A SHORT INTRODUCTION TO STOCHASTIC PDEs

CHARLES-EDOUARD BRÉHIER

Abstract. The main two aims of these lecture notes are: a definition of the space-time white noise and the study of second-order stochastic evolution equations of parabolic type. The main points are the regularity properties, in time and in space, of the solutions.

A few results on the numerical approximation are also given at the end.

The content is based on the lectures delivered at CERMICS in March 2014.

1. Introduction

1.1. **A parabolic semilinear PDE.** The Stochastic PDE considered reads:

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u(t,x) + f(u(t,x)) + \dot{W}(t,x), \quad t>0, x\in\mathcal{D}; \\ u(0,x) = u_0(x), \quad x\in\mathcal{D}; \\ u(t,x) = 0, \quad t>0, x\in\partial\mathcal{D}. \end{cases} \tag{1}$$

The spatial domain $\mathcal{D}$ is a bounded subset of $\mathbb{R}^d$, with a smooth boundary.

The nonlinearity f (to keep things simple) is a smooth function: the well-posedness and regularity problems are only due to the noise forcing. More generally, f is also allowed to depend directly on time t and on position x. When one wants to include spatial derivatives of u, not every f is allowed; for instance, for $d=1$, the theory applies when $f(t,x,u(t,x),\frac{\partial u}{\partial x}(t,x)) = \frac{\partial u}{\partial x}(t,x)$ or $u(t,x)\frac{\partial u}{\partial x}(t,x)$ (Burgers equation), but does not apply when $f(t,x,u(t,x),\frac{\partial u}{\partial x}(t,x)) = \big(\frac{\partial u}{\partial x}(t,x)\big)^2$.

The noise term $\dot{W}$ is introduced to represent a space-time white noise. Formally, $\dot{W}(t,x)$ is a Gaussian random variable, with mean zero, and with correlation:

$$\mathbb{E}[\dot{W}(t,x)\dot{W}(s,y)] = \delta(t-s,x-y).$$

Other notations for $\dot{W}$ might be found in the literature (for instance $\xi(t,x)$).

We will see that the theory for the well-posedness of (1) is limited to $d=1$. To apply the theory described in these lecture notes to higher dimensional equations, noise terms which are colored in space are introduced. For instance:

$$\mathbb{E}[\dot{W}_q(t,x)\dot{W}_q(s,y)] = \delta(t-s)q(x-y).$$

The smoother q, the more regular the noise and the solutions are.

Some natural questions are the following.

- Can we give a mathematical meaning to the notion of space-time white noise?
- What kind of stochastic integration theory can we use?
- What is then the concept of solutions?

For the third question, the answer is weak solutions (see Definition 3.1), which turn to be mild solutions, using Duhamel principle (see Theorem 3.2).

Date: CERMICS, March 2014.

I would like to thank the audience at the lectures for their questions and comments. Moreover I would like to thank especially Claude Le Bris, Mathias Rousset and Gabriel Stoltz for their careful reading of the notes and their useful comments, which helped improving as much as possible the presentation of the theory in these lecture notes.

There are (at least) two approaches (for a comparison, see for instance [3]). To present them, we assume that $f=0$, since as mentioned before the main task is to deal with the stochastic part, and that the initial condition is zero ($u_0=0$):

$$
\begin{cases}
\dfrac{\partial u}{\partial t}=\Delta u(t,x)+\dot{W}(t,x), \quad t>0, x\in\mathcal{D};\\
u(0,x)=0, \quad x\in\mathcal{D};\\
u(t,x)=0, \quad t>0, x\in\partial\mathcal{D}.
\end{cases}
\tag{2}
$$

The general semilinear case (1) follows by the addition of the terms appearing because of the initial condition and of the forcing term f in the mild formulation, by Duhamel principle. Those terms are more regular than the solution of (2).

1.2. **The random-field solution.** Based on Duhamel principle, we would like (in order to agree with the deterministic setting) to define the solution with a formula of the following type:

$$
\begin{aligned}
u(t,x)&=\int_{\mathcal{D}}\int_0^t G(s,t,x,y)\dot{W}(s,y)dsdy\\
&=\int_{\mathcal{D}}\int_0^t G(s,t,x,y)W(ds,dy),
\end{aligned}
\tag{3}
$$

where G denotes the Green function of $\frac{\partial}{\partial t}-\Delta$, on the domain $(0,+\infty)\times\mathcal{D}$ and zero boundary conditions. If the first line gives a consistent formulation with the SPDE (2), the second one is prefered and is the usual notation.

References on this approach are [8], [2]...

The function u is seen as a function of time and space variables (t,x), subjected to a random force that also depends on these two variables.

The approach requires to define stochastic integrals

$$\int_{\mathcal{D}}\int_0^t \psi(s,y)W(ds,dy).$$

for appropriate integrands ψ.

The theory begins with deterministic integrands ψ in the space $L^2([0,T]\times\mathcal{D})$ of scalar-valued square integrable functions. To be able to consider stochastic integrands, a filtration in time is introduced; a notion of so-called "worthy martingale measures" naturally arises, allowing a construction by an approximation procedure with "simple" processes.

1.3. **Functional setting.** We hide the space variable x, and we consider the solution to be a random function of time, with values in an appropriate (infinite dimensional) function space.

As a consequence, we will need stochastic integrals taking values not only in $\mathbb{R}$, but in infinite dimensional linear spaces. Since probability theory works well in Polish spaces, the values must be taken in separable Banach spaces. In fact, we restrict to the Hilbert space setting. To learn more about stochastic integration in Banach spaces, see the survey [4].

The example to keep in mind is $L^2(\mathcal{D})$. In full generality, the Hilbert space will be denoted by H.

We define A an unbounded linear operator on H, taking into account the Dirichlet boundary conditions: $Au=\Delta u$, with the domain $D(A)=H^2(\mathcal{D})\cap H_0^1(\mathcal{D})$.

To consider mild solutions, we introduce the semi-group $(S(t))_{t\geq 0}$; we often write $S(t)=e^{tA}$.

Similarly, we see the noise term as a function of time only; inspired by the notations used for SDEs/diffusion processes in finite dimension, we write the linear SPDE in the form

$$du(t)=Au(t)dt+dW(t). \tag{4}$$

Here W is the so-called cylindrical Wiener process; its definition and its properties, and the study of the solution of (4) are the main subject of these lecture notes.

The main reference for this approach is [6].

In the case of (1), both the random-field approach and the functional one give the same notion of weak/mild solutions. However, sometimes one or the other approach has to be preferred, depending on the regularity of the coefficients, or the definition of the noise correlations. If pointwise in space results are expected and required, it should be better to take the random-field approach. On the other hand, it seems that for the study of the long-time behavior and ergodicity properties it is better to follow the functional approach. Moreover, to consider semilinear variations of (4), general nonlinearities $F : H = L^2(\mathcal{D}) \to H$ can be considered, instead of only Nemytskii operators associated with some function f (for $u \in L^2(\mathcal{D})$, $F(u)(x) = f(u(x))$ for $x \in \mathcal{D}$). This is essential for instance when considering the *averaging principle* for systems of slow-fast SPDEs, see [1] for instance: in general the averaging coefficient $\overline{F}$ is no longer a Nemytskii operator.

1.4. **Other approaches (not developed here).** There is also (as for SDEs) a notion of weak solution in the probability sense, where the noise process is part of the solution and must be built (in the strong sense, the noise process is given and we associate a solution). The usual terminology is the concept of "martingale solutions". The solution is a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, and the martingale problem is formulated with a sufficienlty large set of test functions to get unique solutions (for instance, smooth functions depending on an arbitrary but finite number of modes).

