Chapter 2
Gradient and divergence

2.1 Gradient

If our objective is to define a differential calculus on the Banach space W, why don't we use the notion of Fréchet derivative: A function $F : W \to \mathbf{R}$ is said to be Fréchet differentiable if there exists a continuous linear operator $A : W \to W$ such that

$$\lim_{\epsilon \to 0} \epsilon^{-1} \|F(\omega + \epsilon\omega') - F(\omega) - A\omega'\|_W = 0 \tag{2.1}$$

for any $\omega \in W$ and any $\omega' \in W$. In particular, a Fréchet differentiable function is continuous. One of the most immediate function we can think of is the so-called Itô map which sends a sample-path ω to the corresponding sample-path of the solution of a well defined stochastic differential equation. It is well known (see [Lej09, Section 3.3] for instance) that in dimension higher than one, this map is not continuous. This induces that the notion of Fréchet derivative is not well suited to a differential calculus on the Wiener space. Moreover, since we work on a probability space, measurable functions F from W into $\mathbf{R}$ are random variables, meaning that they are defined up to a negligible set. To avoid any inconsistency in a formula like (2.1), we must ensure that

$$(F = G \ \mu \text{ a.s.}) \Longrightarrow (F(. + \omega') = G(. + \omega') \ \mu \text{ a.s.})$$

for any ω'. With the notations of Theorem 1.8, this requires that $T^{\#}_{\omega'}\mu$ (the pushforward of the measure μ by the translation map $T_{\omega'}$) to be absolutely continuous with respect to the Wiener measure μ. This fact is granted only if ω' belongs to $I_{1,2}$. These two reasons mean that we are to define the directional derivative of F in a restricted class of possible perturbations.

Recall the diagram

$$\begin{array}{ccccc} \mathcal{W}^* & \xrightarrow{\mathfrak{e}^*} & \mathcal{H}^* = (I_{1,2})^* & & \\ & & \downarrow \simeq & & \\ L^2 & \xrightarrow{I^1} & \mathcal{H} = I_{1,2} & \xrightarrow{\mathfrak{e}} & \mathcal{W} \end{array}$$

and that μ is the Wiener measure on $\mathcal{W}$.

Definition 2.1. A function F is said to be cylindrical if there exists an integer n, $f \in \text{Schwartz}(\mathbf{R}^n)$, the Schwartz space on $\mathbf{R}^n$, $(h_1, \cdots, h_n) \in \mathcal{H}^n$ such that

$$F(\omega) = f(\delta h_1, \cdots, \delta h_n).$$

The set of such functionals is denoted by $\mathcal{S}$.

Theorem 2.1. *The set $\mathcal{S}$ is dense in $L^p(\mathrm{W} \to \mathbf{R};\, \mu)$.*

Proof. Let $\mathcal{D}_n$ be the dyadic subdivision of mesh 2^{-n} of $[0,1]$ and $\mathcal{F}_n = \sigma\{B(t), t \in \mathcal{D}_n\}$. Since B has continuous sample-paths, for any $t \in [0,1]$, there exists a sequence $(t_n, n \geq 0)$ such that $t_n \in \mathcal{D}_n$ for any n and $t_n \to t$. Hence, $\vee_n \mathcal{F}_n = \mathcal{F}$ and the L^p convergence theorem for martingales says that

$$\mathbf{E}\left[F \,|\, \mathcal{F}_n\right] \xrightarrow[L^p]{n\to\infty} F.$$

For $\epsilon > 0$, let n such that $\|F - \mathbf{E}\left[F \,|\, \mathcal{F}_n\right]\|_{L^p} < \epsilon$. The Doob Lemma entails that there exists ϕ_n measurable from $\mathbf{R}^{2^n}$ to $\mathbf{R}$ such that

$$\mathbf{E}\left[F \,|\, \mathcal{F}_n\right] = \psi_n(B(t), t \in \mathcal{D}_n)$$

where $t_k^n = k2^{-n}$. Let μ_n be the distribution of the Gaussian vector $(B(t), t \in \mathcal{D}_n)$,

$$\int |\psi_n|^p \,\mathrm{d}\mu_n = \mathbf{E}\left[|\mathbf{E}\left[F \,|\, \mathcal{F}_n\right]|^p\right] \leq \mathbf{E}\left[|F|^p\right] < \infty.$$

That means that ψ_n belongs to $L^p(\mathbf{R}^n \to \mathbf{R};\, \mu_n)$ hence for any $\epsilon > 0$, there exists $\varphi_\epsilon \in \mathcal{S}(\mathbf{R}^{2^n})$ such that $\|\psi_n - \varphi_\epsilon\|_{L^p(\mathbf{R}^n \to \mathbf{R};\, \mu_n)} < \epsilon$. Then, $\varphi_\epsilon(B(t), t \in \mathcal{D}_n)$ belongs to $\mathcal{S}$ and is within distance 2ϵ of F in $L^p(\mathcal{W} \to \mathbf{R};\, \mu)$.

The gradient is first defined on cylindrical functionals.

Definition 2.2. Let $F \in \mathcal{S}$, $h \in \mathcal{H}$, with $F(\omega) = f(\delta h_1, \cdots, \delta h_n)$. Set

$$\nabla F = \sum_{j=1}^{n} \partial_j F(\delta h_1, \cdots, \delta h_n)\ h_j,$$

so that

$$\langle \nabla F, h \rangle_{\mathcal{H}} = \sum_{j=1}^{n} \partial_j F(\delta h_1, \cdots, \delta h_n)\ \langle h_j,\, h \rangle_{\mathcal{H}}\,.$$

This definition is coherent with the natural definition of directional derivative.

Lemma 2.1.

$$\langle \nabla F,\, h\rangle_{\mathcal{H}} = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big(F(\omega+\epsilon h)-F(\omega)\Big).$$

Proof. Due to the fact that $\delta h_i(\omega+h)=\delta h_i(\omega)+\langle h,\, h_i\rangle_{\mathcal{H}}$.

Example 2.1. For any $h\in\mathcal{H}$,

$$\nabla_h B(t)=\langle h, t\wedge .\rangle=\int_0^t \dot h(s)\ \mathrm{d}s=\int_0^1 \mathbf{1}_{[0,t]}(s)\,\dot h(s)\ \mathrm{d}s.$$

Remark that $\mathcal{S}$ is an algebra for the ordinary product.

Lemma 2.2. *For $F\in\mathcal{S}$, $\phi\in\mathcal{C}^\infty$*

$$\begin{aligned}\nabla_h F(\omega) &= \frac{d}{d\epsilon}F(\omega+\epsilon h)\Big|_{\epsilon=0}\\ \nabla(FG) &= F\,\nabla G + G\,\nabla F\\ \nabla\phi(F) &= \phi'(F)\,\nabla F.\end{aligned}$$

Before going futher, we give/recall some elements about tensor products of Banach spaces. Let X and Y two Banach spaces, with respective dual X^* and Y^*. For $x\in X$ and $y\in Y$, $x\otimes y$ is the bilinear form defined by:

$$\begin{aligned}x\otimes y\ :\ X^*\times Y^* &\longrightarrow \mathbf{R}\\ (\eta,\zeta) &\longmapsto \langle \eta, x\rangle_{X^*,X}\,\langle\zeta, y\rangle_{Y^*,Y}\,.\end{aligned}$$

Definition 2.3 (See [Rya02, chapter 2]). The *projective* tensor product of X and Y is the completion of the vector space spanned by the finite linear combinations of some $x\otimes y$ for $x\in X$ and $y\in Y$, equipped with the norm

$$\|z\|_{X\otimes Y}=\inf\left\{\sum_{i=1}^n \|x_i\|_X\,\|y_i\|_Y,\ z=\sum_{i=1}^n x_i\otimes y_i\right\}.$$

We need to take the infimum of all the possible representations of z as a linear combinations of elementary tensor products since such a representation is by no means unique.

Example 2.2. One of the simplest situation we can imagine, is the tensor product of $L^1(\mathbf{R})$ by itself. The function

$$\mathbf{1}_{[0,1]}(s)\otimes\mathbf{1}_{[0,2]}(t)+\mathbf{1}_{[1,2]}(s)\otimes\mathbf{1}_{[1,2]}(t)$$

can be equally written as:

$$\mathbf{1}_{[0,1]}(s) \otimes \mathbf{1}_{[0,1]}(t) + \mathbf{1}_{[0,2]}(s) \otimes \mathbf{1}_{[1,2]}(t).$$

Proposition 2.1. *For X and Y two reflexive Banach spaces, i.e. $(X^*)^* = X$. The dual of $W = X \otimes Y$ is the space $W^* = X^* \otimes Y^*$ with the duality pairing:*

$$\langle w^*, w\rangle_{W*,W} = \sum_{i,j} \langle x_i^*, x_j\rangle_{X*,X}\ \langle y_i^*, y_j\rangle_{Y*,Y}$$

where $w = \sum_j x_j \otimes y_j \in X \otimes Y$ and $w^* = \sum_i x_i^* \otimes y_i^* \in X^* \otimes Y^*$. Moreover,

$$\begin{aligned}\|w^*\|_{W^*} &= \sup_{\|w\|_W=1} \left|\langle w^*, w\rangle_{W^*,W}\right| \\ &= \sup\Big\{\left|\langle w^*, x \otimes y\rangle_{W^*,W}\right|,\ \|x\|_X = 1,\ \|y\|_Y = 1\Big\}. \quad (2.2)\end{aligned}$$

Let X be a Banach space and ν a measure on a space E. The set $L^p(E; X, \nu)$ is the space of functions ψ from E into X such that

$$\int_E \|\psi(x)\|_X^p \ \mathrm{d}\nu(x) < \infty.$$

Theorem 2.2 ([Rya02, page 30]). *For X a Banach space, the space $L^p(E \to \mathbf{R};\, \nu) \otimes X$ is isomorphic to $L^p(E \to X;\, \nu)$.*

Moreover, if $X = L^p(F \to \mathbf{R};\, \rho)$ then $L^p(E \to X;\, \nu)$ is isometrically isomorphic to $L^p(E \times F \to \mathbf{R};\, \nu \otimes \rho)$. Moreover, the set of simple functions, i.e. functions of the form

$$\sum_{j=1}^n f_j(s)\psi_j(x)$$

where $f_j \in L^p(E \to \mathbf{R};\, \nu)$ and $\psi_j \in L^p(F \to \mathbf{R};\, \rho)$, is dense into $L^p(E \times F \to \mathbf{R};\, \nu \otimes \rho)$.

