $\square$

Tensor product spaces of several Hilbert spaces are defined iteratively, a.e., $H^{\otimes 3} = H \otimes (H \otimes H) \cong (H \otimes H) \otimes H$ and so on.

6. Higher Derivatives

The tensor product machinery allows us to iterate the Malliavin derivative. Let $F = f(I(h_1), \dots, I(h_m))$ be a smooth random variable. The second derivative is given by

$$D^2 F = \sum_{i,j=1}^{m} \partial_i \partial_j f(I(h_1), \dots, I(h_m)) h_i \otimes h_j.$$

One can choose two different realizations of the tensor product. If we let

$$h_1 \otimes h_2 \cdot (g_1, g_2) = \langle h_1 \otimes h_2, g_1 \otimes g_2 \rangle_{H\otimes H} = \langle h_1, g_1 \rangle \langle h_2, g_2 \rangle$$

we have by an easy check a valid realization of the tensor product, in view of definition 5.1. This realization give us the representation of D^2F as a bounded bilinear form $D^2F(\omega) : H \times H \to \mathbb{R}$. We have by Parseval's formula that, for an ON-basis $\{e_k\}_{k\in\mathbb{N}}$ and a bilinear functional b,

$$\|b\|^2_{H\otimes H} = \sum_{i,j=1}^{\infty} |\langle b, e_i \otimes e_j \rangle_{H\otimes H}|^2 = \sum_{i,j=1}^{\infty} |b \cdot (e_i, e_j)|^2.$$

Hence $b \in H \otimes H$ if and only if b is Hilbert-Schmidt. This is not the approach we will take. Instead we consider the operator that, by Riesz representation theorem, determines the bilinear form b, i.e. the operator determined by the relation

$$b \cdot (f, g) = \langle Bf, g \rangle, \quad \forall f, g \in H.$$

Clearly B is an Hilbert-Schmidt operator on H, indeed,

$$\|B\|^2_{L_2(H)} = \sum_{k\in\mathbb{N}} |\langle Be_k, e_k \rangle|^2 = \sum_{k\in\mathbb{N}} |b \cdot (e_k, e_k)|^2 < \infty.$$

We thus instead consider the tensor product realization of the previous section, i.e.

$$h_1 \otimes h_2 \cdot f = h_1 \langle h_2, f \rangle$$

and $H \otimes H = L_2(H)$, i.e., the the space of Hilbert-Schmidt operators. Hence, with this approach the second Malliavin derivative is considered an operator. One can think of it as the Malliavin Hessian operator. Iteratively we can define the k:th Malliavin derivative taking values in $H^{\otimes k}$ as

$$D^k F = \sum_{i_1,\dots,i_k=1}^{m} \partial_{i_1} \cdots \partial_{i_k} f(I(h_1), \dots, I(h_m)) h_{i_1} \otimes \cdots \otimes h_{i_k}.$$

Introducing the norms

$$\|F\|_{k,p} = \Big(\mathbb{E}[|F|^p] + \sum_{j=1}^{k} \mathbb{E}[\|D^j F\|_{H^{\otimes j}}]^p \Big)^{\frac{1}{p}}.$$

we can analogous to the case $k = 1$ take the closure and define the higher ordered Watanabe-Sobolev spaces $\mathbb{D}^{k,p}$. It follows directly that $\mathbb{D}^{k,p}$ are monotonously decreasing in both p and k.

7. Hermite Polynomials

In order to develop the Wiener Chaos decomposition in the next section we need some knowledge about Hermite polynomials. The exposition is intensionally very explicit and somehow elementary.

We define, for $n > 0$ the n:th Hermite polynomial $H_n(x)$, $x \in \mathbb{R}$, by

$$H_n(x) = \frac{(-1)^n}{n!} e^{\frac{x^2}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-\frac{x^2}{2}}, \tag{7.1}$$

and $H_0(x) = 1$. The convention here to include the $1/n!$ in the definition is not standard. The following proposition states some important properties of $\{H_n\}_{n\in\mathbb{N}}$.

Proposition 7.1. *Let $\{H_n\}_{n\in\mathbb{N}}$ be given by (7.1). For $n \geq 1$ the following properties holds:*

(i) $H_n'(x) = H_{n-1}(x)$,

(ii) $(n+1)H_{n+1}(x) = xH_n(x) - H_{n-1}(x)$,

(iii) $H_n(-x) = (-1)^n H_n(x)$.

Proof. Consider the function $F(x,t) = \exp(tx - \frac{t^2}{2})$. Making a Maclaurin expansion of F in t we get that

$$
\begin{aligned}
F(x,t) &= \exp\left(\frac{x^2}{2} - \frac{1}{2}(x-t)^2\right) \\
&= e^{\frac{x^2}{2}} \sum_{n=0}^{\infty} \frac{t^n}{n!}\left(\frac{\mathrm{d}^n}{\mathrm{d}t^n} e^{-\frac{(x-t)^2}{2}}\right)\Big|_{t=0} \\
&= e^{\frac{x^2}{2}} \sum_{n=0}^{\infty} t^n \frac{(-1)^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-\frac{x^2}{2}} \\
&= \sum_{n=0}^{\infty} t^n H_n(x).
\end{aligned} \tag{7.2}
$$

Here we have used that $\frac{\mathrm{d}^n}{\mathrm{d}t^n} e^{-\frac{(x-t)^2}{2}}|_{t=0} = (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-\frac{x^2}{2}}$. This can be shown by induction. An easy calculation shows that $\frac{\mathrm{d}}{\mathrm{d}x} F(x,t) = tF(x,t)$, so that

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x,t) = tF(x,t) = \sum_{n=1}^{\infty} t^n H_{n-1}(x).$$

Differentiating term by term in (7.2) we get that

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x,t) = \sum_{n=1}^{\infty} t^n H_n'(x).$$

Comparing the terms gives property (i). For property (ii) we have by property (i) and differentiation of H_n that

$$H_{n-1}(x) = \frac{\mathrm{d}}{\mathrm{d}x} H_n(x) = xH_n(x) - (n+1)H_{n+1}(x).$$

For property (iii) we notice from the definition of F that $F(x,t) = F(-x,-t)$. Now,

$$\sum_{n=1}^{\infty} t^n H_n(x) = F(t,x) = F(-t,-x) = \sum_{n=0}^{\infty} t^n (-1)^n H_n(-x).$$