When the linear operator A also depends on the solution, $A(u)$, one can use a variational approach to solve the SPDE $du(t) = A(u(t))u(t)dt + dW(t)$, with the theory of non-linear monotone operators. See for instance the lecture notes [5].

For other SPDEs (hyperbolic ones,...) other concepts of PDE solutions are currently studied.

Fully non-linear SPDEs have been considered, with the use of viscosity solutions.

Finally, many recent works used connexions with the theory of rough paths.

Some people also call SPDEs (for instance in the community of uncertainty quantification) some equations that might rather be called "PDEs with random coefficients". The noise perturbation, most of the time, takes values in large but finite dimensional spaces, solutions are defined for each realizationof the random coefficients, and the regularity is not modified by the introduction of the noise. This is a different subject, with different analysis and approximation challenges, which is not addressed here.

It is also important to mention that SPDEs with *multiplicative noise* can also be considered, with both random-field and functional approaches. Here, there is a choice to be made for the notion of stochastic integral (Itô or Stratonovitch), which has no effect in the additive case which for simplicity we develop here. Appropriate assumptions must be made on the diffusion coefficients. Basically, the construction of the Itô integral is made by approximation with adapted, left-continuous simple processes, along the same line as for the usual Itô stochastic integral with respect to Brownian Motion (see [7] for instance, or any reference book on stochastic calculus).

In the sequel, we only deal with the functional setting.

First, we want to understand why the cylindrical Wiener process is a suitable mathematical model for space-time white noise (sometimes denoted "STWN" in the sequel). Then we study the mild solutions of the heat equation perturbed with additive noise.

2. The cylindrical Wiener process

Once and for all, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

2.1. **Isonormal Gaussian process.** Let $\mathcal{H}$ be a separable (real) Hilbert space, with scalar product denoted by $|.|_{\mathcal{H}}$, and scalar product $< ., . >_{\mathcal{H}}$.

We mostly deal with the situation when $\dim(\mathcal{H}) = +\infty$, so that $\mathcal{H}$ possesses complete orthonormal systems indexed by $\mathbb{N}^*$. In fact, we could easily include the case $\dim(\mathcal{H}) < +\infty$ in our study, considering finite size complete orthonormal systems; in this case we recover usual objects like Gaussian vectors, Brownian Motion in $\mathbb{R}^N$, etc. The extension to the infinite dimensional situation allows to define properly space-time white noise.

In the definition below, $L^2(\Omega)$ denotes the Hilbert space of square-integrable (almost-surely defined) real random variables defined on Ω.

Definition 2.1. *An $\mathcal{H}$-isonormal Gaussian process is a mapping $\mathcal{W} : \mathcal{H} \to L^2(\Omega)$ (or equivalently a family of random variables $(\mathcal{W}(h))_{h\in\mathcal{H}}$) such that:*

- *for any $n \in \mathbb{N}^*$, and any $(h_1,\ldots,h_n) \in \mathcal{H}^n$, $(\mathcal{W}(h_1),\ldots,\mathcal{W}(h_n))$ is a Gaussian random vector, i.e. for any $(\lambda_1,\ldots,\lambda_n) \in \mathbb{R}^n$, the real random variable $\lambda_1\mathcal{W}(h_1) + \ldots + \lambda_n\mathcal{W}(h_n)$ has a (possibly degenerate) gaussian law and is centered;*
- *for any $h_1, h_2 \in \mathcal{H}$, the covariance of $\mathcal{W}(h_1)$ and $\mathcal{W}(h_2)$ is given by*

$$\mathbb{E}[\mathcal{W}(h_1)\mathcal{W}(h_2)] =< h_1, h_2 >_{\mathcal{H}} .$$

The first point is the Gaussian process property. It is well-known that a Gaussian process $(\mathcal{W}(h))_{h\in\mathcal{H}}$ is caracterized by its mean function

$$m : h \in \mathcal{H} \mapsto \mathbb{E}[\mathcal{W}(h)],$$

and its covariance function

$$c : (h_1, h_2) \in \mathcal{H}^2 \mapsto \mathrm{Cov}(\mathcal{W}(h_1), \mathcal{W}(h_2)) = \mathbb{E}[\mathcal{W}(h_1)\mathcal{W}(h_2)] - \mathbb{E}[\mathcal{W}(h_1)]\mathbb{E}[\mathcal{W}(h_2)].$$

Here $m(h) = 0$ for all $h \in \mathcal{H}$, by the centering assumption.

The mapping $\mathcal{W}$ is in fact linear: we have the equality in $L^2(\Omega)$ and almost surely of random variables $\mathcal{W}(\lambda_1 h_1 + \lambda_2 h_2) = \lambda_1\mathcal{W}(h_1) + \lambda_2\mathcal{W}(h_2)$.

In the literature, $\mathcal{W}$ is also often called the white noise mapping.

The existence of an $\mathcal{H}$-isonormal Gaussian process is ensured by the following result.

Proposition 2.2. *Let $(\eta_n)_{n\in\mathbb{N}^*}$ be a sequence of independent and identically distributed real random variables, with distribution $\mathcal{N}(0,1)$ (standard gaussian random variable).*

Let $(\epsilon_n)_{n\in\mathbb{N}^*}$ a complete orthonormal system of the separable Hilbert space $\mathcal{H}$.

First, we define $\mathcal{W}(\epsilon_n) = \eta_n$.

Then, by linearity, if $h \in \mathcal{H}$ has the expansion $h = \sum_{n\in\mathcal{N}^*} h_n\epsilon_n$, we set $\mathcal{W}(h) = \sum_{n=1}^{+\infty} h_n\eta_n$.

Then $\mathcal{W}$ defines an $\mathcal{H}$-isonormal Gaussian process.

Sketch of proof: First, for $h = \sum_{n\in\mathcal{N}^*} h_n\epsilon_n$, the random variable $\mathcal{W}(h) = \sum_{n=1}^{+\infty} h_n\eta_n$ is well-defined and has a Gaussian law. Indeed, $\mathcal{W}(h) = \lim_{N\to+\infty}\sum_{n=1}^{N} h_n\eta_n$ (the limit holds in $L^2(\Omega)$); for each $N \in \mathbb{N}^*$, the truncated series has a Gaussian law (the η_n are independent and Gaussian), and this property extends to the limit.

The proof of the first point in the definition is a slight generalization of the above argument.

It is also clear that $\mathbb{E}[\mathcal{W}(h)] = 0$ for any $h \in \mathcal{H}$.

Finally, we just have to compute $\mathbb{E}[\mathcal{W}(h)^2]$ for any $h \in \mathcal{H}$. A polarization formula will then allow to obtain the expression for $\mathbb{E}[\mathcal{W}(h^1)\mathcal{W}(h^2)]$.

$$\begin{aligned}
\mathbb{E}[\mathcal{W}(h)^2] = \mathbb{E}\lim_{N\to+\infty}\sum_{1\le k,l\le N} h_k h_l \eta_k\eta_l &= \lim_{N\to+\infty}\mathbb{E}\sum_{1\le k,l\le N} h_k h_l \eta_k\eta_l \\
&= \lim_{N\to+\infty}\sum_{1\le k,l\le N} h_k h_l \mathbb{E}[\eta_k\eta_l] \\
&= \lim_{N\to+\infty}\sum_{1\le k\le N} (h_k)^2 = |h|^2_{\mathcal{H}}.
\end{aligned}$$

$\square$

In the next sections, we specify this construction for $\mathcal{H} = L^2([0,T])$ and $\mathcal{H} = L^2([0,T] \times \mathcal{D})$.