Theorem 2.3. *For $F \in \mathcal{S}$, ∇F belongs to $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ for any $p \geq 1$.*

Proof. Step 1. Assume $p > 1$. Since

$$L^p(\mathcal{W} \to \mathcal{H};\, \mu) \simeq L^p(\mathcal{W} \to \mathbf{R};\, \mu) \otimes \mathcal{H},$$

we have

$$\big(L^p(\mathcal{W} \to \mathcal{H};\, \mu)\ \big)^* \simeq L^q(\mathcal{W} \to \mathbf{R};\, \mu) \otimes \mathcal{H}$$

where $q = p/(p-1)$.

Step 2. Consider the set

$$B_{q,\mathcal{H}} = \{(k,\, G) \in \mathcal{H} \times L^q(\mathcal{W} \to \mathbf{R};\, \mu),\ \|k\|_{\mathcal{H}} = 1,\ \|G\|_{L^q(\mathcal{W}\to\mathbf{R};\,\mu)} = 1\}.$$

Let $F = f(\delta h)$, for $p > 1$, in view of Proposition 2.1, we have to prove that

$$\sup_{(k,G)\in B_{q,\mathcal{H}}} \left| \langle \nabla F, k \otimes G \rangle_{L^p(\mathcal{W}\to\mathcal{H};\,\mu),\, L^q(\mathcal{W}\to\mathcal{H};\,\mu)} \right| < \infty.$$

By the very definition of the duality bracket,

$$\begin{aligned} \left| \langle \nabla F, k \otimes G \rangle_{L^p(\mathcal{W}\to\mathcal{H};\,\mu),\, L^q(\mathcal{W}\to\mathcal{H};\,\mu)} \right| &= \left| \mathbf{E}\left[\langle \nabla F,\, k \rangle_{\mathcal{H}}\, G \right] \right| \\ &= \left| \mathbf{E}\left[f'(\delta h) G \right] \langle k,\, h \rangle_{\mathcal{H}} \right| \\ &\le \| f' \|_\infty \| G \|_{L^q(\mathcal{W}\to\mathbf{R};\,\mu)} \| h \|_{\mathcal{H}} \| k \|_{\mathcal{H}}. \end{aligned}$$

Hence the supremum over $B_{q,\mathcal{H}}$ is finite. The same proof can be applied when $F = f(\delta h_j,\, 1 \le j \le m)$.
Step 2. For $p = 1$, the previous considerations no longer prevail since an L^1 space is not reflexive so that we cannot apply (2.2). However, it is sufficient to see that $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ is included in $L^1(\mathcal{W} \to \mathcal{H};\, \mu)$.

It is an exercise left to the reader to see that the map

$$\begin{aligned} \operatorname{Id} \otimes I^{-1} \,:\, L^p(\mathcal{W} \to \mathbf{R};\, \mu) \otimes \mathcal{H} &\longrightarrow L^p(\mathcal{W} \to \mathbf{R};\, \mu) \otimes L^2([0,1] \to \mathbf{R};\, \lambda) \\ F \otimes h &\longmapsto F \otimes \dot{h} \end{aligned}$$

is continuous. Moreover, Theorem 2.3 means for any $F \in \mathcal{S}$, ∇F belongs to $L^p(\mathcal{W} \to \mathbf{R};\, \mu) \otimes \mathcal{H}$. Hence there exists an element $\dot{\nabla} F$ of $L^p(\mathcal{W} \to \mathbf{R};\, \mu) \otimes L^2([0,1] \to \mathbf{R};\, \lambda)$ such that

$$\langle \nabla F,\, h \rangle_{\mathcal{H}} = \int_0^1 \dot{\nabla}_s F\, \dot{h}(s)\, \mathrm{d}s$$

$$\text{and } \| F \|_{L^p(\mathcal{W}\to\mathcal{H};\,\mu)} = \mathbf{E}\left[\left(\int_0^1 |\dot{\nabla}_s F|^2\, \mathrm{d}s \right)^{p/2} \right]^{1/p}.$$

Proposition 2.2 ([Yos95, page 77]). *A map T is closable if and only if*

$$(x_n, n \ge 0) \in \operatorname{Dom} T,\ \lim_n x_n = 0 \text{ and } \lim_n T x_n = y \Longrightarrow y = 0.$$

Theorem 2.4. *∇ is closable in $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ for $p > 1$.*

The proof is based on the following lemma which is crucial for the sequel.

Lemma 2.3 (Integration by parts). *For F and G cylindrical, for $h \in \mathcal{H}$,*

$$\mathbf{E}\left[G \langle \nabla F,\, h \rangle_{\mathcal{H}} \right] = -\mathbf{E}\left[F\, \langle \nabla G,\, h \rangle_{\mathcal{H}} \right] + \mathbf{E}\left[FG\, \delta h \right]. \tag{2.3}$$

Proof. The Cameron-Martin theorem says that

$$\mathbf{E}\left[F(\omega+\epsilon h)G(\omega+\epsilon h)\right]=\mathbf{E}\left[F(\omega)G(\omega)\exp\left(\delta h-\frac{1}{2}\|h\|_{\mathcal{H}}^2\right)\right]$$

Differentiate both sides with respect to ϵ, at $\epsilon=0$, to obtain

$$\mathbf{E}\left[F\left\langle\nabla G,h\right\rangle_{\mathcal{H}}\right]+\mathbf{E}\left[G\left\langle\nabla F,h\right\rangle_{\mathcal{H}}\right]=\mathbf{E}\left[FG\,\delta h\right],$$

which corresponds to Eqn. (2.3).

Proof (Proof of Theorem 2.4). If F_n tends to 0 in $L^p(\mathcal{W}\to\mathbf{R};\,\mu)$ then the right-hand-side of Eqn. (2.3) tends to 0. On the other hand, by definition of the convergence in $L^p(\mathcal{W}\to\mathcal{H};\,\mu)$,

$$\mathbf{E}\left[G\left\langle\nabla F_n,\,h\right\rangle_{\mathcal{H}}\right]\xrightarrow{n\to\infty}\left\langle\eta,\,h\otimes G\right\rangle_{L^p(\mathcal{W}\to\mathbf{R};\,\mu),\,L^q(\mathcal{W}\to\mathbf{R};\,\mu)}.$$

It means that for any $h\in\mathcal{H}$ and $G\in\mathcal{S}$,

$$\left\langle\eta,\,h\otimes G\right\rangle_{L^p(\mathcal{W}\to\mathbf{R};\,\mu),\,L^q(\mathcal{W}\to\mathbf{R};\,\mu)}=0. \tag{2.4}$$

By density of $\mathcal{S}$ in $L^p(\mathcal{W}\to\mathbf{R};\,\mu)$, (2.4) holds for $G\in L^p(\mathcal{W}\to\mathbf{R};\,\mu)$. According to Theorem 2.2, $\langle\eta,\zeta\rangle_{L^p(W;\mathcal{H}),L^q(\mathcal{W}\to\mathcal{H};\,\mu)}=0$ for any $\zeta\in L^q(\mathcal{W}\to\mathcal{H};\,\mu)$, hence $\eta=0$.

Definition 2.4 (Gross-Sobolev derivative). A functional F belongs to $\mathbb{D}_{p,1}$ if there exists $(F_n,\,n\ge 0)$ which converges to $F\in L^p(\mathcal{W}\to\mathbf{R};\,\mu)$, such that $(\nabla F_n,\,n\ge 0)$ is Cauchy in $L^p(W;\mathcal{H})$. Then, ∇F is defined as the limit of this sequence.

Remark 2.1. The closability of ∇ ensures that the limit does not depend on the approximating sequence.

Lemma 2.4. *Let $p>1$. Assume that there exists $(F_n,\,n\ge 0)$ which converges in $L^p(\mathcal{W}\to\mathbf{R};\,\mu)$ to F such that $\sup_n\|\nabla F_n\|_{L^p(\mathcal{W}\to\mathcal{H};\,\mu)}$ is finite. Then, $F\in\mathbb{D}_{p,1}$.*

For this result, we need to invoke two theorems of functional analysis.

Definition 2.5 (Weak convergence). A sequence $(x_n,\,n\ge 0)$ is said to be weakly convergent in a Banach space X, if for every $\eta\in X^*$, $(\langle\eta,x_n\rangle_{X^*,X}\,,\,n\ge 0)$ is convergent.

Remark 2.2. Since $|\langle\eta,x_n-x\rangle_{X^*,X}|\le\|\eta\|_{X^*}\|x_n-x\|_X$, strong convergence implies weak convergence but the converse is false. For instance, let $(e_n,\,n\ge 0)$ a complete orthonormal basis in a Hilbert space X, on the one hand $\|e_n\|_X=1$. On the other hand, according to Parseval equality, for $\eta\in X^*=X$, $\|\eta\|_X^2=\sum_n|\langle\eta,e_n\rangle_X|^2$. Hence, $(\langle\eta,x_n\rangle_{X^*,X}\,,\,n\ge 0)$ converges weakly to 0. The convergence cannot hold in the strong sense.