Comparing terms again gives property (iii). □

Consider the Ornstein-Uhlenbeck operator L^1, introduced in Section 1,

$$L^1 f(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) - x\frac{\mathrm{d}}{\mathrm{d}x} f(x)$$

acting on smooth enough functions $f : \mathbb{R} \to \mathbb{R}$. Now, we let L^1 act on the n:th Hermite polynomial.

$$L^1 H_n(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} H_n(x) - x\frac{\mathrm{d}}{\mathrm{d}x} H_n(x) = H_{n-2}(x) - xH_{n-1}(x) = nH_n(x).$$

Here property (i) and (ii) of Proposition 7.1 has been used. It is clear that the Hermite polynomials are the eigenfunctions for the Ornstein-Uhlenbeck operator in dimension 1. Hermite polynomial H_n has the corresponding eigenvalue n. We will show that $\{H_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for $L_2(\mathbb{R},\nu^1)$, where ν^1 is the standard normal law.

We will move into a probabilistic setting and actually prove a slightly more general result for Gaussian random variables. We need the following Lemma.

LEMMA 7.2. *Let X and Y have a joint Gaussian distribution with $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ and $\mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1$. Then*

$$\mathbb{E}\Big[\exp\Big(sX - \frac{s^2}{2}\Big)\exp\Big(tY - \frac{t^2}{2}\Big)\Big] = \exp[st\mathbb{E}[XY]], \quad \forall s,t\in\mathbb{R}. \tag{7.3}$$

PROOF. Denote the correlation by $\rho = \mathbb{E}[XY]$. We compute, using standard techniques,

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(sX - \frac{s^2}{2}\Big)\exp\Big(tY - \frac{t^2}{2}\Big)\Big]\\
&\quad= \int\int e^{sx-\frac{s^2}{2}}e^{ty-\frac{y^2}{2}}e^{-(x^2-2\rho xy+y^2)}\frac{\mathrm{d}x\,\mathrm{d}y}{2\pi\sqrt{1-\rho^2}}\\
&\quad= \int\int e^{-(x-s)^2/2}e^{-(y-t)^2/2}e^{\rho xy}\frac{\mathrm{d}x\,\mathrm{d}y}{2\pi\sqrt{1-\rho^2}}\\
&\quad= e^{st\rho}\int\int e^{-(z^2-2\rho zw+w^2)/2}e^{\rho(tz+sw)}\frac{\mathrm{d}z\,\mathrm{d}w}{2\pi\sqrt{1-\rho^2}}\\
&\quad= \exp(st\mathbb{E}[XY])\mathbb{E}[\exp(\rho(tX+sY))].
\end{aligned}$$

The factor $\mathbb{E}[\exp(\rho(tX+sY))] = 1$ since $\rho(tX+sY)$ is Gaussian with zero mean and $\mathbb{E}[e^G] = e^{\mathbb{E}[G]}$ for any Gaussian random variable G. ☐

We are now ready to prove the orthogonality relationship for Hermite polynomials of Gaussian random variables.

LEMMA 7.3. *Let X and Y have a joint Gaussian distribution with $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ and $\mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1$. Then, for any non-negative $n \neq m$, $H_n(X)$ and $H_m(Y)$ are orthogonal, i.e., $\mathbb{E}[H_n(X)H_m(Y)] = 0$. Moreover*

$$\mathbb{E}[H_n(X)H_n(Y)] = \frac{1}{n!}\mathbb{E}[XY]^n.$$

If we let $X = Y$ we obtain the classical orthogonality property for $\{H_n\}_n$ on $L_2(\mathbb{R},\nu^1)$.

PROOF. Differentiating the left hand side of (7.3) utilizing the expansion (7.2) yields

$$
\begin{aligned}
&\frac{\partial^{m+n}}{\partial s^m \partial t^n}\mathbb{E}\Big[\sum_{i=1}^{\infty} s^i H_i(X) \sum_{j=1}^{\infty} t^j H_j(Y)\Big]\Big|_{s,t=0} \\
&= \mathbb{E}\Big[\sum_{i=m}^{\infty} i(i-1)\cdots(i-n+1)s^{i-m}H_i(X) \\
&\quad \times \sum_{j=n}^{\infty} j(j-1)\cdots(j-n+1)t^{j-n}H_j(Y)\Big]\Big|_{s,t=0} \\
&= \mathbb{E}[m!H_m(X)n!H_n(Y)]
\end{aligned} \tag{7.4}
$$

Differentiating the right hand side of (7.3) it can easily be shown that

$$
\frac{\partial^{m+n}}{\partial s^m \partial t^n}\exp(st\mathbb{E}[XY])\Big|_{s,t=0} = n!E[XY]^n \tag{7.5}
$$

if $n = m$ and zero otherwise. Equating (7.4) and (7.5) finishes the proof.

□

8. The Wiener Chaos Decomposition

In this section a useful decomposition, due to Wiener, of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ known as Wiener Chaos decomposition will be proved. It will be of central importance when doing spectral theory for the Ornstein-Uhlenbeck operator and it's corresponding semigroup. We will also need it when proving an important property of the adjoint of the Malliavin derivative. The Hermite polynomials will be a useful tool for us.

We return to the setting of an isonormal Gaussian process $\{I(h) : h \in H\}$ where H is our separable Hilbert space.

Definition 8.1. *The n:th Wiener Chaos $\mathcal{H}_n$ is the closed subspace of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ given by $\{H_n(I(h)) : h \in H, \|h\| = 1\}$. Here H_n is the n:th Hermite polynomial. We denote the orthogonal projection onto $\mathcal{H}_n$ by J_n.*

It is clear that $\mathcal{H}_0$ is the space of constants and $\mathcal{H}_1$ is our Gaussian Hilbert space $\{I(h) : h \in H\}$. Lemma 7.3 tells us that $\mathcal{H}_n \perp \mathcal{H}_m$, whenever $n \neq m$. The following theorem is the so called Wiener-Chaos decomposition of $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 8.2. *The following decomposition of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ holds.*

$$
L_2(\Omega, \mathcal{F}, \mathbb{P}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n.
$$