2.2. **Brownian Motion and white noise in time.** The main aim of this section is to recall the definition of Brownian Motion, to show that it can be obtained thanks to an isonormal process, and to interpret this property as the following statement: "The white noise is the derivative of the Brownian Motion".

Formally, a Stochastic Differential Equation writes (for instance):

$$\frac{dx_t}{dt} = f(x_t) + \sigma(x_t)\dot{W}_t,$$

where x_t denotes the value of a process at time t.

The term $\dot{W}$ is meant to represent a gaussian white noise in time: $\mathbb{E}\dot{W}_t = 0$ and $\mathbb{E}\dot{W}_t\dot{W}_s = \delta(t-s)$.

It turns out that the SDE has to be interpreted at an integrated-in-time level, in the Itô sense:

$$dx_t = f(x_t)dt + \sigma(x_t)dB_t,$$

where B is a Brownian Motion; formally $\dot{W}_t = \frac{dB_t}{dt}$.

Recall the definition of a Brownian Motion:

Definition 2.3. *Let $T \in \mathbb{R}^+$ given. A Brownian Motion is a stochastic process $(t,\omega) \in [0,T] \times \Omega \mapsto B_t(\omega)$, such that:*

(1) *$B_0 = 0$ a.s.;*
(2) *the increments are stationary: for any $0 \leq s \leq t \leq T$, $B_t - B_s \sim B_{t-s}$;*
(3) *the increments are independent: for any $n \in \mathbb{N}^*$, for any $0 = t_0 \leq t_1 \leq \ldots \leq t_n \leq T$, $(B_{t_{i+1}} - B_{t_i})_{0 \leq i \leq n-1}$ are independent;*
(4) *the process is Gaussian: for any $n \in \mathbb{N}^*$, for any $0 \leq t_1 \leq \ldots \leq t_n \leq T$ and any $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, $\lambda_1 B_{t_1} + \ldots + \lambda_n B_{t_n}$ is a real Gaussian random variable;*
(5) *for any $t \geq 0$, $B_t \sim \mathcal{N}(0,t)$;*
(6) *the trajectories are almost surely continuous: for $\mathbb{P}$-almost every $\omega \in \Omega$, $t \in [0,T] \mapsto B_t(\omega)$ is continuous.*

Usually, the variable ω is not written.

Brownian Motion can be constructed thanks to the Kolmogorov extension Theorem (given finite dimensional laws, a process can be defined). The continuity property (6) is a consequence of the Kolmogorov-Centzov continuity criterion, see [7], [6] for instance. In fact, a process satisfying conditions (1) to (5) has a version with continuous trajectories.

We recall the notion of a modification (or version) of a stochastic processes. Given some index set $\mathbb{T} \subset \mathbb{R}^d$, and two stochastic processes $(X_t)_{t \in \mathbb{T}}$ and $(Y_t)_{t \in \mathbb{T}}$, *i.e.* measurable mappings $\mathbb{T} \times \Omega \to E$ (E being some Banach space), we say that Y is a modification (or version) of X if for any $t \in \mathbb{T}$ we have $\mathbb{P}(X_t = Y_t) = 1$. Notice that if X and Y are two versions of the same stochastic process, they have the same law (since it is caracterized by finite-dimensional distributions). The Kolmogorov-Centzov continuity criterion is a famous and simple result ensuring existence of a modification of a stochastic process $(X_t)_{t \in \mathbb{T}}$ with almost surely (Hölder)-continuous paths, under a condition of the moments of increments $X_t - X_s$, for $s, t \in \mathbb{T}$; the precise range of Hölder exponents depends on dimension d of the index variable t and on the estimate: if there exists $\alpha, \beta > 0$ and a constant $C \in \mathbb{R}^+$ such that

$$\mathbb{E}\|X(t) - X(s)\|_E^\alpha \leq C|t-s|_{\mathbb{R}^d}^{d+\beta},$$

then the process $(X_t)_{t \in \mathbb{T}}$ admits a version with almost surely γ-Hölder continous paths, for any $0 < \gamma < \frac{\beta}{\alpha}$.

It is worth noting that two versions of a stochastic process with (almost surely) continuous paths satisfy the stronger result: $\mathbb{P}(X_t = Y_t, \forall t \in \mathbb{T}) = 1$.

Here we give another construction of Brownian Motion:

Proposition 2.4. *Let $\mathcal{W}$ be an $L^2([0,T])$-isonormal Gaussian process. We set $B_t = \mathcal{W}(\mathbb{1}_{[0,t]})$. Then (up to a continuous version) $B = (B_t)_{t\in[0,T]}$ is a Brownian Motion.*

Sketch of proof: The fact that B is a Gaussian process directly follows from the definition of an isonormal process. The stationarity is also straightforward, as well as the fact that $B_t \sim \mathcal{N}(0,t)$.

It is sufficient to prove that $\mathbb{E}[B_tB_s] = t\wedge s (:= \inf(t,s))$ for $0 \leq s,t \leq T$: the independence then follows from the fact that the process is Gaussian and from $\mathbb{E}[B_s(B_t - B_s)] = 0$ for $0 \leq s \leq t \leq T$. We just check that

$$\mathbb{E}[B_tB_s] = \mathbb{E}[\mathcal{W}(\mathbb{1}_{[0,t]})\mathcal{W}(\mathbb{1}_{[0,s]})] = < \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]} >_{L^2([0,T])} = s\wedge t.$$

□

In this case, the Kolmogorov-Centzov criterion ensures that there exists a version of Brownian Motion with γ-Hölder continuous paths, for any $\gamma < 1/2$: we use the expression of the covariance and the Gaussian property to control moments of any order: $\mathbb{E}|B_t - B_s|^{2n} \leq C_n|t-s|^n$ for any $n \in \mathbb{N}^*$, so that all exponents γ such that $\gamma < \frac{n-1}{2n}$ for some $n \in \mathbb{N}^*$ are allowed, and hence all $\gamma < 1/2$.

When Brownian Motion is considered, it is always meant that it is a version with almost surely continuous paths.

Notice that in this context, we get the Itô isometry property (for deterministic integrands): for $h \in L^2([0,T])$, $\int_0^T h(t)dB_t := \mathcal{W}(h)$ is well-defined, is Gaussian and satisfies

$$\mathbb{E}\Big|\int_0^T h(t)dB_t\Big|^2 = \int_0^T h(t)^2dt. \tag{5}$$

2.3. **Space-time white noise and the cylindrical Wiener process.** *Heuristics: for our SPDE, we want to define a quantity $W(t)$ such that $\frac{dW(t)}{dt}$ is space-time white noise. We have seen that $W(t)$ must have the behavior of a Brownian Motion in time, but of a white noise in space. We thus consider a white noise mapping, and "integrate" in time.*

We consider the separable Hilbert space $\mathcal{H} = L^2([0,T]\times\mathcal{D})$, and we give $\mathcal{W}$ a $\mathcal{H}$-isonormal Gaussian process.

Definition 2.5. *For $t \in (0,T]$, we define for $\phi \in L^2(\mathcal{D})$*

$$\mathcal{W}_t(\phi) = \mathcal{W}(\mathbb{1}_{[0,t]}\otimes\phi).$$

Then $\frac{1}{\sqrt{t}}\mathcal{W}_t(.)$ is an $L^2(\mathcal{D})$-isonormal Gaussian process.

An $L^2(\mathcal{D})$ isonormal processes being understood as a way to represent white noise in the space variable $x \in \mathcal{D}$, for a fixed time t, $\mathcal{W}_t$ is thus a spatial white noise.