Proposition 2.3 (Eberlein-Shmulyan,[Yos95, page 141]). *Let X be a reflexive Banach space, i.e. $(X^*)^* = X$. Then, any strongly bounded sequence admits a weakly convergent subsequence.*

Remark 2.3. For any measure, L^p spaces are reflexive only for $p \neq 1, \infty$. We do have that the dual of L^1 is L^∞ but the dual of L^∞ is larger than L^1.

Proposition 2.4 (Mazur, [Yos95, page 120]). *Let $(x_n,\, n \geq 0)$ be a weakly convergent subsequence in a Banach space X and set x its limit. Then, for any $\epsilon > 0$, there exist n and $(\alpha_i,\, 1 \leq i \leq n)$ such that $\alpha_i \geq 0$, $\sum_i \alpha_i = 1$ and*

$$\| \sum_{i=1}^{n} \alpha_i x_{n_i} - x \|_X \leq \epsilon.$$

Proof (Proof of Lemma 2.4). Since $\sup_n \|\nabla F_n\|_{L^p(\mathcal{W} \to \mathcal{H};\, \mu)}$ is finite, there exists a subsequence (see Proposition 2.3) which we still denote by $(\nabla F_n, n \geq 0)$ weakly convergent in $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ to some limit denoted by η. For $k > 0$, let n_k be such that $\|F_m - F\|_{L^p} < 1/k$ for $m \geq n$. The Mazur's Theorem 2.4 implies that there exists a convex combination of elements of $(\nabla F_m, m \geq n_k)$ such that

$$\| \sum_{i=1}^{M_k} \alpha_i^k \nabla F_{m_i} - \eta \|_{L^p(\mathcal{W} \to \mathcal{H};\, \mu)} < 1/k.$$

Moreover, since the α_i^k are positive and sums to 1,

$$\| \sum_{i=1}^{M_k} \alpha_i^k F_{m_i} - F \|_{L^p} \leq 1/k.$$

We have thus constructed a sequence

$$F^k = \sum_{i=1}^{M_k} \alpha_i^k F_{m_i}$$

such that F^k tends to F in L^p and ∇F^k converges in $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ to a limit. By the construction of $\mathbb{D}_{p,1}$, this means that F belongs to $\mathbb{D}_{p,1}$ and that $\nabla F = \eta$.

Definition 2.6. The space $\mathbb{D}_{p,1}$ is the closure of $\mathcal{S}$ for the norm

$$\|F\|_{p,1} = \mathbf{E}\left[|F|^p\right]^{1/p} + \mathbf{E}\left[\|\nabla F\|_{\mathcal{H}}^p\right]^{1/p}.$$

Corollary 2.1. *Let F belong to $\mathbb{D}_{p,1}$ and G to $\mathbb{D}_{q,1}$ with $q = p/(p-1)$. If $h \in \mathcal{H}$, then Eqn. (2.3) holds:*

$$\mathbf{E}\left[G \left\langle \nabla F,\, h \right\rangle_{\mathcal{H}}\right] = -\mathbf{E}\left[F \,\left\langle \nabla G,\, h \right\rangle_{\mathcal{H}}\right] + \mathbf{E}\left[FG\, \delta h\right].$$

Proof. According to Lemma 2.3, it is true for F and G in $\mathcal{S}$. Let $(F_n, n \geq 0)$ a sequence of elements of $\mathcal{S}$ converging to F in $\mathbb{D}_{p,1}$. Since G belongs to $\mathcal{S}$, G and $\nabla_h G$ belong to $L^q(\mathcal{W} \to \mathbf{R};\, \mu)$. By Hölder inequality, we see that (2.3) holds for $F \in \mathbb{D}_{p,1}$ and $G \in \mathcal{S}$. Repeat the same approach with an approximation of $G \in \mathbb{D}_{q,1}$ by elements of $\mathcal{S}$.

We can now generalize the basic formulas to elements of $\mathbb{D}_{p,1}$.

Theorem 2.5. *For $F \in \mathbb{D}_{p,1}$ and $G \in \mathbb{D}_{q,1}$ (with $1/p + 1/q = 1/r$ for $r > 1$), for ϕ Lipschitz continuous, the product FG belongs to $\mathbb{D}_{r,1}$ and*

$$\begin{aligned}\nabla(FG) &= F\,\nabla G + G\,\nabla F\\ \nabla\phi(F) &= \phi'(F)\,\nabla F.\end{aligned}$$

For the next theorem, we need to introduce the two families of projections: For any $t \in [0,1]$

$$\begin{aligned}\pi_t \,:\, \mathcal{H} &\longrightarrow \mathcal{H} \\ h = I^1(\dot h) &\longmapsto I^1(\dot h \mathbf{1}_{[0,t]})\end{aligned} \qquad\qquad \begin{aligned}\dot\pi_t \,:\, L^2 &\longrightarrow L^2 \\ \dot h &\longmapsto \dot h \mathbf{1}_{[0,t]}.\end{aligned}$$

We have

$$\|\pi_t h\|^2_{\mathcal{H}} = \int_0^1 \dot h(s)^2 \mathbf{1}_{[0,t]}(s)\, \mathrm{d}s \leq \|\dot h\|^2_{L^2} = \|h\|^2_{\mathcal{H}},$$

meaning that π_t is a continuous contraction on $\mathcal{H}$. Moreover,

$$\pi_t(s \wedge .) = I^1\big(\dot\pi_t(\mathbf{1}_{[0,s]})\big) = I^1\big(\mathbf{1}_{[0,s]}\mathbf{1}_{[0,t]})\big) = I^1\big(\mathbf{1}_{[0,s\wedge t]})\big) = (t \wedge s) \wedge .$$

so that

$$\pi_t(s \wedge .) = \begin{cases} s \wedge . & \text{if } s \leq t \\ t \wedge . & \text{otherwise.}\end{cases} \tag{2.5}$$

Lemma 2.5. *Let $F \in \mathbb{D}_{p,1}$ and $\mathcal{F}_t = \sigma\{\omega(s),\ s \leq t\}$. Then, $\mathbf{E}[F \,|\, \mathcal{F}_t]$ belongs to $\mathbb{D}_{p,1}$ and we have*

$$\pi_t \mathbf{E}[\nabla F \,|\, \mathcal{F}_t] = \nabla \mathbf{E}[F \,|\, \mathcal{F}_t] \tag{2.6}$$

Furthermore, if F is $\mathcal{F}_t$-measurable then $\dot\nabla_s F = 0$ for all $s > t$.

Proof. Step 1. First consider that F is cylindrical. For the sake of simplicity, imagine that

$$F = f\big(B(t_1), B(t_2)\big) \text{ with } t_1 < t < t_2.$$

Then,

$$\begin{aligned}\mathbf{E}\left[F \mid \mathcal{F}_t\right] &= \mathbf{E}\left[f\big(B(t_1), B(t_2) - B(t) + B(t)\big)\right] \\ &= \int_{\mathbf{R}} f\big(B(t_1), B(t) + x\big) p_{t-t_2}(x)\ \mathrm{d}x \\ &= \tilde{f}\big(B(t_1), B(t)\big),\end{aligned} \tag{2.7}$$

where p_{t-t_2} is the density of a centered Gaussian distribution of variance $(t_2 - t)$ and

$$\tilde{f}(u, v) = \int_{\mathbf{R}} f\big(u, v + x\big) p_{t-t_2}(x)\ \mathrm{d}x \text{ belongs to Schwartz}(\mathbf{R}^2).$$

On the one hand,

$$\nabla_s \mathbf{E}\left[F \mid \mathcal{F}_t\right] = \partial_1 \tilde{f}\big(B(t_1), B(t)\big)\, t_1 \wedge s + \partial_2 \tilde{f}\big(B(t_1), B(t)\big)\, t \wedge s. \tag{2.8}$$

On the other hand,

$$\begin{aligned}\mathbf{E}\left[\nabla_s F \mid \mathcal{F}_t\right] = \mathbf{E}&\left[\partial_1 f\big(B(t_1), B(t_2)\big) \mid \mathcal{F}_t\right]\ t_1 \wedge s \\ &+ \mathbf{E}\left[\partial_2 f\big(B(t_1), B(t_2)\big) \mid \mathcal{F}_t\right]\ t_2 \wedge s.\end{aligned} \tag{2.9}$$

The same reasoning as in (2.7) leads to

$$\begin{aligned}\mathbf{E}\left[\partial_i f(B(t_1), B(t_2)) \mid \mathcal{F}_t\right] &= \int_{\mathbf{R}} \partial_i f\big(B(t_1), B(t) + x\big) p_{t-t_2}(x)\ \mathrm{d}x \\ &= \partial_i \tilde{f}\big(B(t_1), B(t)\big),\end{aligned} \tag{2.10}$$

for $i \in \{1, 2\}$. In view of (2.10), Eqn. (2.9) becomes

$$\mathbf{E}\left[\nabla_s F \mid \mathcal{F}_t\right] = \sum_{i=1}^{2} \partial_i \tilde{f}\big(B(t_1), B(t)\big)\ t_i \wedge s. \tag{2.11}$$