For the proof we need the following lemma

Lemma 8.3. *If $\mathbb{E}[Xe^{I(h)}] = 0$ for all $h \in H$. Then $X = 0$.*

proof of lemma. If X satisfies the given assumption, then clearly

$$\mathbb{E}\Big[X \exp\Big(\sum_{i=1}^{m} s_i W(h_i)\Big)\Big] = 0 \tag{8.1}$$

for any $s_1, \ldots, s_m \in \mathbb{R}$ and $h_1, \ldots, h_m \in H$. Let us fix $m \geq 1$ and $h_1, \ldots, h_m \in H$. We have that

$$\begin{aligned} \mathbb{E}\Big[X \exp\Big(\sum_{i=1}^{m} s_i W(h_i)\Big)\Big] &= \int_\Omega \exp\Big(\sum_{i=1}^{m} s_i W(h_i)\Big) X \,\mathrm{d}\mathbb{P} \\ &:= \int_{\mathbb{R}^m} \exp\Big(\sum_{i=1}^{m} s_i x_i\Big) \,\mathrm{d}\mu(\mathbf{x}) \end{aligned} \tag{8.2}$$

Here, for $A \in \mathcal{B}_{\mathbb{R}^m}$,

$$\begin{aligned} \mu(A) &= \Big([I(h_1), \ldots, I(h_m)]^{-1} \circ \int_{\cdot} X \,\mathrm{d}\mathbb{P}\Big)(A) \\ &= \int_{[I(h_1),\ldots,I(h_m)]^{-1}(A)} X \,\mathrm{d}\mathbb{P} \\ &= \int_\Omega \chi_A(I(h_1), \ldots, I(h_m)) X \,\mathrm{d}\mathbb{P} \\ &= \mathbb{E}[\chi_A(I(h_1), \ldots, I(h_m))X]. \end{aligned}$$

Hence (8.1) and (8.2) imply that the Laplace transform of the signed measure μ equals to zero. Hence, μ is the zero measure on $\mathbb{R}^m$. Since $\mathcal{F}$ is generated by $\{I(h) : h \in H\}$ we have that $\mathbb{E}[X\chi_F] = 0$, for all $F \in \mathcal{F}$. This implies that $X = 0$. $\square$

PROOF OF THEOREM 8.2. Let $X \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ be orthogonal to all Wiener-Chaos $\mathcal{H}_n$, $n \in \mathbb{N}$. Then $\mathbb{E}[XH_n(I(h))] = 0$, for all $n \in \mathbb{N}$ and $h \in H$. Since x^n can be expressed as a linear combination of Hermite polynomials of order less than or equal to n, $\mathbb{E}[XI(h)^n] = 0$, for all $n \in \mathbb{N}$ and $h \in H$. Expressing $\exp(I(h))$ as a power series we also have that $\mathbb{E}[X\exp(I(h))] = 0$, for all $n \in \mathbb{N}$ and $h \in H$. But, by Lemma 8.3, $X = 0$ and we are done. $\square$

Let $\mathcal{P}_n^0$ be the set of all random variables of the form $p(I(h_1), \ldots, I(h_m))$ where p is a polynomial of degree not greater that n, $m \in \mathbb{N}$ and $h_1, \ldots, h_m \in H$. We denote its closure in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ by $\mathcal{P}_n$.

LEMMA 8.4. *For any $n \in \mathbb{N}$,*

$$\mathcal{P}_n = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n.$$

PROOF. The inclusion $\mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n \subset \mathcal{P}_n$ is obvious. For the converse inclusion it is enough to prove that $\mathcal{P}_n \perp \mathcal{H}_k$ for all $k \geq n$. For this it suffices to show that $\mathbb{E}[p(I(h_1),\ldots,I(h_m))H_k(I(h))] = 0$, for $h, h_1\ldots,h_m \in H$, $\|h\| = 1$ and $p \in \mathcal{P}_n^0$, since $\mathcal{P}_n^0 \subset \mathcal{P}_n$ is dense. By a Gram-Schmidt process and by linearity of $h \mapsto I(h)$ we can replace $p(I(h_1),\ldots,I(h_m))$ by $q(W(e_1),\ldots,W(e_m),I(h))$ for some $q \in \mathcal{P}_n^0$ and $\{e_i\}_{i=1}^m \subset H$ orthonormal. By independence it then suffices to prove that $\mathbb{E}[I(h)^r H_k(I(h))] = 0$, for $1 \le r \le n < k$. This holds since $I(h)^r$ can be written as a linear combination of Hermite polynomials of $I(h)$. □

Let Λ be the set of all sequence $a = (a_1, a_2, \ldots)$, where $a_i \in \mathbb{N}$, $i \in \mathbb{N}$, with $a_i = 0$ except for a finite number of $i \in \mathbb{N}$. We refer to Λ as a set of multiindices. Let $|a| = \sum_{i=1}^\infty a_i$ and $a! = \sum_{i=1}^\infty a_i!$. Let, for $a \in \Lambda$ and $\{e_i\}_{i\in\mathbb{N}} \subset H$ being an orthonormal basis,

$$\Phi_a = \sqrt{a!}\prod_{i=1}^\infty H_{a_i}(W(e_i)).$$

This is a well defined product since all but finitely many $H_{a_i} = 1$.

Proposition 8.5. *The family of random variables $\{\Phi_a\}_{a\in\Lambda}$ is an orthonormal basis for $L_2(\Omega,\mathcal{F},\mathbb{P})$. For any $n \ge 1$,*

$$\{\Phi_a : a \in \Lambda, |a| = n\}$$

is an orthonormal basis for $\mathcal{H}_n$.

Proof. We prove the second statement. The first is then a consequence of the Wiener-Chaos decomposition, Theorem 8.2. Let $a, b \in \Lambda$. Then, by independence and Lemma 7.3,

$$\begin{aligned}
\mathbb{E}[\Phi_a\Phi_b] &= \sqrt{a!b!}\mathbb{E}\Big[\prod_{i=1}^\infty H_{a_i}(W(e_i))H_{b_i}(W(e_i))\Big] \\
&= \sqrt{a!b!}\prod_{i=1}^\infty \mathbb{E}[H_{a_i}(W(e_i))H_{b_i}(W(e_i))] \\
&= \begin{cases} 1 & \text{if} \quad a = b \\ 0 & \text{if} \quad a \ne b \end{cases}
\end{aligned}$$