Proposition 2.6. *Let $(e_n)_{n\in\mathbb{N}^*}$ be a complete orthonormal system of $L^2(\mathcal{D})$.*

Define $\beta_n : t \in [0,T] \mapsto \mathcal{W}(\mathbb{1}_{[0,t]}\otimes e_n)$. Then:

- *$(\beta_n)_{n\in\mathbb{N}}$ are independent Brownian Motions;*
- *for any $\phi = \sum_{n\in\mathbb{N}^*}\phi_n e_n \in L^2(\mathcal{D})$, we have (almost surely)*

$$\mathcal{W}_t(\phi) = \sum_{n\in\mathbb{N}^*}\beta_n(t)\phi_n.$$

Proof First, for each $n \in \mathbb{N}^*$, it is easily checked that β_n is a centered Gaussian process, since $\beta_n(t_m) = \mathcal{W}(\mathbb{1}_{[0,t_m]}\otimes e_n)$.

We have the following formula: for $n, m \in \mathbb{N}^*$, $0 \leq s, t \leq T$,

$$\begin{aligned}
\mathbb{E}[\beta_n(t)\beta_m(s)] &= \mathbb{E}[\mathcal{W}(\mathbb{1}_{[0,t]} \otimes e_n)\mathcal{W}(\mathbb{1}_{[0,s]} \otimes e_m)] \\
&=< \mathbb{1}_{[0,t]} \otimes e_n, \mathbb{1}_{[0,s]} \otimes e_m >_{L^2([0,T]\times\mathcal{D})} \\
&= t \wedge s < e_n, e_m >^2_{L^2(\mathcal{D})} = (t \wedge s)\delta_{n,m}.
\end{aligned}$$

The independence of $\beta_n(t)$ and $\beta_m(s)$ for $n \neq m$ follows directly; by the Gaussian property, it is in fact equivalent to the independence of $(\beta_n)_{n\in\mathbb{N}}$.

Taking $n = m$ in the above formula proves that β_n is a Brownian Motion.

Finally, the expression of $\mathcal{W}_t(\phi)$ is clear for $\phi \in \mathrm{Span}\{e_1, \ldots, e_N\}$, and follows taking a limit $N \to +\infty$. □

We are now able to introduce the following object, in a more general setting, and to show how it relates to the notion of space-time white noise.

Definition 2.7. *Let $(e_n)_{n\in\mathbb{N}^*}$ be a complete orthonormal system of a separable Hilbert space H, and $T \in \mathbb{R}^+$.*

Let $(\beta_n)_{n\in\mathbb{N}^*}$ an i.i.d. sequence of Brownian Motions.

We set

$$W(t) = \sum_{n\in\mathbb{N}^*} \beta_n(t)e_n. \tag{6}$$

W is called a cylindrical Wiener process in H.

A good example to think about is of course $H = L^2(\mathcal{D})$.

Another example is $H = \mathbb{R}^N$, which is a finite dimensional Hilbert space; the setting also applies (with $\{1, \ldots, N\}$ instead of $\mathbb{N}^*$), and then we just obtain a N-dimensional Brownian Motion. The word "cylindrical" precisely means that $W(t)$ is defined so that when truncating at a level N we recover the usual N-dimensional notion.

However, the series from which $W(t)$ is defined does not converge in H, for a fixed $t > 0$ since the strong law of large numbers ensures that when $N \to +\infty$

$$\sum_{n=1}^{N} |\beta_n(t)|^2 \sim Nt.$$

This negative result is compensated by the fact that the series is convergent in any larger Hilbert spaces $\tilde{H}$ such that $H \subset \tilde{H}$ and the embedding is Hilbert-Schmidt. We recall the definition of such operators.

Definition 2.8. *Let $H, \tilde{H}$ two separable Hilbert spaces, and $L \in \mathcal{L}(H, \tilde{H})$ a bounded linear operator. We say that L is a Hilbert-Schmidt operator if there exists a complete orthonormal system $(b_n)_{n\in I}$ of H, $I \subset \mathbb{N}^*$, such that $\sum_{n\in I} |Lb_n|^2_{\tilde{H}} < +\infty$.*

In this situation, the value $\sum_{n\in I} |Lb_n|^2_{\tilde{H}}$ is finite for any complete orthonormal system, and does not depend on the choice of such a basis.

Moreover L is Hilbert-Schmidt if and only if LL^* and/or L^*L are trace-class operators:

$$\sum_{n\in I} |Lb_n|^2_{\tilde{H}} = \sum_{n\in I} < Lb_n, Lb_n >_{\tilde{H}} = \sum_{n\in I} < L^*Lb_n, b_n >_{\tilde{H}} .$$

We denote by $\mathcal{L}_2(H, \tilde{H})$ the vector space of Hilbert-Schmidt operators from H to $\tilde{H}$; it is a Hilbert space for the following norm:

$$\|L\|^2_{\mathcal{L}_2(H,\tilde{H})} = Tr(LL^*) = \sum_{n\in I} |Lb_n|^2_{\tilde{H}}.$$

A typical example of Hilbert-Schmidt operator is the inclusion map $i : L^2(D) \to H^{-s}(\Omega)$ for $s > d/2$.

The convergence statement on $W(t)$ is a consequence of the following more general result:

Proposition 2.9. *If $L \in \mathcal{L}_2(H, \tilde{H})$, then $LW(t) = \sum_{n\in I} \beta_n(t) L e_n$ is a well-defined $L^2(\Omega)$-random variable in $\tilde{H}$. Moreover the result does not depend on the choice of the complete orthonormal system $(e_n)_{n\in I}$ of H.*

Example 1: if $h \in H$, the application $< h, . >: H \to \mathbb{R}$ is Hilbert-Schmidt, and its norm is equal to $\| < h, . > \|_{\mathcal{L}_2(H)} = |h|_H$. As a consequence, $< h, W(t) >$ is a well-defined real random variable. We check that $< h, W(t) >= \mathcal{W}_t(h)$, and we see that $W(t)$ is a very natural object related to white noise, which allows to recover the white-noise mapping $\mathcal{W}$.

Example 2: assume that $H \subset \tilde{H}$ and that the inclusion map $i : H \to \tilde{H}$ is a linear, Hilbert-Schmidt operator. Then $W(t)$ can be identified with $i(W(t)) \in \tilde{H}$, which is a well-defined object in $\tilde{H}$. The properties of this object do not depend on the choice of $\tilde{H}$ and i. Moreover, given H a separable Hilbert space, there always exists such a choice: if $(e_n)_{n\in I}$ and $(\tilde{e}_n)_{n\in I}$ are complete orthonormal systems of H and $\tilde{H}$ respectively, define $i : H \to \tilde{H}$ by linearity and $i(e_n) = \frac{1}{n}\tilde{e}_n$.

Therefore, in the sequel we only work with $W(t)$, and we have to be careful to give a meaning to the associated quantities.

2.4. **Stochastic integral.** Recall that we want to define stochastic integrals where the noise $W(t)$ is given by a H-cylindrical Wiener process, and also taking values in H (or possibly a different Hilbert space).

For our purpose, we only need deterministic integrands. The extension to progressively measurable processes (or predictable?) follows the same lines as for the usual real-valued Itô stochastic integral with respect to Brownian Motion - thanks to the approximation by elementary processes. See for instance [7].

We in fact rely on this construction and extend it to get Hilbert space-valued integrals.

Let $L : t \in [0, T] \mapsto \mathcal{L}(H, \tilde{H})$, and W be a cylindrical Wiener process in H.

Proposition 2.10. *If $\int_0^T \|L(t)\|^2_{\mathcal{L}_2(H,\tilde{H})} dt < +\infty$, then we can define the stochastic integral*

$$\int_0^T L(t) dW(t) = \sum_{i,j\in I,J} \int_0^T < L(t)e_j, \tilde{e}_i >_{\tilde{H}} d\beta_j(t)\tilde{e}_i,$$

where $(e_j)_{j\in J}$ and $(\tilde{e}_i)_{i\in I}$ are complete orthonormal systems of H and $\tilde{H}$ respectively.