Thus, according to (2.5),

$$\begin{aligned}\pi_t \mathbf{E}\left[\nabla_s F \mid \mathcal{F}_t\right] &= \sum_{i=1}^{2} \partial_i \tilde{f}\big(B(t_1), B(t)\big)\ \pi_t(t_i \wedge .)(s) \\ &= \sum_{i=1}^{2} \partial_i \tilde{f}\big(B(t_1), B(t)\big)\ (t_i \wedge t) \wedge s \\ &= \nabla_s \mathbf{E}\left[F \mid \mathcal{F}_t\right].\end{aligned}$$

Step 2. For the general case, let $(F_n, n \geq 0)$ a sequence of elements of $\mathcal{S}$ converging to F in $\mathbb{D}_{p,1}$. We can construct a sequence of cylindrical functions which are $\mathcal{F}_t$ measurable and converge in $\mathbb{D}_{p,1}$ to $\mathbf{E}\left[F \mid \mathcal{F}_t\right]$. For any n, there exist $t_1^n < \ldots < t_{k_n}^n$ such that $F_n = f_n(B(t_1^n), \cdots, B(t_{k_n}^n))$. If $t_{j_0}^n \leq t < t_{j_0+1}^n$, for $l \geq j_0+1$, replace $B(t_l^n)$ by

$$\big(B(t_l^n) - B(t_{l-1}^n)\big) + \ldots + \big(B(t_{j_0+1}^n) - B(t)\big) + B(t).$$

Let W^n the Gaussian vector whose coordinates are the independent Gaussian random variables $(B(t_{k_n}^n) - B(t_{k_n-1}^n), \cdots, B(t_{j_0+1}^n) - B(t))$ and

$$\begin{aligned} \kappa_n \, : \, \mathbf{R}^{k_n} &\longrightarrow \mathbf{R}^{k_n} \\ w = (w_i,\, 1 \leq i \leq k_n) &\longmapsto w_i \text{ if } i \leq j_0, \\ &\longmapsto w_i + B(t) + \sum_{l=1}^{i-j_0} W_l^n \text{ if } i > j_0. \end{aligned}$$

Hence

$$\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right] = \mathbf{E}\left[(f_n \circ \kappa_n)\big(B(t_1^n), \cdots, B(t_{j_0})\big) \,|\, B(t_1^n), \cdots, B(t)\right].$$

Starting from this identity, we can reproduce the latter reasoning and see that (2.6) holds for such functionals.

Step 3. It remains to prove that $\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right]$ converges to $F = \mathbf{E}\left[F \,|\, \mathcal{F}_t\right]$ in $\mathbb{D}_{p,1}$. By Jensen inequality,

$$\mathbf{E}\left[|\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right] - \mathbf{E}\left[F \,|\, \mathcal{F}_t\right]|^p\right] \leq \mathbf{E}\left[|F_n - F|^p\right] \xrightarrow{n\to\infty} 0.$$

According to Proposition 2.1, the dual of $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ is $L^q(\mathcal{W} \to \mathcal{H};\, \mu)$ and

$$\begin{aligned} &\|\nabla\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right] - \nabla\mathbf{E}\left[F_m \,|\, \mathcal{F}_t\right]\|_{L^p(\mathcal{W}\to\mathcal{H};\,\mu)} \\ &= \sup\left\{\Big|\mathbf{E}\left[\langle\nabla\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right] - \nabla\mathbf{E}\left[F_m \,|\, \mathcal{F}_t\right], h\otimes G\rangle\right]\Big|,\ \|h\|_{\mathcal{H}} = 1,\ \|G\|_{L^q} = 1\right\}. \end{aligned}$$

Then, (2.6) implies that

$$\begin{aligned} &|\mathbf{E}\left[\langle\nabla\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right] - \nabla\mathbf{E}\left[F_m \,|\, \mathcal{F}_t\right], h\otimes G\rangle\right]| \\ &\qquad = |\mathbf{E}\left[\langle\pi_t\nabla\mathbf{E}\left[F_n - F_m \,|\, \mathcal{F}_t\right], h\rangle_{\mathcal{H}}\, G\right]| \\ &\qquad = |\mathbf{E}\left[\langle\nabla(F_n - F_m),\, \pi_t h\rangle_{\mathcal{H}}\, \mathbf{E}\left[G \,|\, \mathcal{F}_t\right]\right]| \\ &\qquad \leq \|\nabla(F_n - F_m)\|_{L^p(W;\mathcal{H})}\|h\|_{\mathcal{H}}\|G\|_{L^q(\mathcal{W}\to\mathbf{R};\,\mu)}. \end{aligned}$$

Since $(\nabla F_n,\, n \geq 0)$ is a Cauchy sequence in $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$, so does the sequence $(\nabla\mathbf{E}\left[F_n \,|\, \mathcal{F}_t\right],\, n \geq 0)$, hence it is a converging sequence. Since ∇ is closable, the limit can only be $\nabla\mathbf{E}\left[F \,|\, \mathcal{F}_t\right]$.

Step 4. Let $H_t^\perp = \bigcap_{s\in[t,1]\cap\mathbf{Q}} \ker(\epsilon_s - \epsilon_t)$; it is a denumerable intersection of closed subspaces of $\mathcal{H}$, hence it is closed in $\mathcal{H}$. By sample-paths continuity of the elements of $\mathcal{H}$, $\dot{h}(s) = 0$ for $s > t$ means that $h(s) = h(t)$ for any $s > t$ and $s \in \mathbf{Q}$, which is equivalent to $h \in H_t^\perp$. There exists a subsequence, we still denote by $(F_n, n \geq 0)$, such that $\nabla \mathbf{E}[F_n \mid \mathcal{F}_t]$ converges almost-surely in $\mathcal{H}$ to $\nabla \mathbf{E}[F \mid \mathcal{F}_t]$. Since $H_t^{\perp}$ is closed, $\nabla \mathbf{E}[F \mid \mathcal{F}_t]$ belongs to $H_t^{\perp}$.

As we saw above, an element U of $L^p(\mathcal{W} \to \mathcal{H};\, \mu)$ can be represented as

$$U(\omega, t) = \int_0^t \dot{U}(\omega, s)\ \mathrm{d}s, \text{ for all } t \in [0,1] \tag{2.12}$$

where $\dot{U}$ is measurable from $W \times [0,1]$ onto $\mathbf{R}$.

Definition 2.7. An $\mathcal{H}$-valued random variable is said to be adapted whenever there exists a process $\dot{U}$ adapted in the classical sense such that (2.12) holds.

We denote by $L^2_a(W; \mathcal{H})$ the set of $\mathcal{H}$-valued adapted, random variables such that

$$\mathbf{E}\left[\int_0^1 |\dot{U}(s)|^2\ \mathrm{d}s\right] = \mathbf{E}\left[\|U\|^2_{\mathcal{H}}\right] < \infty.$$

It is a closed subspace of $L^2_a(W; \mathcal{H})$.

Similarly $\mathbb{D}^a_{2,1}(\mathcal{H})$ is the subset of $L^2(W; \mathcal{H})$ such that

$$\mathbf{E}\left[\iint |\dot{\nabla}_r \dot{U}(s)|^2\ \mathrm{d}r\ \mathrm{d}s\right] = \mathbf{E}\left[\|\nabla U\|^2_{L^2([0,1];\mathcal{H})}\right] < \infty.$$

Proposition 2.5. *Let X and Y two Banach spaces and M a dense subset of X. Consider $(A_n,\, n \geq 0)$ a sequence of continuous linear maps from X to Y such that:*

1. $\sup_{n \geq 0} \|A_n\| < \infty$,
2. *for any $x \in M$, the sequence $(A_n x,\, n \geq 0)$ is Cauchy in Y.*

Then, for any $x \in X$ and not only in M, the sequence $(A_n x,\, n \geq 0)$ is convergent in F and the linear map defined by $Ax = \lim_{n \to \infty} A_n x$ is continuous from X into Y. Moreover, $\|A\| \leq \sup_n \|A_n\|$.

Theorem 2.6. *Let U belongs to $\mathbb{D}^a_{p,1}(\mathcal{H})$, and $\mathcal{D}_n$ the dyadic partition of $[0,1]$ of step 2^{-n}. Then,*

$$\dot{U}_{\mathcal{D}_n}(t) = \sum_{i=1}^{2^n-1} 2^n \Big(\int_{(i-1)2^{-n}}^{i\,2^{-n}} \dot{U}(r)\ dr\Big)\ \mathbf{1}_{(i2^{-n},(i+1)2-n]}(t)$$

converges in $\mathbb{D}^a_{p,1}(\mathcal{H})$ to U.