So, the families $\{\Phi_a : a \in \Lambda, |a| = n\}$ and $\{\Phi_a : a \in \Lambda, |a| = m\}$ are orthogonal. Its respective members are mutually orthogonal. For $n = 1$, $\{\Phi_a : a \in \Lambda, |a| = 1\} = \{W(e_k)\}_{k\in\mathbb{N}}$. Clearly $\mathrm{cl}\{W(e_k)\}_{k\in\mathbb{N}} = \{I(h) : h \in H\} = \mathcal{H}_1$ so $\{\Phi_a : a \in \Lambda, |a| = 1\}$ is an ON-basis for $\mathcal{H}_1$. Assume inductively that $\{\Phi_a : a \in \Lambda, |a| = n-1\}$ is an ON-basis for $\mathcal{H}_{n-1}$. Clearly $\{\Phi_a : a \in \Lambda, |a| = n\} \subset \mathcal{P}_n \setminus \mathcal{P}_{n-1} = \mathcal{H}_n$ since Φ_a is a polynomial of degree at most n and since $\mathcal{P}_{n-1} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_{n-1} \perp \{\Phi_a : a \in \Lambda, |a| = n\}$ by the inductive assumption. Since $\mathcal{P}_n^0$ is dense in $\mathcal{P}_n$ it is enough to show that any $p \in \mathcal{P}_n^0$ can be approximated by polynomials of members of the family $\{W(e_k)\}_{k\in\mathbb{N}}$. This is clear since $\{e_k\}_{k\in\mathbb{N}}$ is a basis for H. □

REMARK 8.6. Later it will be shown that $\{\Phi_a\}_{a\in\Lambda}$ is the eigenvectors for the Ornstein-Uhlenbeck operator, still not defined. $\mathcal{H}_n$ is the eigenspace corresponding to the eigenvalue n for this operator.

9. More About the Malliavin Derivative

In Section 4 we introduced the derivative operator and proved some important properties of it. With the Wiener Chaos machinery as a tool we shall now prove more. The first result is a necessary and sufficient criterion for $F \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ to be in $\mathbb{D}^{k,2}$.

PROPOSITION 9.1. *Let $F \in L_2(\Omega)$. Then $F \in \mathbb{D}^{k,2}$ if and only if*

$$\sum_{n=1}^{\infty} n^k \mathbb{E}[|J_n F|^2] < \infty.$$

In this case, $D^k J_n F = J_{n-k}(D^k F)$.

PROOF. Consider $k = 1$. Let $F \in L_2(\Omega)$, $\{\Phi_a\}_{a\in\Lambda} \subset L_2(\Omega)$ being the ON-basis we defined in Section 8 and $|a| = n$.

$$D\Phi_a = \sqrt{a!}\sum_{j=1}^{\infty}\prod_{i=1,i\neq j}^{\infty} H_{a_i}(W(e_i))H_{a_j-1}(W(e_j))e_j$$

It is clear that $D\Phi_a \in \mathcal{H}_{n-1}$. A similar calculation to (8) shows that $\{D\Phi_a : a \in \Lambda, |a| = n\}$ is an ON-basis for $\mathcal{H}_{n-1}(H)$. This space is defined as the closure in $L_2(\Omega, H)$ of $\{\sum_{i=1}^m F_i h_i : m \geq 1, F_i \in \mathcal{H}_n, i = 1, 2, \ldots, m, h_1, h_2, \ldots, h_m \in H\}$. By the Pythagorean theorem, independence of $W(e_i)$ and $W(e_j)$ for $i \neq j$ and Lemma 7.3, we have

$$\mathbb{E}[\|D\Phi_a\|^2] = a!\sum_{j=1}^{\infty}\prod_{i=1,i\neq j}^{\infty}\mathbb{E}[H^2_{a_i}(W(e_i))]\mathbb{E}[H^2_{a_j-1}(W(e_j))] =$$
$$\sum_{j=1}^{\infty}\frac{\prod_{i=1}^{\infty} a_i!}{\prod_{i=1,i\neq j}^{\infty} a_i!(a_j-1)!} = \sum_{j=1}^{\infty}\frac{a_j!}{(a_j-1)!} = |a| = n.$$

Attacking the norm of DF, we get using the Pythagorean theorem and Parseval's formula that

$$\begin{aligned}\mathbb{E}[\|DF\|^2] &= \mathbb{E}\Big[\Big\|\sum_{n=1}^{\infty}\sum_{a\in\Lambda,|a|=n}\langle F, \Phi_a\rangle D\Phi_a\Big\|^2\Big] \\ &= \sum_{n=1}^{\infty}\sum_{a\in\Lambda,|a|=n}|\langle J_n F, \Phi_a\rangle|^2\mathbb{E}[\|D\Phi_a\|^2] \\ &= \sum_{n=1}^{\infty} n\mathbb{E}[|J_n F|^2].\end{aligned}$$

This is what we wanted to prove for $k=1$. Iterating we get that

$$\mathbb{E}[\|D^kF\|_{H^{\otimes k}}] = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)\mathbb{E}[|J_nF|^2].$$

This is finite if and only if

$$\sum_{n=1}^{\infty} n^k \mathbb{E}[|J_nF|^2] < \infty$$

and we are done. $\square$

The following two corollaries are proved for $\mathbb{D}^{1,1}$ by Nualart. Notice that since $\{\mathbb{D}^{1,p}\}_{p\geq 1}$ is a family of spaces, monotonously decreasing in p, $F \in \mathbb{D}^{1,p}$, for $p \geq 1$, implies $F \in \mathbb{D}^{1,1}$. We here settle with $p \geq 2$.