The result does not depend on the choice of those systems.

Moreover we have the Itô isometry property $\mathbb{E}|\int_0^T L(t)dW(t)|^2_{\tilde{H}} = \int_0^T \|L(t)\|^2_{\mathcal{L}_2(H,\tilde{H})} dt$.

Elements of proof We only check the Itô isometry property:

$$\begin{aligned}
\mathbb{E}\Big|\int_0^T L(t)dW(t)\Big|_{\tilde{H}}^2 &= \sum_{i\in I}\mathbb{E}\Big|\sum_{j\in J}\int_0^T < L(t)e_j, \tilde{e}_i > d\beta_j(t)\Big|^2\\
&= \sum_{i\in I}\sum_{j\in J}\mathbb{E}\Big|\int_0^T < L(t)e_j, \tilde{e}_i > d\beta_j(t)\Big|^2\\
&= \sum_{i\in I}\sum_{j\in J}\Big|\int_0^T < L(t)e_j, \tilde{e}_i > dt\Big|^2\\
&= \sum_{j\in J}\int_0^T |L(t)e_j|^2 dt\\
&= \int_0^T \|L(t)\|^2_{\mathcal{L}_2(H,\tilde{H})}dt.
\end{aligned}$$

Notice that we have used the Itô isometry (5) for the stochastic integral with respect to a one-dimensional Brownian Motion β_j.

□

3. Solutions of linear SPDEs perturbed by space-time white noise: stochastic convolution

3.1. **General setting.** In an abstract form, we want to solve SPDEs written in the Hilbert space H

$$du(t) = Au(t)dt + BdW(t), \tag{7}$$

with an initial condition $u(0) = u_0 \in H$, $\big(W(t)\big)_{t\in[0,T]}$ a cylindrical Wiener process in H (it could be in another space U), $B \in \mathcal{L}(H)$, and $A : H \to H$. We assume that A generates a strongly continuous semi-group $(S(t))_{t\in[0,T]}$.

The linear operator B is here to take into account possible correlations in space. The case $B = I$ corresponds to space-time white noise. We often set $Q = BB^*$, to denote the covariance operator.

Definition 3.1. *u is a weak solution of the SPDE (7) if for any $\xi \in D(A^*)$ and any $t > 0$ we have*

$$< u(t), \xi >=< u_0, \xi > + \int_0^t < u(s), A^*\xi > ds + < BW(t), \xi > .$$

Theorem 3.2. *Assume*

$$\int_0^T \|S(s)B\|^2_{\mathcal{L}_2(H,H)}ds < +\infty. \tag{8}$$

Then (7) admits a unique weak solution, which satisfies:

$$u(t) = S(t)u_0 + \int_0^t S(t-s)BdW(s).$$

A function u satisfying such a formula is a mild (or integral) solution.

Moreover $\mathbb{E}|u(t)|_H^2 = \int_0^t \|S(s)B\|^2_{\mathcal{L}_2(H,H)}ds$.

See Section 5.2 in [6], for the application of Duhamel principle in this setting.

When $u_0 = 0$, the solution is denoted by W_A and is called the stochastic convolution (or $W_{A,Q}$ if we want to make precise the dependence with respect to B or Q). When $Q = I$, we simply write W_A.

The condition (8) is precisely the one required to be able to define the stochastic integral in H. If it is removed, there exists no H-valued solution.

3.2. **The particular case of the heat equation in a bounded domain.** We consider the heat equation in the smooth bounded domain $\mathcal{D} \subset \mathbb{R}^d$, with homogeneous Dirichlet boundary conditions.

Therefore A is unbounded, self-adjoint, negative, with a compact inverse. Thus there exists $(e_k)_{k\in\mathbb{N}^*}$ a complete orthonormal system of $H = L^2(\mathcal{D})$ and a non-decreasing sequence $(\lambda_k)_{k\in\mathbb{N}^*}$ of positive real numbers such that $Ae_k = -\lambda_k e_k$. We denote the semi-group $S(t) = e^{tA}$.

According to the previous Theorem, the SPDE (7) admits a solution in H if (and only if, in fact)

$$\int_0^T \|S(s)B\|^2_{\mathcal{L}_2(H,H)} ds < +\infty.$$

Let us consider space-time white noise: $B = I$. Then

$$\begin{aligned} \int_0^T \|S(s)B\|^2_{\mathcal{L}_2(H,H)} ds &= \int_0^T \mathrm{Tr}(e^{2sA}) ds \\ &= \sum_{k=1}^{+\infty} \frac{1 - e^{-\lambda_k T}}{\lambda_k}. \end{aligned}$$

Since $\lambda_k \sim ck^{2/d}$, the above series is convergent if and only if $d = 1$.

Theorem 3.3. *The linear heat equation on a bounded domain, perturbed by space-time white noise, (7), admits a (mild) solution in $L^2(\mathcal{D})$ if and only if $d = 1$.*

If $d > 1$, on the one hand, it is not difficult to see that the stochastic convolution with space-time white noise still can be defined as a process with values in spaces of distributions (for instance: Sobolev spaces of negative order). This statement rises problems when you consider semi-linear equations.

On the other hand, another way to work with such SPDEs is to consider correlations in space, so that the equation is perturbed by noise which is white in time and colored in space. In fact, a sufficient condition on $Q = BB^*$ such that the stochastic convolution belongs to $L^2(\mathcal{D})$ is: $\mathrm{Tr}((-A)^{-1+\epsilon}Q) < +\infty$ - where fractional powers of $-A$ are defined later (by the spectral theorem, since A is self-adjoint with compact resolvent).

Notice that, expanded in the complete orthonormal system $(e_k)_{k\in\mathbb{N}^*}$ of H, the components of the stochastic convolution W_A are independent one-dimensional Ornstein-Uhlenbeck processes, depending on the respective eigenvalues $(\lambda_k)_{k\in\mathbb{N}^*}$: for any $t \geq 0$

$$W_A(t) = \sum_{k=1}^{+\infty} \int_0^t e^{-\lambda_k(t-s)} d\beta_k(s) e_k.$$

3.3. **Semi-linear equations with Lipschitz nonlinearity.** It is not difficult (thanks to a usual fixed point argument) to define weak/mild solutions of semilinear SPDEs

$$du(t) = Au(t)dt + F(u(t))dt + BdW(t),$$

where $F : H \to H$ is Lipschitz.

We could also consider equations with multiplicative noise

$$du(t) = Au(t)dt + F(u(t))dt + B(u(t))dW(t),$$

where $B : H \to \mathcal{L}(H)$ satisfies a Lipschitz condition in an appropriate norm (depending on A).

The weakening of those conditions depends crucially on the regularity properties (in time and in space) of the solutions. Since (in general) the less regular part is the stochastic one, in the next Section we study regularity in time and in space for the linear SPDE (7).

4. Regularity properties

The first remark is that since we deal with a parabolic second-order (S)PDE, regularities in time and in space are related: "one derivative in time is equivalent to two derivatives in space", and the same holds for "fractions". This statement holds for deterministic PDEs, as well as in our setting, as shown below.

Instead of giving a general statement, depending on the conditions on A and B, we focus on one case: dimension 1, and space-time white noise. We will also provide (as a remark) a comparison with the case when Q is trace-class (i.e. B is Hilbert-Schmidt, so that $BW(t) \in H$).

Before, we introduce some notation (classical in the deterministic setting), which allow to simplify the proofs.