Proof. Since indicator functions with disjoint support are orthogonal in $L^2([0,1])$, we have

$$\int_0^1 |\dot U_{\mathcal{D}_n}(t)|^2 \,\mathrm{d}t = \sum_{i=1}^{2^n-1} \Big(2^n \int_{(i-1)2^{-n}}^{i\,2^{-n}} \dot U(r)\,\mathrm{d}r\Big)^2 \int_0^1 \mathbf{1}_{(i2^{-n},(i+1)2{-n}]}(t)\,\mathrm{d}t$$
$$\le \sum_{i=1}^{2^n-1} \int_{(i-1)2^{-n}}^{i\,2^{-n}} |\dot U(r)|^2 \,\frac{\mathrm{d}r}{2^{-n}}\, 2^{-n} = \int_0^1 |\dot U(r)|^2 \,\mathrm{d}r,$$

according to the Jensen inequality. Hence,

$$\mathbf{E}\left[\Big(\int_0^1 |\dot U_{\mathcal{D}_n}(t)|^2 \,\mathrm{d}t\Big)^{p/2}\right] \le \mathbf{E}\left[\Big(\int_0^1 |\dot U(r)|^2 \,\mathrm{d}r\Big)^{p/2}\right],$$

or in other words that the maps

$$\begin{aligned} \mathcal{D}_n \,:\, L^p(\mathcal{W}\to\mathcal{H};\,\mu) &\longrightarrow L^p(\mathcal{W}\to\mathcal{H};\,\mu) \\ U &\longmapsto I^1(\dot U_{\mathcal{D}_n}) \end{aligned}$$

are continuous and satisfy $\|\mathcal{D}_n\| \le 1$. Let M be the subset of $L^p(\mathcal{W}\to\mathcal{H};\,\mu)$ composed of processes such that $\dot U$ has continuous sample-paths with $\mathbf{E}\left[\|\dot U\|_\infty^p\right] < \infty$. For such a process

$$\|\dot U - \dot U_{\mathcal{D}_n}\|^2_{L^2([0,1])} \le \sum_{i=1}^{2^n-1} \int_{(i-1)2^{-n}}^{i\,2^{-n}} \Big(2^n \int_{(i-1)2^{-n}}^{i\,2^{-n}} |\dot U(r) - \dot U(t)| \,\mathrm{d}r\Big)^2 \,\mathrm{d}t$$
$$\le \sum_{i=1}^{2^n-1} \int_{(i-1)2^{-n}}^{i\,2^{-n}} 2^n \int_{(i-1)2^{-n}}^{i\,2^{-n}} |\dot U(r) - \dot U(t)|^2 \,\mathrm{d}r\,\mathrm{d}t,$$

by the Jensen inequality. Since $\dot U$ is a.s. continuous, for $t \in ((i-1)2^{-n},\ i\,2^{-n}]$,

$$2^n \int_{(i-1)2^{-n}}^{i\,2^{-n}} |\dot U(r) - \dot U(t)|^2 \,\mathrm{d}r \xrightarrow[\text{a.s.}]{n\to\infty} 0.$$

Since $\mathbf{E}\left[\|\dot U\|_\infty^p\right]$ is finite, the dominated convergence theorem entails that

$$\mathbf{E}\left[\Big(\sum_{i=1}^{2^n-1} \int_{(i-1)2^{-n}}^{i\,2^{-n}} 2^n \int_{(i-1)2^{-n}}^{i\,2^{-n}} |\dot U(r) - \dot U(t)|^2 \,\mathrm{d}r\,\mathrm{d}t\Big)^{p/2}\right] \xrightarrow{n\to\infty} 0.$$

Apply the Proposition 2.5 to see that $\dot U_{\mathcal{D}_n}$ tends to $\dot U$ in $L^2(\mathcal{W}\otimes[0,\,1] \to \mathbf{R};\,\mu\otimes\lambda)$.

Remark that if $\dot U$ is adapted then so does $\dot U_{\mathcal{D}_n}$. Moreover, there exists a subsequence $\dot U_{\pi_{n_k}}$ which converges almost-surely to $\dot U$ and that guarantees the adaptability of $\dot U$.

If $U \in \mathbb{D}_{2,1}$, $\dot{\nabla}_r \dot{U}_t$ can be approximated in $L^2(\mathcal{W} \otimes [0,\, 1]^2 \to \mathbf{R};\, \mu \otimes \lambda^{\otimes 2})$ by

$$\sum_{i=1}^{2^n-1} 2^n \Big(\int_{(i-1)2^{-n}}^{i\,2^{-n}} \dot{\nabla}_r \dot{U}(s) \, \mathrm{d}s \Big) \, \mathbf{1}_{(i2^{-n},(i+1)2-n]}(t).$$

Then, the same proof as before shows this approximation converges in $L^2(\mathcal{W} \otimes [0,\, 1]^2 \to \mathbf{R};\, \mu \otimes \lambda^{\otimes 2})$ to $\dot{\nabla} \dot{U}$.

Theorem 2.7. *For $U \in \mathbb{D}^a_{2,1}(\mathcal{H})$, the Itô integral of $\dot{U}$ belongs to $\mathbb{D}_{2,1}$ and for any $h \in \mathcal{H}$,*

$$\Big\langle \nabla \Big(\int \dot{U}(s) \ dB(s) \Big),\, h \Big\rangle_{\mathcal{H}} = \int_0^1 \dot{U}(s) \dot{h}(s) \ ds + \int_0^1 \Big\langle \nabla \dot{U}(s),\, h \Big\rangle_{\mathcal{H}} \ dB(s) \quad (2.13)$$

converges in $\mathbb{D}^a_{2,1}(\mathcal{H})$ to U.

Proof. From the previous theorem, we know that $\Big\langle \nabla \dot{U}(s),\, h \Big\rangle_{\mathcal{H}}$ is adapted and square integrable so that its stochastic integral is well defined. For $U(t) = U_a \, I^1(\mathbf{1}_{(a,b]})(t)$ with $U_a \in \mathcal{F}_a$ and $U_a \in \mathbb{D}_{2,1}$, on the one hand, since ∇ is a derivation operator, we have

$$\begin{aligned} \Big\langle \nabla \Big(\int \dot{U}(s) \ \mathrm{d}B(s) \Big),\, h \Big\rangle_{\mathcal{H}} &= \Big\langle \nabla \Big(U_a \left(B(b) - B(a) \right) \Big),\, h \Big\rangle_{\mathcal{H}} \\ &= \langle \nabla U_a,\, h \rangle_{\mathcal{H}} \ \big(B(b) - B(a) \big) + \int_0^1 U_a \, \mathbf{1}_{(a,b]}(s) \, \dot{h}(s) \ \mathrm{d}s \\ &= \int_0^1 \langle \nabla U_a,\, h \rangle_{\mathcal{H}} \, \mathbf{1}_{(a,b]}(s) \ \mathrm{d}B(s) + \int_0^1 U_a \, \mathbf{1}_{(a,b]}(s) \, \dot{h}(s) \ \mathrm{d}s \\ &= \int_0^1 \dot{U}(s) \dot{h}(s) \ \mathrm{d}s + \int_0^1 \Big\langle \nabla \dot{U}(s),\, h \Big\rangle_{\mathcal{H}} \ \mathrm{d}B(s). \end{aligned}$$

By linearity, Eqn. (2.13) holds for simple processes as in Theorem 2.6. Since for U with continuous sample-paths, U_π tends in $L^2(W \times [0,1], \mu \otimes \lambda)$ to U, in virtue of Lemma 2.4, it remains to prove that

$$\sup_\pi \mathbf{E} \left[\| \nabla \int \dot{U}_\pi(s) \ \mathrm{d}B(s) \|^p_{\mathcal{H}} \right] < \infty.$$

Consider the map

$$\begin{aligned}\int_0^1 \nabla \dot{U}(s)\, \mathrm{d}B(s) \,:\, \mathcal{H} &\longrightarrow L^2_a(\mu)\\ h &\longmapsto \left\langle \int_0^1 \nabla \dot{U}(s),\, h \right\rangle_{\mathcal{H}} \mathrm{d}B(s).\end{aligned}$$

We expect to compute

$$\mathbf{E}\left[\sup_{\|h\|_{\mathcal{H}}=1} \left| \int_0^1 \left\langle \nabla \dot{U}(s),\, h \right\rangle_{\mathcal{H}} \mathrm{d}B(s) \right|^p \right].$$

Without the supremum inside the expectation, we would usually refer to the Burkholder-Davis-Gundy inequality. In order to deal with the supremum, we note that

$$t \longmapsto \int_0^t \nabla \dot{U}(s)\, \mathrm{d}B(s)$$

is an Hilbert valued martingale and it satisfies also a BDG inequality:

$$\begin{aligned}\mathbf{E}\left[\sup_{t\le 1} \left| \int_0^t \nabla \dot{U}(s)\, \mathrm{d}B(s) \right|^p \right] &\le c_p\, \mathbf{E}\left[\left(\int_0^1 \|\nabla \dot{U}(s)\|^2_{\mathcal{H}}\, \mathrm{d}s \right)^{p/2} \right]\\ &= c_p \mathbf{E}\left[\left(\int_0^1 \int_0^1 |\dot{\nabla}_r \dot{U}(s)|^2\, \mathrm{d}r\, \mathrm{d}s \right)^{p/2} \right] = c_p \|\nabla U\|^p_{L^p(W;\mathcal{H}\otimes\mathcal{H})}.\end{aligned}$$

Combining (2.13) with this upper-bound, we get

$$\mathbf{E}\left[\|\nabla \int \dot{U}_\pi(s)\, \mathrm{d}B(s)\|^p_{\mathcal{H}} \right] \le c \left(\|U_\pi\|^p_{L^p(\mathcal{W}\to\mathcal{H};\,\mu)} + \|\nabla U_\pi\|^p_{L^p(W;\mathcal{H}\otimes\mathcal{H})} \right)$$

We conclude with Theorem 2.6.

For cylindrical functions, we can clearly define higher order derivative following the same rule. The only difficulty is to realize that the second (respectively k-th) order gradient belongs to $\mathcal{H}^{\otimes(2)}$ (respectively $\mathcal{H}^{\otimes(k)}$): For instance, for $F = f(\delta h_j,\, 1 \le j \le n)$,

$$\begin{aligned}\left\langle \nabla^{(2)} F, h \otimes k \right\rangle &= \sum_{j,l=1}^n \partial_{j,l} f(\delta h_j,\, 1 \le j \le n)\ \langle h_j,\, h\rangle_{\mathcal{H}} \otimes \langle h_l,\, k\rangle_{\mathcal{H}}\\ &= \left\langle \nabla\big(\langle \nabla F, h\rangle_{\mathcal{H}}\big),\, k\right\rangle_{\mathcal{H}}.\end{aligned}$$

Definition 2.8. For any $p > 1$ and $k \ge 1$, $\mathbb{D}_{p,k}$ is the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{p,k} = \|F\|_p + \sum_{j=1}^{k} \|\nabla^{(j)} F\|_{L^p(W;\mathcal{H}^{\otimes(j)})}.$$

The space of *test functions* is $\mathbb{D} = \cap_{p>1} \cap_{k\geq 1} \mathbb{D}_{p,k}$.