Corollary 9.2. *Let $F \in \mathbb{D}^{1,2}$ with $DF = 0$. Then $DF = \mathbb{E}[F]$.*

Proof. Since $DF = 0$ the only Wiener chaos that can be nonzero is $\mathcal{H}_0$, consisting of constant random variables. $\square$

Corollary 9.3. *Let $A \in \mathcal{F}$. Then, the indicator function $\chi_A \in \mathbb{D}^{1,2}$ if and only if $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.*

Proof. By applying the chain rule to $\phi(x) = x^2$ it yields that

$$D\chi_A = D\chi_A^2 = 2\chi_A D\chi_A.$$

This is possible only if $D\chi_A = 0$. By Corollary 9.2 we have that $\chi_A = \mathbb{E}[\chi_A] = \mathbb{P}(A)$. $\square$

10. The Divergence Operator

We want to define the adjoint operator δ to $D : \mathbb{D}^{1,2} \to L_2(\Omega, H)$. For this purpose we need the functional $l_u(F) = \mathbb{E}[\langle DF, u\rangle_H] = \langle DF, u\rangle_{L_2(\Omega,H)}$ to be L_2-continuous on $\mathbb{D}^{1,2}$. The domain of δ is

$$\mathrm{Dom}(\delta) = \{u \in L_2(\Omega, H) : |\mathbb{E}[\langle DF, u\rangle_H]| \leq C_u(\mathbb{E}|F|^2)^{\frac{1}{2}}, \forall F \in \mathbb{D}^{1,2}\}$$

Since l_u is continuous on a dense domain $\mathbb{D}^{1,2} \subset L_2(\Omega)$ for every $u \in \mathrm{Dom}(\delta)$ there is an extension by continuity $\tilde{l}_u$ on the whole of $L_2(\Omega)$. Riesz representation theorem guarantees the existence of an element $G \in L_2(\Omega)$ such that $l_u(F) = \tilde{l}_u(F) = \mathbb{E}[FG] = \langle F, G\rangle_{L_2(\Omega)}$. It is easy to prove that G is uniquely and linearly determined by u, so we define $\delta(u) := G$. Hence we have for $F \in \mathbb{D}^{1,2}$ and $u \in \mathrm{Dom}(\delta)$

$$\mathbb{E}[\langle DF, u\rangle_H] = \mathbb{E}[F\delta(u)]. \tag{10.1}$$

or equivalently

$$\langle DF, u\rangle_{L_2(\Omega,H)} = \langle F, \delta(u)\rangle_{L_2(\Omega)}.$$

Now, let us explore the properties of the operator δ and for $u \in \mathrm{Dom}(\delta)$ the real valued random variables $\delta(u)$. By letting $F = 1$ in (10.1) we get that $\mathbb{E}[\delta(u)] = 0$ for $u \in \mathrm{Dom}(\delta)$. Moreover, by definition δ is linear. We define the class $\mathcal{S}_H$ of H-valued random variables of the form

$$u = \sum_{i=1}^{n} F_i h_i, \tag{10.2}$$

where $F_i \in \mathcal{S}$, $h_i \in H$, $i = 1, \ldots, n$ and $n \in \mathbb{N}$. We claim that for $u \in \mathcal{S}_H$

$$\delta(u) = \sum_{j=1}^{n} F_j W(h_j) - \langle DF_j, h_j\rangle_H. \tag{10.3}$$

This follows by taking the adjoint and using equation (4.2) for $G \in \mathcal{S}$

$$\begin{aligned}\mathbb{E}[G\delta(u)] = \mathbb{E}[\langle DG, u\rangle_H] &= \sum_{j=1}^{n} \mathbb{E}[\langle F_j DG, h_j\rangle_H] = \sum_{j=1}^{n} \mathbb{E}[\langle GDF_j - DF_j G, h_j\rangle_H] \\ &= \sum_{j=1}^{n} \mathbb{E}[GF_j W(h_j)] - \mathbb{E}[G\langle DF_j, h_j\rangle_H] = \mathbb{E}\Big[G\Big(\sum_{j=1}^{n} F_j W(h_j) - \langle DF_j, h_j\rangle_H\Big)\Big]\end{aligned}$$

for all $G \in \mathcal{S}$. Since $\mathcal{S} \subset L_2(\Omega)$ is dense (10.3) holds true. Notice that when $u = h \in H$ the rightmost sum in (10.3) vanishes and we have $\delta(h) = I(h)$. This was expected in view of Remark 4.4.

Recall that the second Malliavin derivative of a random variable $F \in \mathbb{D}^{2,2}$ is a random variable taking its value in the space of Hilbert-Schmidt operators, $H \otimes H$. It is also the Malliavin derivative of the H-valued random variable DF. Hence, the first Malliavin derivative of a smooth enough, H-valued, random variable is $H \otimes H$-valued. For $u \in \mathcal{S}_H$ of the form (10.2) we have the following explicit expressions for the derivative and directional derivative respectively,

$$Du = \sum_{i=1}^{n} h_i \otimes DF_i$$

and

$$D^h u = \sum_{i=1}^{n} D^h F_i h_i = \sum_{i=1}^{n} \langle DF_i, h\rangle_H h_i = \Big(\sum_{i=1}^{n} h_i \otimes DF_i\Big) h = (Du)h.$$

We will use that

$$D\langle u,h\rangle_H = D\sum_{i=1}^n F_i\langle h_i,h\rangle_H = \Big(\sum_{i=1}^n DF_i\otimes h_i\Big)h = (Du)^*h, \tag{10.4}$$

where * denotes the adjoint. Here it has been used that $(u\otimes v)^* = v\otimes u$. This is easy to prove.

The space $\mathbb{D}^{1,2}(H)$ is defined as the closure of $\mathcal{S}_H$ under the norm

$$\|u\|_{1,2,H} = \Big(\mathbb{E}[\|u\|^2] + \mathbb{E}[\|Du\|^2_{H\otimes H}]\Big)^{\frac{1}{2}}.$$

Also, for $h\in H$, $\mathbb{D}^{h,2}(H)$ is the space all $u\in L_2(\Omega,H)$ such that $D^h u\in L_2(\Omega,H)$.

We state the basic properties of δ in the next Proposition.

PROPOSITION 10.1. *The operator δ satisfies:*

(i) $\mathbb{E}[\delta(u)] = 0, \quad \forall u\in Dom(\delta)$,

(ii) $u\mapsto\delta(u)$ *is linear,*

(iii) $\delta(h) = I(h)$ *for* $h\in H$,

(iv) $\mathbb{E}[\delta(u)\delta(v)] = \mathbb{E}[\langle u,v\rangle_H] + \mathbb{E}[Tr(Du\circ Dv)], \quad \forall u,v\in\mathbb{D}^{1,2}(H)\subset Dom(\delta)$,

(v) $D^h\delta(u) = \langle u,h\rangle_H + \delta(D^h u), \quad \forall u\in\mathbb{D}^{1,2}(H)$ *with* $D^h u\in Dom(\delta)$.

In order to prove Proposition 10.1 we need the following lemma.

LEMMA 10.2. *Let $G\in L_2(\Omega)$. If there exists $Y\in L_2(\Omega)$ such that*

$$\mathbb{E}[G\delta(hF)] = \mathbb{E}[YF],$$

for all $F\in\mathbb{D}^{1,2}$, then $G\in\mathbb{D}^{h,2}$ and $D^hG = Y$.