4.1. A few notations.

Recall that A is unbounded, self-adjoint, negative, with a compact inverse. Thus there exists $(e_k)_{k\in\mathbb{N}^*}$ a complete orthonormal system of $H = L^2(\mathcal{D})$ and a non-decreasing sequence $(\lambda_k)_{k\in\mathbb{N}^*}$ of positive real numbers such that $Ae_k = -\lambda_k e_k$. We denote the semi-group $S(t) = e^{tA}$.

For $u = \sum_{k\in\mathbb{N}^*} u_k e_k$, we have $e^{tA}u = \sum_{k\in\mathbb{N}^*} e^{-\lambda_k t} u_k e_k$.

Definition 4.1. *Let $\alpha \geq 0$. We define*

$$D(-A)^\alpha = \left\{ u = \sum_{k=1}^{+\infty} u_k e_k \in H; |u|_\alpha^2 := \sum_{k=1}^{+\infty} (\lambda_k)^{2\alpha} |u_k|^2 < +\infty \right\},$$

and for $u \in D(-A)^\alpha$

$$(-A)^\alpha u = \sum_{k\in\mathbb{N}^*} \lambda_k^\alpha u_k e_k \in H.$$

If $\alpha \leq 0$, we define $D(-A)^\alpha$ as the dual of $D(-A)^{-\alpha}$, and $|u|_\alpha^2 := \sum_{k=1}^{+\infty} (\lambda_k)^{2\alpha} |u_k|^2$ for $u = \sum_{k=1}^{+\infty} u_k e_k \in H$.

Examples (in the special case): $D(-A)^0 = L^2(0,1)$, $D(-A) = H^2(0,1) \cap H_0^1(0,1)$, $D(-A)^{1/2} = H_0^1(0,1)$, $D(-A)^{-1/2} = H^{-1}(0,1)$, $D(-A)^\alpha = H^{2\alpha}(0,1)$ if $0 < \alpha < 1/4$; $D(-A)^{1/4+\epsilon} \subset L^\infty(0,1)$, $\epsilon > 0$.

The following regularization properties of the semi-group are used below:

Proposition 4.2. *For any $\sigma \in [0,1]$, there exists $C_\sigma > 0$ such that we have:*

(1) *for any $t > 0$ and $x \in H$*

$$|e^{tA}x|_{(-A)^\sigma} \leq C_\sigma t^{-\sigma} e^{-\frac{\lambda}{2}t} |x|_H.$$

(2) *for any $0 < s < t$ and $x \in H$*

$$|e^{tA}x - e^{sA}x|_H \leq C_\sigma \frac{(t-s)^\sigma}{s^\sigma} e^{-\frac{\lambda}{2}s} |x|_H.$$

(3) *for any $0 \leq s < t$ and $x \in D(-A)^\sigma$*

$$|e^{tA}x - e^{sA}x|_H \leq C_\sigma (t-s)^\sigma e^{-\frac{\lambda}{2}s} |x|_{(-A)^\sigma}.$$

Here $\lambda := \lambda_1 = \min(\lambda_i)_{i\in\mathbb{N}^*}$.

Elements of proof

We write $x = \sum_{k=1}^{+\infty} x_k e_k$ a generic element of H.

For (1), we write

$$
\begin{aligned}
|e^{tA}x|^2_{(-A)^\sigma} &= \sum_{k=1}^{+\infty} e^{-2\lambda_k t}\lambda_k^{2\sigma}|x_k|^2 \\
&\leq e^{-\lambda t}\sum_{k=1}^{+\infty}\frac{t^{2\sigma}}{t^{2\sigma}}e^{-\lambda_k t}\lambda_k^{2\sigma}|x_k|^2 \\
&\leq t^{-2\sigma}e^{-\lambda t}C_\sigma\sum_{k=1}^{+\infty}|x_k|^2,
\end{aligned}
$$

where $C_\sigma = \sup_{\tau\geq 0} e^{-\tau}\tau^{2\sigma}$.

We now focus on (2) and (3):

$$
\begin{aligned}
|e^{tA}x - e^{sA}x|_H^2 &= \sum_{k=1}^{+\infty}\big(e^{-t\lambda_k} - e^{-s\lambda_k}\big)^2|x_k|^2 \\
&= \sum_{k=1}^{+\infty} e^{-2\lambda_k s}\big(e^{-(t-s)\lambda_k} - 1\big)^2|x_k|^2 \\
&\leq 2^{1-\sigma}\sum_{k=1}^{+\infty} e^{-2\lambda_k s}|t-s|^{2\sigma}\lambda_k^{2\sigma}|x_k|^2.
\end{aligned}
$$

The key estimate $|e^{-x} - e^{-y}| \leq 2^{1-\sigma}|x-y|^\sigma$ for any $x, y > 0$ and $\sigma \in [0,1]$ is obtained by interpolation of the corresponding ones when $\sigma = 0$ and $\sigma = 1$.

To obtain (2), we use $e^{-2\lambda_k s}\lambda_k^{2\sigma} \leq e^{-\lambda s}s^{-\sigma}C_\sigma$. The last estimate (3) is a simple consequence of $e^{-2\lambda_k s} \leq e^{-2\lambda s}$.

□

4.2. **Regularity in time.** We start with a very general result, based on an argument called the "factorization method". See [6].

Proposition 4.3. *It there exists $\alpha > 0$ such that $\int_0^T t^{-\alpha}\|S(t)B\|_{\mathcal{L}_2(H,H)}dt < +\infty$, then W_A admits a continuous version.*

This argument only proves continuity, and we can obtain better results of Hölder regularity using the well-known Kolmogorov-Centzov criterion.

Recall that we consider the heat equation in $(0,1)$ ($d = 1$) with Dirichlet boundary conditions.

We first start with $Q = I$ (Space-Time White Noise):

Theorem 4.4. *For any $\alpha > 0$ and any $T \in \mathbb{R}^+$, there exists $C(T,\alpha) \in \mathbb{R}^+$ such that for any $0 \leq s \leq t \leq T$ we have*

$$
\mathbb{E}|W_A(t) - W_A(s)|_H^2 \leq C(T,\alpha)|t-s|^{1/2-\alpha}.
$$

As a consequence, W_A admits a version which is $1/4-\alpha$ Hölder continuous paths with values in H, for any $\alpha > 0$.

Proof We decompose, for $0 \leq s \leq t \leq T$

$$
\begin{aligned}
W_A(t) - W_A(s) &= \int_0^s \big(e^{(t-r)A} - e^{(s-r)A}\big)dW(r) \\
&+ \int_s^t e^{(t-r)A}dW(r),
\end{aligned}
$$

where on the right-hand side the stochastic integrals are independent.

For the first term, we write

$$\begin{aligned}
\mathbb{E}|\int_0^s \big(e^{(t-r)A}-e^{(s-r)A}\big)dW(r)|_H^2 &= \int_0^s \mathrm{Tr}\big((e^{(t-s)A}-I)e^{2(s-r)A}(e^{(t-s)A}-I)\big)dr\\
&= \int_0^s \mathrm{Tr}\big((e^{(t-s)A}-I)(-A)^{-1/4+\alpha/2}(-A)^{-1/2-\alpha/2}(-A)^{1-\alpha/2}e^{2(s-r)A}(-A)^{-1/4+\alpha/2}(e^{(t-s)A}-I)\big)dr\\
&\leq \mathrm{Tr}((-A)^{-1/2-\alpha/2})|(e^{(t-s)A}-I)(-A)^{-1/4+\alpha/2}|_{\mathcal{L}(H)}^2\int_0^s |(-A)^{1-\alpha/2}e^{2(s-r)A}|_{\mathcal{L}(H)}dr\\
&\leq C(t-s)^{2(1/4-\alpha/2)}s^{\alpha/2}.
\end{aligned}$$

Let us comment on the way we have chosen to decompose the product. First, $\mathrm{Tr}((-A)^{-1/2-\alpha/2})<+\infty$ if and only if $\alpha>0$. We then write that $\mathrm{Tr}((-A)^{-1/2-\alpha/2}L)\leq \mathrm{Tr}((-A)^{-1/2-\alpha/2})|L|_{\mathcal{L}(H)}$ for any $L\in\mathcal{L}(H)$.