2.2 Divergence

For a matrix $M \in \mathcal{M}_{n,p}(\mathbf{R})$, its adjoint, which turns to coincide with its transpose, is defined by the identity:

$$\langle Mx,\, y\rangle_{\mathbf{R}^p} = \langle x, M^* y\rangle_{\mathbf{R}^n}.$$

We see that to define an adjoint, we need to have a notion a scalar product or more generally of a duality bracket. This means that if M is continuous from a Banach E into a Banach F, its adjoint is a continuous map from F^* into E^* defined by the identity:

$$\langle Mx,\, y\rangle_{F,F^*} = \langle x, M^* y\rangle_{E,E^*}.$$

For any $q > 1$, the Gross-Sobolev derivative, which we denoted by ∇, is continuous between the two spaces:

$$\mathbb{D}_{q,1} \subset L^q(\mathcal{W} \to \mathbf{R};\, \mu) \longrightarrow L^q(\mathcal{W} \to \mathcal{H};\, \mu).$$

Therefore its adjoint is a map from

$$\Big(L^q(W \to \mathcal{H};\, \mu)\Big)^* = L^p(\mathcal{W} \to \mathcal{H};\, \mu) \\ \longrightarrow \Big(L^q(\mathcal{W} \to \mathbf{R};\, \mu)\Big)^* = L^p(\mathcal{W} \to \mathbf{R};\, \mu)$$

with $1/p + 1/q = 1$ and must satisfy the identity

$$\langle \nabla F, U\rangle_{L^q(\mathcal{W}\to\mathcal{H};\,\mu),L^p(\mathcal{W}\to\mathcal{H};\,\mu)} = \langle F, \nabla^* U\rangle_{L^q(\mathcal{W}\to\mathbf{R};\,\mu),L^p(\mathcal{W}\to\mathbf{R};\,\mu)} \\ \iff \mathbf{E}\left[\langle \nabla F, U\rangle_{\mathcal{H}}\right] = \mathbf{E}\left[F\,\delta U\right].$$

An additional difficulty comes from the fact that ∇ is not defined on the whole of $L^q(\mathcal{W} \to \mathbf{R};\, \mu)$ but only on the subset $\mathbb{D}_{q,1}$, hence we need to take some restrictions in the definition of the adjoint.

Definition 2.9. Let $p > 1$. Let $\operatorname{Dom}_p \nabla^*$ be the set of $\mathcal{H}$-valued random variables U for which there exists $c_p(U)$ such that for any $F \in \mathbb{D}_{q,1}$,

$$|\mathbf{E}\left[\langle \nabla F,\, U\rangle_{\mathcal{H}}\right]| \leq c_p(U)\, \|F\|_{L^q(\mathcal{W}\to\mathbf{R};\,\mu)}.$$

In this case, we define $\nabla^* U$ as the unique element of $L^p(\mathcal{W} \to \mathbf{R};\, \mu)$ such that

$$\mathbf{E}\left[\langle \nabla F,\, U\rangle_{\mathcal{H}}\right] = \mathbf{E}\left[F\, \nabla^* U\right].$$

Remark 2.4 (∇^* coincides with the Wiener integral on $\mathcal{H}$). Recall that δ is the Wiener integral. We now show that $\delta = \nabla^*|_{\mathcal{H}}$. For any $F \in \mathcal{S}$, according to (2.3), we have

$$\mathbf{E}\left[\langle \nabla F,\, h\rangle_{\mathcal{H}}\right] = \mathbf{E}\left[F \delta h\right] \tag{2.14}$$

and δh is a Gaussian random variable of variance $\|h\|^2_{\mathcal{H}}$, thus belongs to any $L^q(\mathcal{H} \to \mathbf{R};\, \mu)$ for any $q > 1$. Hence,

$$\left|\mathbf{E}\left[\langle \nabla F,\, h\rangle_{\mathcal{H}}\right]\right| \le c\, \|h\|_{\mathcal{H}} \|F\|_{L^p(\mathcal{W}\to\mathbf{R};\, \mu)}.$$

This means that h belongs to $\operatorname{Dom}_p \delta$ and (2.14) entails that $\nabla^* h = \delta h$. Henceforth, in the following, we will use the notation δ instead of ∇^* and we keep for further reference the fundamental formula

$$\mathbf{E}\left[\langle \nabla F,\, U\rangle_{\mathcal{H}}\right] = \mathbf{E}\left[F\, \delta U\right] \tag{2.15}$$

for any $F \in \mathbb{D}_{q,1}$ and $U \in \operatorname{Dom}_p \delta$.

In usual deterministic calculus, if a is a constant, then trivially

$$\int au(s)\, \mathrm{d}s = a \int u(s)\, \mathrm{d}s. \tag{2.16}$$

For Itô integrals, this property does not hold any longer since we may have a problem of adaptability: If a is a random variable, not belonging to $\mathcal{F}_0$ and u is an adapted process with all the required integrability properties, then the process $(a\, u(s),\, s \ge 0)$ is not adapted so that $\int au(s)\, \mathrm{d}B(s)$ is not even well defined. For the divergence, since we got rid of the adaptability hypothesis, we can prove a formula analog to (2.16) which is a simple consequence of the fact that ∇ is a derivation operator.

Theorem 2.8. *Let $U \in \operatorname{Dom}_p \delta$ and $a \in \mathbb{D}_{q,1}$ with $1/p + 1/q = 1/r$. Then, $aU \in \operatorname{Dom}_r \delta$ and*

$$\delta(aU) = a\, \delta U - \langle \nabla a, U\rangle_{\mathcal{H}}. \tag{2.17}$$

Proof. Step 1. We first prove that the right-hand-side belongs to $L^r(\mathcal{W} \to \mathbf{R};\, \mu)$.

$$\mathbf{E}\left[|a\delta U|^r\right] \le \mathbf{E}\left[|a|^p\right]^{r/p} \mathbf{E}\left[|\delta U|^q\right]^{r/q} \tag{2.18}$$

and

$$\begin{aligned}\mathbf{E}\left[|\langle\nabla a,\, U\rangle_{\mathcal{H}}|^r\right] &\le \mathbf{E}\left[\|\nabla a\|_{\mathcal{H}}^r\|U\|_{\mathcal{H}}^r\right]\\ &\le \mathbf{E}\left[\|\nabla a\|_{\mathcal{H}}^q\right]^{r/q}\mathbf{E}\left[\|U\|_{\mathcal{H}}^p\right]^{r/p}\\ &\le \|a\|_{\mathbb{D}_{p,1}}^r\|U\|_{\mathbb{D}_{q,1}}^r. \end{aligned}\tag{2.19}$$

STEP 2. Denote $r^* = r/(r-1)$. For $F \in \mathbb{D}_{r^*,1}$, since ∇ is a true derivation,

$$\begin{aligned}\mathbf{E}\left[\langle\nabla F, aU\rangle_{\mathcal{H}}\right] &= \mathbf{E}\left[\langle a\nabla F, U\rangle_{\mathcal{H}}\right]\\ &= \mathbf{E}\left[\langle\nabla(aF) - F\nabla a, U\rangle_{\mathcal{H}}\right]\\ &= \mathbf{E}\left[F\, a\delta U\right] - \mathbf{E}\left[F\langle\nabla a, U\rangle_{\mathcal{H}}\right]. \end{aligned}\tag{2.20}$$

According to (2.18) and (2.19), (2.20) implies that

$$|\mathbf{E}\left[\langle\nabla F, aU\rangle_{\mathcal{H}}\right]| \le \|a\|_{\mathbb{D}_{p,1}}\|U\|_{\mathbb{D}_{q,1}}\,\|F\|_{L^{r^*}(\mathcal{W}\to\mathcal{H};\,\mu)}$$

Hence, aU belongs to $\operatorname{Dom}_r\delta$.
STEP 3. At last, (2.20) implies (2.17) by identification.

We have already seen that the Itô integral coincides with the Wiener integral for deterministic integrands provided that we identify h and $\dot{h}$. We now show that modulo the same identification, the divergence of adapted processes coincides with their Itô integral.

Corollary 2.2 (Divergence extends Itô integral). *Let $U \in \mathbb{D}_{2,1}^a(\mathcal{H})$. Then, U belong to $\operatorname{Dom}_2\delta$ and*

$$\delta U = \int_0^1 \dot{U}(s)\ dB(s), \tag{2.21}$$

where the stochastic integral is taken in the Itô sense.