PROOF OF LEMMA. We expand G into its Wiener-Chaos expansion. Recall that J_n denotes the orthogonal projection onto the n:th Wiener-Chaos $\mathcal{H}_n$. Monotone convergence and duality yields

$$\begin{aligned}\mathbb{E}[YF] &= \mathbb{E}[G\delta(hF)] = \mathbb{E}[\Big(\sum_{i=1}^\infty J_nG\Big)\delta(hF)] = \sum_{i=1}^\infty \mathbb{E}[J_nG\delta(hF)] \\ &= \sum_{i=1}^\infty \mathbb{E}[\langle DJ_nG,hF\rangle_H] = \sum_{i=1}^\infty \mathbb{E}[F\langle DJ_nG,h\rangle_H] = \sum_{i=1}^\infty \mathbb{E}[FD^h(J_nG)].\end{aligned}$$

Hence, we have that

$$Y = \sum_{n=0}^\infty J_nY = \sum_{n=1}^\infty J_{n-1}Y = \sum_{n=1}^\infty D^h(J_nG)$$

and $J_{n-1}Y = D^h(J_nG)$ for all $n \in \mathbb{N}$. In view of Proposition 9.1 this implies that

$$\begin{aligned} J_{n-1}Y &= D^h(J_nG) = \langle DJ_nG, h\rangle_H = \langle J_{n-1}(DG), h\rangle_H = J_{n-1}\langle DG, h\rangle_H \\ &= J_{n-1}D^hG. \end{aligned}$$

for all $n \in \mathbb{N}$. We have proven that $Y = D^hG$. □

Proof of proposition 10.1. Properties (i) and (ii) are already proved. The proof goes as follows. We prove (v) for $u \in \mathcal{S}_H$ and use it to prove (iv). Finally we prove (v) in full generality using (iv) and Lemma 10.2.

Let $u \in \mathcal{S}_H$ have the form

$$u = \sum_{i=1}^{n} F_i h_i.$$

In the next calculation we use the symmetry $\langle D\langle DF, h\rangle_H, g\rangle_H = \langle D\langle DF, g\rangle_H, h\rangle_H$. This is easily checked, by direct calculations, taking $F = f(I(h_1), \ldots, I(h_n))$, for $f \in C_p^\infty(\mathbb{R}^n)$. From applying (10.3), the product rule and then (10.3) again we have that

$$\begin{aligned} D^h\delta(u) &= D^h\Big(\sum_{j=1}^{n} F_jW(h_j) - \sum_{j=1}^{n}\langle DF_j, h_j\rangle_H\Big) \\ &= \sum_{j=1}^{n} F_j\langle DW(h_j), h\rangle_H + \sum_{j=1}^{n}\langle DF_j, h\rangle_H W(h_j) - \sum_{j=1}^{n}\langle D\langle DF_j, h_j\rangle_H, h\rangle_H \\ &= \Big\langle \sum_{j=1}^{n} F_jh_j, h\Big\rangle_H + \sum_{j=1}^{n}\langle DF_j, h\rangle_H W(h_j) - \sum_{j=1}^{n}\langle D\langle DF_j, h\rangle_H, h_j\rangle_H \\ &= \langle u, h\rangle_H + \delta\Big(\sum_{j=1}^{n}\langle DF_j, h\rangle_H h_j\Big) = \langle u, h\rangle_H + \delta(D^hu). \end{aligned}$$

Hence (v) holds for $u \in \mathcal{S}_H$.

Next, consider (iv). Let $\{e_k\}_{k\in\mathbb{N}} \subset H$ be an ON-basis and $u, v \in \mathcal{S}_H$. We use duality, (v) for $u \in \mathcal{S}_H$ and some analysis.

$$\begin{aligned} \mathbb{E}[\delta(u)\delta(v)] &= \mathbb{E}[\langle u, D\delta(v)\rangle_H] = \mathbb{E}\Big[\Big\langle \sum_{i=1}^{\infty}\langle u, e_i\rangle_H e_i, D(\delta(v))\Big\rangle_H\Big] \\ &= \mathbb{E}\Big[\sum_{i=1}^{\infty}\langle u, e_i\rangle_H D^{e_i}(\delta(v))\Big] = \sum_{i=1}^{\infty}\mathbb{E}[\langle u, e_i\rangle_H(\langle v, e_i\rangle_H + \delta(D^{e_i}v))] \\ &= \mathbb{E}[\langle u, v\rangle_H] + \sum_{i=1}^{\infty}\mathbb{E}[\langle D\langle u, e_i\rangle_H, D^{e_i}v\rangle_H] \end{aligned}$$

The second term is, using 10.4,

$$\mathbb{E}\Big[\sum_{i=1}^{\infty}\langle (Du)^* e_i, (Dv)e_i\rangle_H\Big] = \mathbb{E}\Big[\sum_{i=1}^{\infty}\langle (DuDv)e_i, e_i\rangle_H\Big] = \mathbb{E}[\mathrm{Tr}(Du\circ Dv)].$$

We have proved property (iv) for $u, v \in \mathcal{S}_H$. Notice that $\mathrm{Tr}(AB) = \langle A^*, B\rangle_{H\otimes H}$. Then by Cauchy-Schwarz inequality $\mathrm{Tr}(Du\circ Du) \leq \|Du\|^2_{H\otimes H}$. It follows that

$$\mathbb{E}[\delta(u)^2] \leq \mathbb{E}[\|u\|^2_H] + \mathbb{E}[\|Du\|^2_{H\otimes H}] = \|u\|^2_{1,2,H}. \tag{10.5}$$

Hence, $\mathbb{D}^{1,2} \subset \mathrm{Dom}(\delta)$. From the construction of $\mathcal{S}_H$ there exists, for every $u \in \mathbb{D}^{1,2}(H)$, a sequence $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{S}_H$ such that $u_n \to u$ in $\mathbb{D}^{1,2}(H)$. But then, by (10.5), $\{\delta(u_n)\}_{n\in\mathbb{N}}$ converges. Call the limit $\delta(u)$. By approximating with elements of $\mathcal{S}_H$ one can prove formula (iv) for $u, v \in \mathbb{D}^{1,2}$.