The next factor to look at is the integral: using point (1) of Proposition 4.2,

$$\int_0^s |(-A)^{1-\alpha/2}e^{2(s-r)A}|_{\mathcal{L}(H)}dr\int_0^s \frac{C}{(s-r)^{1-\alpha/2}}dr<+\infty$$

if and only if $\alpha>0$.

The factor giving the order of convergence is the remaining one, and we use point (3) of Proposition 4.2 to conclude.

For the second term, we have with the Itô isometry

$$\begin{aligned}
\mathbb{E}|\int_s^t e^{(t-r)A}dW(r)|_H^2 &= \int_s^t \mathrm{Tr}(e^{2(t-r)A})dr\\
&= \int_s^t \mathrm{Tr}((-A)^{-1/2-\alpha}(-A)^{1/2+\alpha}e^{2(t-r)A})dr\\
&\leq \mathrm{Tr}((-A)^{-1/2-\alpha})\int_s^t |(-A)^{1/2+\alpha}e^{2(t-r)A}|_{\mathcal{L}(H)}dr\\
&\leq C(\alpha)\int_s^t \frac{1}{(t-r)^{1/2+\alpha}}dr=C(T,\alpha)(t-s)^{1/2-\alpha}.
\end{aligned}$$

□

It is useful to compare the previous result with the regularity of solutions of SDEs driven by Brownian Motion: since $\mathbb{E}|B_t-B_s|^2=|t-s|$, there exists (Kolmogorov-Centzov continuity criterion) a modification with $1/2-\alpha$ Hölder continuous trajectories, for any $\alpha>0$.

The regularity of solutions of SDEs is recovered with smoother noise:

Theorem 4.5. *Assume $Tr(Q)<+\infty$.*

For any $\alpha>0$ and any $T\in\mathbb{R}^+$, there exists $C(T,\alpha)\in\mathbb{R}^+$ such that for any $0\leq s\leq t\leq T$ we have

$$\mathbb{E}|W_{A,Q}(t)-W_{A,Q}(s)|_H^2\leq C(T,\alpha)|t-s|^{1-\alpha},$$

where $W_{A,Q}(t)=\int_0^t e^{(t-s)A}BdW(s)$ and $Q=BB^*$. As a consequence, $W_{A,Q}$ admits a version which is $1/2-\alpha$ Hölder continuous paths with values in H, for any $\alpha>0$.

The general case is between those two results, depending on Q. In higher dimensions, such results can also be obtained with slightly more general arguments.

Notice that even in Theorem 4.5, one requires $\alpha>0$ for the $L^2(\Omega)$ estimate, whereas in the finite-dimensional (SDE) case one can take $\alpha=0$. However, there is no difference for the regularity of the trajectories.

4.3. **Regularity in space.** Again we state results for $Q = I$ (STWN) and $\mathrm{Tr}(Q) < +\infty$.

Theorem 4.6. *For any $t > 0$, $\mathbb{E}|W_A(t)|_\alpha^2 < +\infty$ if and only if $\alpha < 1/4$.*
If Tr$(Q) < +\infty$ and $\alpha < 1/2$, then $\mathbb{E}|W_{A,Q}(t)|_\alpha^2 < +\infty$.

Again, in general dimension d, and for general Q, results can also be written.

In particular, when $Q = I$, $W_A(t) \notin D(A)$: solutions are not strong (in the PDE sense). It is even worse: $W_A(t) \notin D(-A)^{1/2}$.

Can you give a meaning to $\partial_x(W_A(t)^2)$ and then to solutions of the viscous Burgers equation? Thanks to appropriate Sobolev embeddings and regularity estimates on the semi-group, the answer is yes.

Can you give a meaning to $\big(\partial_x W_A(t)\big)^2$, to be able to consider the KPZ equation? The answer is much more involved and has required a lot of (recent) work...

Proof of Theorem 4.6

$$\begin{aligned}\mathbb{E}|W_{A,Q}(t)|_\alpha^2 &= \mathbb{E}|(-A)^\alpha W_{A,Q}(t)|_H^2\\ &= \int_0^t \mathrm{Tr}((-A)^\alpha e^{(t-s)A} Q e^{(t-s)A}(-A)^\alpha) ds\\ &\leq \int_0^t |e^{2(t-s)A}(-A)^{1-2\epsilon}|_{\mathcal{L}(H)} ds \mathrm{Tr}((-A)^{2\alpha-1+2\epsilon}Q),\end{aligned}$$

for $\epsilon > 0$. We just need to find conditions such that the trace is finite.

When $Q = I$, it is the case for $2\alpha - 1 + 2\epsilon < -1/2$, which is satisfied for a sufficiently small $\epsilon > 0$, when $\alpha < 1/4$.

When $\mathrm{Tr}(Q) < +\infty$, one concludes using the fact that $(-A)^{2\alpha-1+2\epsilon}$ is bounded when $2\alpha-1+2\epsilon \leq 0$, which is satisfied for a sufficiently small $\epsilon > 0$, when $\alpha < 1/2$. $\square$

In fact, $W_A(t)$ takes values not only in $L^2(0,1)$, but also in the space of continuous functions $\mathcal{C}([0,1])$, since the series converges pointwise uniformly in x, at a fixed time t, in the $L^2(\Omega)$ sense, and we write:

$$W_A(t,x) = \sum_{k=1}^{+\infty} \int_0^t e^{-\lambda_k(t-s)} d\beta_k(s) e_k(x). \tag{9}$$

We use the following property of the complete orthonormal system $(e_k)_{k\in\mathbb{N}^*}$ associated with the Laplace operator in $(0,1)$ with Dirichlet boundary conditions:

$$\begin{gathered} e_k \in \mathcal{C}^1([0,1]) \quad \text{for any } k \in \mathbb{N}^*\\ \sup_{k\in\mathbb{N}^*} \sup_{x\in[0,1]} |e_k(x)| < +\infty\\ \sup_{k\in\mathbb{N}^*} \sup_{x\in[0,1]} \frac{|\nabla e_k(x)|}{\sqrt{\lambda_k}} < +\infty.\end{gathered}$$

We can obtain, in the case of space-time white noise:

Theorem 4.7. *For any $\alpha > 0$ and any $T \in \mathbb{R}^+$, there exists $C(T,\alpha) \in \mathbb{R}^+$ such that for any $0 \leq s \leq t \leq T$ we have*

$$\mathbb{E}|W_A(t,x) - W_A(s,y)|^2 \leq C(T,\alpha)\big(|t-s|^{1/2-\alpha} + |x-y|^{1-\alpha}\big).$$

As a consequence, W_A admits a version which is $1/4-\alpha$ Hölder continuous in time and $1/2-\alpha$ Hölder continuous in space, for any $\alpha > 0$.

4.4. **Law of $W_A(t)$.** We want to understand the law of the H-valued random variable $W_A(t)$: as expected by intuition, it has to be Gaussian. Before stating this result, we give the definition and the main properties of Gaussian laws in Hilbert spaces.