Proof. The principle of the proof is to establish (2.21) for adapted simple processes and then pass to the limit.
STEP 1. For $0 \le s < t \le 1$, let

$$\dot{U}(r) = \theta_s\,\mathbf{1}(s,t](r),\ \text{i.e. } U(r) = \theta_s\left(t\wedge r - s\wedge r\right),$$

where $\theta_s \in \mathbb{D}_{2,1}$ and θ_r is $\mathcal{F}_s$-measurable. According to Theorem 2.8, U is in $\operatorname{Dom}_2\delta$ and

$$\begin{aligned}\delta(U) &= \theta_s\,\delta(t\wedge . - s\wedge .) - \langle\nabla\theta_s,\, t\wedge . - s\wedge .\rangle_{\mathcal{H}}\\ &= \theta_s\left(B(t)-B(s)\right) - \int_0^1 \dot{\nabla}_\tau\theta_s\,\mathbf{1}(s,t](\tau)\tau\end{aligned}$$

Now recall that according to Lemma 2.5, since $\theta_s \in \mathcal{F}_s$,

$$\dot{\nabla}_\tau\theta_s = 0 \text{ if } \tau > r,$$

hence

$$\delta(U) = \theta_s \big(B(t) - B(s)\big) = \int_0^1 \dot{U}(r) \, \mathrm{d}B(r). \tag{2.22}$$

Step 2. If $\dot{U}$ is adapted, the random variable

$$2^n \left(\int_{(i-1)2^{-n}}^{i2^{-n}} \dot{U}(r) \, \mathrm{d}r \right) \text{ belongs to } \mathcal{F}_{i2^{-n}}.$$

Hence, with the notations of Theorem 2.6, we have by linearity

$$\delta(U_{\mathcal{D}_n}) = \int_0^1 \dot{U}_{\mathcal{D}_n}(r) \, \mathrm{d}B(r).$$

Step 3. It remains to show that we can pass to the limit in both sides of (2.21). The application δ is continuous from $\mathbb{D}_{2,1}^a(\mathcal{H}) \subset \mathbb{D}_{2,1}(\mathcal{H})$ into $L^2(\mathcal{W} \to \mathbf{R};\, \mu)$. Hence, Theorem 2.6 entails that

$$\delta(U_{\mathcal{D}_n}) \xrightarrow[n\to\infty]{L^2(\mathcal{W}\to\mathbf{R};\, \mu)} \delta(U).$$

Furthermore, the Itô integral is an isometry hence a continuous map from $L_a^2(\mathcal{W} \times [0,1] \to \mathbf{R}; \mu)$ into $L^2(\mathcal{W} \to \mathbf{R};\, \mu)$. Hence,

$$\int_0^1 \dot{U}_{\mathcal{D}_n}(r) \, \mathrm{d}B(r) \xrightarrow[n\to\infty]{L^2(\mathcal{W}\to\mathbf{R};\, \mu)} \int_0^1 \dot{U}(r) \, \mathrm{d}B(r).$$

The proof is thus complete.

The Itô isometry states that for U adapted

$$\mathbf{E}\left[\left(\int_0^1 \dot{U}(s) \, \mathrm{d}B(s) \right)^2 \right] = \mathbf{E}\left[\int_0^1 |\dot{U}(s)|^2 \, \mathrm{d}s \right].$$

One of the most elegant formula given by the Malliavin calculus is the generalization of this identity to non-adapted integrands.

Remark 2.5. If $U \in \mathbb{D}_{2,1}(\mathcal{H})$ then $\dot{\nabla}\dot{U}$ is a.s. an Hilbert-Schmidt map on $L^2([0,1] \times [0,1], \lambda \otimes \lambda)$. Indeed, by the definition of the norm in $\mathbb{D}_{2,1}(\mathcal{H})$,

$$\|U\|_{\mathbb{D}_{2,1}}^2 = \mathbf{E}\left[\|\nabla U\|_{\mathcal{H}\otimes\mathcal{H}}^2\right] = \mathbf{E}\left[\int_0^1 \int_0^1 |\dot{\nabla}_r \dot{U}(s)|^2 \, \mathrm{d}r \, \mathrm{d}s \right].$$

This ensures the almost-sure finiteness of

$$\int_0^1 \int_0^1 |\dot{\nabla}_r \dot{U}(s)|^2 \, \mathrm{d}r \, \mathrm{d}s,$$

meaning that $\dot{\nabla}\dot{U}$ is Hilbert-Schmidt with probability 1.

Lemma 2.6. *If U belongs to $\mathbb{D}^a_{2,1}(\mathcal{H})$ then* $\operatorname{trace}(\nabla U \circ \nabla U) = 0$.

Proof. According to Lemma 1.4,

$$\operatorname{trace}(\nabla U \circ \nabla U) = \iint_{[0,1]^2} \dot{\nabla}_r \dot{U}(s) \dot{\nabla}_s \dot{U}(r) \, \mathrm{d}r \, \mathrm{d}s.$$

Since $\dot{U}(s)$ is $\mathcal{F}_s$-measurable, $\dot{\nabla}_r \dot{U}(s) = 0$ if $r > s$. Similarly, $\dot{\nabla}_s \dot{U}(r) = 0$ if $s > r$. Hence, the product is zero $\lambda \otimes \lambda$ almost-surely. It follows that the integral is null.

Theorem 2.9 (L^2 norm of divergence). *The space $\mathbb{D}_{1,2}(\mathcal{H})$ is included in* $\operatorname{Dom}_2 \delta$ *and for $U \in \mathbb{D}_{1,2}(\mathcal{H})$,*

$$\mathbf{E}\left[\delta U^2\right] = \mathbf{E}\left[\|U\|^2_{\mathcal{H}}\right] + \mathbf{E}\left[\operatorname{trace}(\nabla U \circ \nabla U)\right]. \tag{2.23}$$

Lemma 2.7. *For $k \geq 1$, for $V \in \mathbb{D}_{2,1}(\mathcal{H}^{\otimes(k)})$, for $x \in \mathcal{H}^{\otimes(k)}$, for $h \in \mathcal{H}$,*

$$\langle \nabla \langle V, x \rangle_{\mathcal{H}^{\otimes(k)}}, \, h \rangle_{\mathcal{H}} = \langle \nabla V, \, x \otimes h \rangle_{\mathcal{H}^{\otimes(k+1)}}$$

Proof. For the sake of simplicity, we give the proof for $k = 1$. The general case is handled similarly. Going back to the definition of the scalar product in $\mathcal{H}$, we have

$$\langle \nabla \langle V, x \rangle_{\mathcal{H}^{\otimes(k)}}, \, h \rangle_{\mathcal{H}} = \int_0^1 \dot{\nabla}_s \left(\int_0^1 \dot{V}(r) \, \dot{x}(r) \, \mathrm{d}r \right) \dot{h}(s) \, \mathrm{d}s.$$

Approximate the inner integral by Riemann sums and pass to the limit to show that

$$\dot{\nabla}_s \left(\int_0^1 \dot{V}(r) \, \dot{x}(r) \, \mathrm{d}r \right) = \int_0^1 \dot{\nabla}_s \dot{V}(r) \, \dot{x}(r) \, \mathrm{d}r,$$

first for V such that $(r, s) \longmapsto \dot{\nabla}_s \dot{V}(r)$ is continuous and then by density for all $V \in \mathbb{D}_{2,1}(\mathcal{H})$. Hence the result.

Lemma 2.8. *For $U \in \mathbb{D}_{2,2}(\mathcal{H})$, for any $h, k, l \in \mathcal{H}$,*

$$\begin{aligned} \left\langle \nabla^{(2)} \langle U, h \rangle_{\mathcal{H}}, \, k \otimes l \right\rangle_{\mathcal{H} \otimes \mathcal{H}} &= \left\langle \nabla^{(2)} U, \, h \otimes k \otimes l \right\rangle_{\mathcal{H} \otimes \mathcal{H}} \\ &= \left\langle \nabla^{(2)} U, \, h \otimes l \otimes k \right\rangle_{\mathcal{H} \otimes \mathcal{H}}. \end{aligned}$$

That means the map $\nabla^{(2)} U$ is a symmetric operator.

Proof. Step 1. The first equality is a consequence of Lemma 2.7.
Step 2. For $F \in \mathcal{S}$, $F = f(\delta h_1, \cdots, \delta h_M)$, in virtue of the Schwarz theorem for crossed derivatives of functions of several variables,

$$\begin{aligned}\left\langle \nabla^{(2)} F,\, k \otimes l \right\rangle_{\mathcal{H}\otimes\mathcal{H}} &= \sum_{i,j=1}^{n} \partial^2_{ij} f(\delta h_1, \cdots, \delta h_M) \, \langle k, h_i \rangle_{\mathcal{H}} \, \langle l, h_j \rangle_{\mathcal{H}} \\ &= \sum_{i,j=1}^{n} \partial^2_{ji} f(\delta h_1, \cdots, \delta h_M) \, \langle l, h_j \rangle_{\mathcal{H}} \, \langle k, h_i \rangle_{\mathcal{H}} \\ &= \left\langle \nabla^{(2)} F,\, l \otimes k \right\rangle_{\mathcal{H}\otimes\mathcal{H}}.\end{aligned}$$

The proof is thus complete.