Now, consider (v). Let $u \in \mathbb{D}^{1,2}(H)$ with $D^h u \in \mathrm{Dom}(\delta)$ and $\{e_i\}_{i\in\mathbb{N}} \subset H$ be an orthonormal basis. For all $F \in \mathbb{D}^{1,2}$ and $h \in H$, we have, by (iv) that,

$$\mathbb{E}[\delta(u)\delta(hF)] = \mathbb{E}[\langle u, h\rangle_H F] + \mathbb{E}[\mathrm{Tr}(Du\circ D(hF)]$$

We calculate and then use duality on the second term.

$$\begin{aligned}
\mathbb{E}[\mathrm{Tr}(Du\circ D(hF)] &= \mathbb{E}\Big[\sum_{i=1}^{\infty}\langle Du\circ(h\otimes DF)e_i, e_i\rangle_H\Big] \\
&= \mathbb{E}\Big[\sum_{i=1}^{\infty}\langle (Du)h, e_i\rangle_H \langle DF, e_i\rangle_H\Big] \\
&= \mathbb{E}[\langle D^h u, DF\rangle_H] \\
&= \mathbb{E}[\delta(D^h u)F].
\end{aligned}$$

Hence, we have that, for all $F \in \mathbb{D}^{1,2}$, that

$$\mathbb{E}[\delta(u)\delta(hF)] = \mathbb{E}[(\langle u, h\rangle_H + \delta(D^h u))F].$$

Then, by Lemma 10.2 for $G = \delta(u)$ and $Y = \langle u, h\rangle_H + \delta(D^h u)$ it follows that

$$D^h\delta(u) = \langle u, h\rangle_H + \delta(D^h u)$$

and $\delta(u) \in \mathbb{D}^{h,2}$. The proof is now complete. □

The following proposition will be a useful tool for factoring out scalar random variables.

PROPOSITION 10.3. *If $F \in \mathbb{D}^{1,2}$ and $u \in$ Dom(δ) such that $uF \in L_2(\Omega, H)$, then $Fu \in$ Dom(δ). Moreover*

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

holds given that the right hand side belongs to $L_2(\Omega)$.

When $u = h \in H$, then by Proposition 10.1 (iii) we have that

$$\delta(Fh) = FI(h) - D^h F.$$

This result is true for $F \in \mathbb{D}^{h,2}$, see Nualart [**?**].

PROOF. Let $G \in \mathcal{S}_0$, i.e., $G \in \mathcal{S}$ and with compact support. By duality and the product rule

$$\begin{aligned}
\mathbb{E}[G\delta(Fu)] &= \mathbb{E}[\langle FDG, u\rangle_H] \\
&= \mathbb{E}[\langle D(FG) - GDF, u\rangle_H] \\
&= \mathbb{E}[G(F\delta(u) - \langle DF, u\rangle_H)].
\end{aligned}$$

Since $\mathcal{S}_0 \subset L_2(\Omega)$ is dense the result follows. $\square$

11. Malliavin Calculus for Cylindrical Wiener Process

Let U be a real separable Hilbert space. The process $\{W(t)\}_{t\in[0,T]}$ is a U-valued cylindrical Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$. The Itô integral is well defined for predictable, Hilbert-Schmidt valued integrands $\{\Phi(t)\}_{t\in[0,T]}$ satisfying

$$\mathbb{E}\Big[\int_0^T \|\Phi(s)\|^2_{L_2(U)}\,\mathrm{d}s\Big] < \infty.$$

Here, the Hilbert-Schmidt norm for an operator S is given by

$$\|S\|^2_{L_2(U)} = \sum_{k=1}^{\infty} \|Se_k\|^2_U,$$

where $\{e_k\}_{k\in\mathbb{N}}$ is an orthonormal basis for U.

We now want to develop Malliavin calculus for this setup. For $\Phi_1, \dots, \Phi_n \in L_2([0,T], L_2(U))$ and $f \in C_p^\infty(U^n, U)$ we would like the following "smooth" chain rule to be valid.

$$\begin{aligned}
&D_s f\Big(\int_0^T \Phi_1\,\mathrm{d}W, \dots, \int_0^T \Phi_n\,\mathrm{d}W\Big) \\
&\quad = \sum_{i=1}^n \partial_i f\Big(\int_0^T \Phi_1\,\mathrm{d}W, \dots, \int_0^T \Phi_n\,\mathrm{d}W\Big)\Phi_i(s),
\end{aligned} \tag{11.1}$$

Here $\partial_i f$ is the Frechèt derivative in the i:th variable. But, the U-valued Itô integral neither is an isonormal Gaussian process nor has a density. Consequently we cannot use our machinery from earlier directly or prove an integration by parts formula such as Lemma 4.3 for this case. We need an other starting point.

Let $\{e_k\}_{k\in\mathbb{N}} \subset U$ be an ON-basis and, for $m \in \mathbb{N}$, P_m be the projection onto the subspace spanned by $\{e_1, \dots, e_m\}$. For $\Phi \in L_2([0,T], L_2(U))$ define $\Phi_m = P_m \Phi P_m$. For simplicity we consider $F = f(\int_0^T \Phi_m\,\mathrm{d}W)$, i.e. for $n = 1$. Then,

$$
\begin{aligned}
F &= f\Big(\int_0^T \Phi_m \,\mathrm{d}W\Big) = f\Big(\sum_{k=1}^m \int_0^T \langle \Phi_m^* e_k,\, \mathrm{d}W\rangle_U e_k\Big) \\
&= \sum_{l=1}^\infty \Big\langle f\Big(\sum_{k=1}^m \int_0^T \langle \Phi_m^* e_k,\, \mathrm{d}W\rangle_U e_k\Big), e_l\Big\rangle_U e_l \qquad (11.2)\\
&:= \sum_{l=1}^\infty \tilde f_l\Big(\int_0^T \langle \Phi_m^* e_1,\, \mathrm{d}W\rangle_U, \dots, \int_0^T \langle \Phi_m^* e_m,\, \mathrm{d}W\rangle_U\Big) e_l
\end{aligned}
$$

Here $\tilde f_l \in C_p^\infty(\mathbb{R}^m, \mathbb{R})$. This motivates why to start with smooth random variables of the type

$$
F = f\Big(\int_0^T \langle \phi_1,\, \mathrm{d}W\rangle_U, \dots, \int_0^T \langle \phi_n,\, \mathrm{d}W\rangle_U\Big),
$$

where $\phi_1, \dots, \phi_n \in L_2([0,T];U)$, and $f \in C_p^\infty(\mathbb{R}^n;\mathbb{R})$. Clearly

$$
W(\phi) := \int_0^T \langle \phi,\, \mathrm{d}W\rangle_U, \quad \phi \in L_2([0,T],U) := H
$$

is an isonormal Gaussian process, indeed for $l_u := \langle u, \cdot\rangle_U$, the Itô isometry gives that

$$
\begin{aligned}
\mathbb{E}[W(\phi_1)W(\phi_2)] &= \mathbb{E}\Big[\int_0^T \langle \phi_1(s),\, \mathrm{d}W(s)\rangle_U \int_0^T \langle \phi_2(s),\, \mathrm{d}W(s)\rangle_U\Big] \\
&= \mathbb{E}\Big[\int_0^T l_{\phi_1(s)} \,\mathrm{d}W(s) \int_0^T l_{\phi_2(s)} \,\mathrm{d}W(s)\Big] \\
&= \int_0^T \langle l_{\phi_1(s)}, l_{\phi_2(s)}\rangle_{L_2(U,\mathbb{R})} \,\mathrm{d}s \\
&= \int_0^T \sum_{k=1}^\infty \langle \phi_1(s), e_k\rangle_U \langle \phi_2(s), e_k\rangle_U \,\mathrm{d}s \\
&= \int_0^T \langle \phi_1(s), \phi_2(s)\rangle_U \,\mathrm{d}s \\
&= \langle \phi_1, \phi_2\rangle_H.
\end{aligned}
$$

The Malliavin calculus developed in previous sections thus applies here. In the rest of this section we will adopt the notations from earlier to this specific setting. Let F be of the form as in (11.2). Then

$$
\begin{aligned}
D_s F &= \sum_{l=1}^\infty \sum_{i=1}^m \partial_i\Big(\Big\langle f\Big(\sum_{k=1}^m \int_0^T \langle \Phi^*(s) e_k,\, \mathrm{d}W(s)\rangle_U e_k\Big), e_l\Big\rangle_U\Big) e_l \otimes (\Phi^*(s) e_i) \\
&= \sum_{l=1}^\infty \sum_{i=1}^m \Big\langle f'\Big(\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big), e_l\Big\rangle_U e_l \otimes (\Phi^*(s) e_i)
\end{aligned}
$$

Differently from before we define $D_s^h F = D_s F h$ for $h \in U$. For a complete analogue we should have $D^h = \langle DF, h\rangle_{L_2(0,T,U)}$, $h \in L_2([0,T],U)$. This will never be used. We continue with $D_s^h F$.

$$\begin{aligned} D_s^h F &= \sum_{i=1}^m \sum_{l=1}^\infty \partial_i \Big\langle f\big(\sum_{k=1}^m \int_0^T \langle \Phi^*(s)e_k,\, \mathrm{d}W(s)\rangle_U e_k\Big), e_l\Big\rangle_U e_l \langle \Phi^*(s)e_i, h\rangle_U \\ &= \sum_{i=1}^m f'\Big(\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big) e_i \langle e_i, \Phi(s)h\rangle_U \\ &= f'\Big(\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big)\Phi(s)h \end{aligned}$$

This implies that

$$D_s F = f'\Big(\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big)\Phi(s),$$

as we wanted in (11.1).

Let us define $\mathbb{D}^{1,2}(U)$ to be the random variables of the form

$$G = \sum_{k=1}^\infty F_k e_k \in L_2(\Omega, U),$$

where $\{e_k\}_{k\in\mathbb{N}} \subset U$ is an ON-basis and $F_i \in \mathbb{D}^{1,2}$, $i = 1, 2, \ldots$ which moreover satisfy

$$\begin{aligned} \mathbb{E}\Big[\int_0^T \|D_s G\|_{L_2(U)}\,\mathrm{d}s\Big] &= \mathbb{E}\Big[\int_0^T \Big\|\sum_{k=1}^\infty e_k \otimes D_s F_k\Big\|_{L_2(U)}^2 \mathrm{d}s\Big] \\ &= \mathbb{E}\Big[\int_0^T \sum_{n=1}^\infty \sum_{k=1}^\infty \|D_s F_k\|_U^2 \langle e_k, e_n\rangle_U\Big]\,\mathrm{d}s \\ &= \sum_{k=1}^\infty \int_0^T \mathbb{E}[\|D_s F_k\|_U^2]\,\mathrm{d}s < \infty \end{aligned}$$

Lemma 11.1. *Assume that $F \in \mathbb{D}^{1,2}(U)$ and let $\Phi \in L^2(\Omega\times[0,T], \mathcal{L}_2(U))$ be an adapted process. Then*

$$\mathbb{E}\Big\langle F, \int_0^T \Phi(s)\, dW(s)\Big\rangle = \mathbb{E}\int_0^T Tr\big\{\Phi^*(s) D_s F\big\}\, ds.$$

Proof. Let $\{e_k\}_{k\in\mathbb{N}} \subset U$ be an ON-basis. We use Fubini's theorem and Malliavin integration by parts.

$$
\begin{aligned}
\mathbb{E}\Big\langle F, \int_0^T \Phi(s)\,\mathrm{d}W(s)\Big\rangle &= \mathbb{E}\Big\langle F, \sum_{j\in\mathbb{N}} \int_0^T \langle \Phi^*(s)e_j,\, \mathrm{d}W(s)\rangle e_j\Big\rangle \\
&= \sum_{j\in\mathbb{N}} \mathbb{E}\langle F, e_j\rangle_H \int_0^T \langle \Phi^*(s)e_j,\, \mathrm{d}W(s)\rangle_H \\
&= \sum_{j\in\mathbb{N}} \mathbb{E} \int_0^T \langle D_s\langle F, e_j\rangle_H, \Phi^*(s)e_j\rangle_H \,\mathrm{d}s \\
&= \mathbb{E} \int_0^T \sum_{j\in\mathbb{N}} \langle (D_sF)^* e_j, \Phi^*(s)e_j\rangle_H \,\mathrm{d}s \\
&= \mathbb{E} \int_0^T \mathrm{Tr}\{\Phi^*(s)D_sF\}\,\mathrm{d}s.
\end{aligned}
$$

Here we used that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

$\square$

Bibliography