Proof (Proof of Theorem 2.9). For $U \in \mathbb{D}_{1,2}(\mathcal{H})$, U takes its values in $\mathcal{H}$ so that we can write

$$U = \sum_{n \ge 0} \langle U,\, h_n \rangle_{\mathcal{H}} \ h_n,$$

for $(h_n,\, n \ge 0)$ a complete orthonormal basis of $\mathcal{H}$. The series

$$U_N = \sum_{n=0}^{N} \langle U,\, h_n \rangle_{\mathcal{H}} \ h_n \text{ and } \nabla U_N = \sum_{n=0}^{N} \nabla \langle U,\, h_n \rangle_{\mathcal{H}} \ h_n$$

converge in $L^2(\mathcal{W} \to \mathcal{H};\, \mu)$ and $L^2(\mathcal{W} \to \mathcal{H} \otimes \mathcal{H};\, \mu))$ respectively.
According to (2.17),

$$\delta U_N = \sum_{n=0}^{N} \langle U,\, h_n \rangle_{\mathcal{H}} \ \delta h_n - \sum_{n=0}^{N} \langle \nabla U,\, h_n \otimes h_n \rangle_{\mathcal{H}\otimes\mathcal{H}}.$$

Thus,

$$\begin{aligned}\nabla \delta U_N = \sum_{n=0}^{N} \{ \langle \nabla U,\, h_n \rangle_{\mathcal{H}} \ \delta h_n + \langle U,\, h_n \rangle_{\mathcal{H}} \ h_n \\ - \nabla \Big(\langle \nabla U,\, h_n \otimes h_n \rangle_{\mathcal{H}\otimes\mathcal{H}} \Big) \Big\}.\end{aligned}$$

Consequently, in virtue of Lemma 2.7,

$$\begin{aligned}\mathbf{E}\left[\delta U_N \ \delta U_N \right] &= \sum_{n,k \ge 0}^{N} \mathbf{E}\left[\langle U,\, h_k \rangle_{\mathcal{H}} \, \langle \nabla U,\, h_n \otimes h_k \rangle_{\mathcal{H}\otimes\mathcal{H}} \ \delta h_n \right] \\ &+ \sum_{n,k \ge 0}^{N} \mathbf{E}\left[\langle U,\, h_n \rangle_{\mathcal{H}} \, \langle U,\, h_k \rangle_{\mathcal{H}} \, \langle h_n,\, h_k \rangle_{\mathcal{H}} \right] \\ &- \sum_{n,k \ge 0}^{N} \mathbf{E}\left[\langle U,\, h_k \rangle_{\mathcal{H}} \left\langle \nabla^{(2)} U,\, h_n \otimes h_n \otimes h_k \right\rangle_{\mathcal{H}^{\otimes(3)}} \right] \\ &= A_1 + A_2 - A_3.\end{aligned}$$

On the one hand, Parseval equality yields

$$A_2 = \sum_{n,k\geq 0} \mathbf{E}\left[\langle U,\, h_n\rangle_{\mathcal{H}} \langle U,\, h_k\rangle_{\mathcal{H}} \langle h_n,\, h_k\rangle_{\mathcal{H}}\right] \\ = \sum_{n\geq 0} \mathbf{E}\left[\langle U,\, h_n\rangle_{\mathcal{H}}^2\right] = \mathbf{E}\left[\|U\|_{\mathcal{H}}^2\right].$$

Apply once more the integration by parts formula in A_1:

$$\begin{aligned} A_1 &= \sum_{n,k\geq 0}^{N} \mathbf{E}\left[\langle \nabla U,\, h_k\otimes h_n\rangle_{\mathcal{H}\otimes\mathcal{H}} \langle \nabla U,\, h_n\otimes h_k\rangle_{\mathcal{H}\otimes\mathcal{H}}\right] \\ &+ \sum_{n,k\geq 0}^{N} \mathbf{E}\left[\langle U,\, h_k\rangle_{\mathcal{H}} \left\langle \nabla^{(2)} U,\, h_n\otimes h_k\otimes h_n\right\rangle_{\mathcal{H}^{\otimes(3)}}\right] \\ &= \operatorname{trace}(\nabla U_N \circ \nabla U_N) + A_3, \end{aligned}$$

since $\nabla^{(2)}$ is a symmetric operator, cf. Lemma 2.8.
STEP 3. In brief, we have proved so far that

$$\mathbf{E}\left[\delta U_N^2\right] = \|U_N\|^2_{L^2(W;\mathcal{H})} + \mathbf{E}\left[\operatorname{trace}(\nabla U_N \circ \nabla U_N)\right].$$

Then, Eqn. (1.18) entails that

$$\mathbf{E}\left[\delta(U_N - U_K)^2\right] \leq \|U_N - U_K\|^2_{L^2(\mathcal{W}\to\mathcal{H};\,\mu)} \\ + \|\nabla U_N - \nabla U_K\|^2_{L^2(\mathcal{W}\to\mathcal{H}\otimes\mathcal{H};\,\mu)}.$$

Thus, the sequence $(\delta U_N,\, N\geq 0)$ is Cauchy in $L^2(\mathcal{W}\to\mathbf{R};\,\mu)$ thus convergent towards a limit temporarily denoted by $\eta\in L^2(\mathcal{W}\to\mathbf{R};\,\mu)$. Hence, for $F\in\mathbb{D}_{1,2}$,

$$\mathbf{E}\left[\langle \nabla F,\, U\rangle_{\mathcal{H}}\right] = \lim_{N\to\infty} \mathbf{E}\left[\langle \nabla F,\, U_N\rangle_{\mathcal{H}}\right] = \lim_{N\to\infty} \mathbf{E}\left[F\,\delta U_N\right] = \mathbf{E}\left[F\,\eta\right].$$

By the very definition of the divergence, this means that $U\in\operatorname{Dom}_2\delta$ and $\delta U = \eta = \lim_{N\to\infty}\delta U_N$.

Exercises

Exercise 2.1. For $h\in\mathcal{H}$, show that

$$\nabla^{(k)} \exp(\delta h - \frac{1}{2}\|h\|^2_{\mathcal{H}}) = \exp(\delta h - \frac{1}{2}\|h\|^2_{\mathcal{H}})\; h^{\otimes(k)}. \tag{2.24}$$

Exercise 2.2. Let $F \in \mathbb{D}_{p,1}$ and $\epsilon > 0$. Set $\phi_\epsilon(x) = \sqrt{x^2 + \epsilon^2}$.

1. Show that $\phi_\epsilon(F) \in \mathbb{D}_{p,1}$.
2. Show that $|F| \in \mathbb{D}_{p,1}$ and that

$$\dot{\nabla}_s |F| = \begin{cases} \dot{\nabla}_s F & \text{if } F > 0 \\ 0 & \text{if } F = 0 \\ -\dot{\nabla}_s F & \text{if } F < 0. \end{cases}$$

3. If $G \in \mathbb{D}_{p,1}$, compute $\nabla(F \vee G)$.

Let B be the standard Brownian motion on $[0, 1]$ and $M = \sup_{t\in[0,1]} B(s)$. Let $\mathbf{Q} \cap [0, 1] = \{t_n, n \geq 0\}$. Consider

$$M_n = \sup_{s\in\{t_1,\cdots,t_n\}} B(s).$$

We admit that B attains its maximum at a unique almost-surely. Let
$T = \text{argmax}_{s\in[0,1]} B(s)$.

4. Show that M_n belongs to $\mathbb{D}_{p,1}$ and compute $\dot{\nabla} M_n$.
5. Prove that $M \in \mathbb{D}_{p,1}$ and that $\dot{\nabla} M = \mathbf{1}_{[0,T]}$.

Exercise 2.3 (Iterated divergence). For $U \in \mathcal{S}(\mathcal{H})$, i.e.

$$U = \sum_{j=1}^{n} f_j(\delta h_1, \cdots, \delta h_m) v_j$$

where $(v_1, \cdots, v_n)$ belong to $\mathcal{H}$ and f_j in the Schwartz space on $\mathbf{R}^m$. Let $\delta^{(2)}$ defined by the duality

$$\mathbf{E}\left[\delta^2 u^{\otimes(2)} G\right] = \mathbf{E}\left[\left\langle u^{\otimes(2)}, \nabla^{(2)} G\right\rangle_{\mathcal{H}\otimes\mathcal{H}}\right]$$

for any $G \in \mathbb{D}_{2,2}$. Show that

$$\delta^2(U^{\otimes(2)}) = \big(\delta U\big)^2 - \|U\|^2_{\mathcal{H}} - \text{trace}(\nabla U \circ \nabla U) - 2\delta(\langle \nabla U, U\rangle_{\mathcal{H}}).$$

Exercise 2.4 (Stratonovitch integral). The Itô integral has a major drawback: Its differential is not given by the usual formula but by the Itô formula. On the other hand, the Stratonovitch integral does satisfy the usual rule of differentiation but does not give to a martingale ! We see in this exercise that the Stratonovitch integral can be computed with δ and ∇. For $T_n = \{0 = t_0 < t_1 = 1/n < \ldots < t_n = 1\}$, let

$$\mathrm{d}B_{T_n}(t) = \sum_{i=0}^{n-1} \frac{B(t_{i+1}) - B(t_i)}{t_{i+1} - t_i} \, \mathbf{1}_{[t_i,\, t_{i+1}]}(t) \, \mathrm{d}t$$

be the linear affine interpolation of B. For any $\mathcal{H}$-valued random variable U, consider the Riemann-like sum

$$S_{T_n}^U = \sum_{i=0}^{n-1} \frac{B(t_{i+1}) - B(t_i)}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} \dot{U}(t) \, \mathrm{d}t.$$

The process U is said to be Stratonovitch integrable if the sequence $(S_{T_n}^U, \, n \geq 0)$ converges in probability as n goes to infinity.

Assume that U belongs to $\mathbb{D}_{1,2}(\mathcal{H})$ and that the map

$$\begin{aligned} [0,1] \times [0,1] &\longrightarrow \mathbf{R} \\ (s,t) &\longmapsto \dot{\nabla}_s \dot{U}(t) \end{aligned}$$

is continuous.

1. Show that U is Stratonovitch integrable and

$$\lim_{n \to \infty} S_{T_n}^U = \delta U + \int_0^1 \dot{\nabla}_r \dot{U}(r) \, \mathrm{d}r.$$

Indication: Verify that

$$S_{T_n}^U = \sum_{i=0}^{n-1} \frac{1}{t_{i+1} - t_i} \delta(I^1(\mathbf{1}_{[t_i, t_{i+1}]})) \int_{t_i}^{t_{i+1}} \dot{U}(t) \, dt.$$

Apply (2.17).

2. Find

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \left(\dot{U}(t_i) + \dot{U}(t_{i+1}) \right) \left(B(t_{i+1}) - B(t_i) \right).$$