Definition 4.8. *Let X be a H-valued random variable, where H is a separable Hilbert space. X is said to be Gaussian if for any $h \in H$, the real random variable $< X, h >_H$ has a Gaussian distribution $\mathcal{N}(\mathbb{E} < X, h >, Var(< X, h >))$.*

Theorem 4.9.
- *There exists $m \in H$ and $Q \in \mathcal{L}(H)$, self-adjoint and nonnegative, such that the following properties hold.*

$$\mathbb{E} < X, h > = < m, h > \qquad \text{for any } h \in H$$
$$Cov(< X, h_1 >< X, h_2 >) = < Qh_1, h_2 > \ \text{ for any } h_1, h_2 \in H.$$

m is called the mean, Q is called the covariance operator.
- *The law of X is caracterized by m and Q; we denote $\mathcal{N}(m, Q)$.*
- *Q is trace-class: $Tr(Q) < +\infty$.*
- *We have $\mathbb{E}[|X|_H^2] = Tr(Q)$: more generally, for any $m \in \mathbb{N}^*$, there exists $C_m \in \mathbb{R}^+$ (not depending on Q) such that $\mathbb{E}[|X|_H^{2m}] \leq C_m\big(Tr(Q)\big)^m$.*
- *There exits $s_Q > 0$ such that for any $0 \leq s < s_Q$ we have $\mathbb{E}\exp(s|x|_H^2) < +\infty$ (Fernique Theorem).*

For any $m \in H$ and any $Q \in \mathcal{L}(H)$ which is trace-class, self-adjoint and nonnegative, there exists a Gaussian random variable X with law $\mathcal{N}(m, Q)$. It can be constructed in the following way (Karhunen-Loève expansion).

Since Q is self-adjoint and compact, there exists a sequence $(q_k)_{k\in\mathbb{N}^*}$ of nonnegative real numbers, and a complete orthonormal system $(f_k)_{k\in\mathbb{N}^*}$ of H, with $Qf_k = q_k f_k$. One sets

$$X = m + \sum_{k\in\mathbb{N}^*} \sqrt{q_k}\gamma_k f_k,$$

where $(\gamma_k)_{k\in\mathbb{N}^*}$ is a iid sequence of standard $\mathcal{N}(0, 1)$ real Gaussian random variables.

When it is well-defined, the stochastic convolution at a given time t is a H-valued Gaussian random variable.

Theorem 4.10. *Let $B \in \mathcal{L}[H)$, and $Q = BB^*$, and $T > 0$.*

Assume that $\int_0^T \|e^{sA}B\|_{\mathcal{L}_2(H)}^2 ds < +\infty$.

Then, for $0 \leq t \leq T$, $W_{A,Q}(t) = \int_0^t e^{(t-s)A}BdW(s)$ is a H-valued Gaussian random variable, with mean 0, and with covariance operator $\int_0^t e^{sA}Qe^{sA}ds$.

Proof If $h = \sum_{n\in\mathbb{N}^*} h_n e_n \in H$, we have by construction of the stochastic integral

$$\begin{aligned}< W_{A,Q}(t), h > &= < \sum_{i,j\in\mathbb{N}^*} \int_0^t < e^{(t-s)A}Be_j, e_i > d\beta_j(s)e_i, \sum_{i\in\mathbb{N}^*} h_i e_i > \\ &= \sum_{i,j\in\mathbb{N}^*} \int_0^t < e^{(t-s)A}Be_j, e_i > d\beta_j(s)h_i \\ &= \sum_{j\in\mathbb{N}^*} \int_0^t < e^{(t-s)A}Be_j, h > d\beta_j(s),\end{aligned}$$

which shows that as a limit of real Gaussian random variables $< W_A(t), h >$ is Gaussian.

It is straightforward that $\mathbb{E} < W_{A,Q}(t), h >= 0$. Now we compute the covariance: for $h_1, h_2 \in H$,

$\mathbb{E} < W_{A,Q}(t), h_1 >< W_{A,Q}(t), h_1 >$

$$
\begin{aligned}
&= \mathbb{E} \sum_{j_1, j_2 \in \mathbb{N}^*} \int_0^t < e^{(t-s)A} B e_{j_1}, h_1 > d\beta_{j_1}(s) \int_0^t < e^{(t-s)A} B e_{j_2}, h_2 > d\beta_{j_2}(s) \\
&= \mathbb{E} \sum_{j \in \mathbb{N}^*} \int_0^t < e^{(t-s)A} B e_j, h_1 > d\beta_j(s) \int_0^t < e^{(t-s)A} B e_j, h_2 > d\beta_j(s) \\
&= \sum_{j \in \mathbb{N}^*} \int_0^t < e^{(t-s)A} B e_j, h_1 >< e^{(t-s)A} B e_j, h_2 > ds \\
&= \int_0^t < B^* e^{sA} h_1, B^* e^{sA} h_2 > \\
&=< (\int_0^t e^{sA} Q e^{sA} ds) h_1, h_2 > .
\end{aligned}
$$

□

Let us finally consider the long-time behavior in the STWN case ($Q = I$).

Proposition 4.11. *When $t \to +\infty$, $W_A(t)$ converges in law to a Gaussian random variable $\mathcal{N}(0, \int_0^{+\infty} e^{2sA} ds = \frac{(-A)^{-1}}{2})$.*

Moreover, if $X \sim \mathcal{N}(0, \frac{(-A)^{-1}}{2})$, we can write $X = \sum_{k \in \mathbb{N}^*} \frac{1}{\sqrt{2\lambda_k}} \gamma_k e_k$; we see that almost surely $X \in \mathcal{C}([0,1])$.

Recall that $e_k(x) = \sqrt{2} \sin(k\pi x)$.

The process $x \in [0,1] \mapsto X(x) = \sum_{k \in \mathbb{N}^*} \frac{1}{\sqrt{\lambda_k}} \gamma_k \sin(k\pi x)$ is Gaussian, with mean 0, and the covariance function: for any $x, y \in [0,1]$

$$
Cov(X(x), X(y)) = (x \wedge y)(1 - x \vee y).
$$

The process (with variable x) X has the law of a Brownian Bridge (defined by $B_x - xB_1$, where B is a Brownian Motion).

Proof The convergence is checked by looking at $< W_A(t), h >$ for all $h \in H$.

Notice that $(-A)^{-1}$ is trace-class, so that the expansion is the general Karhunen-Loeve expansion of Gaussian random variables.

For a fixed $x \in [0,1]$, the series $\sum_{k=1}^{+\infty} \frac{1}{\sqrt{\lambda_k}} \gamma_k \sin(k\pi x)$ converges in probability, and is therefore Gaussian as a limit of gaussian random variables: indeed, by independence of the $(\gamma_k)_{k \in \mathbb{N}^*}$

$$
\begin{aligned}
\mathbb{E}[\big(\sum_{k=N}^{+\infty} \frac{1}{\sqrt{\lambda_k}} \gamma_k \sin(k\pi x)\big)^2] &= \mathbb{E} \sum_{k=N}^{+\infty} \frac{1}{\lambda_k} \gamma_k^2 \sin^2(k\pi x) \\
&\leq \sum_{k=N}^{+\infty} \frac{1}{\lambda_k}.
\end{aligned}
$$

Moreover, with the same computation using independence, and since $\lambda_k = k^2\pi^2$, we have

$$
\begin{aligned}
\mathbb{E}|X(x) - X(y)|^2 &\leq \sum_{k=1}^{+\infty} \frac{1}{\lambda_k} |\sin(k\pi x) - \sin(k\pi y)|^2 \\
&\leq C|x - y|^{1-\alpha} \sum_{k=1}^{+\infty} \frac{1}{k^{1+\alpha}},
\end{aligned}
$$

and an application of the Kolmogorov-Centzov criterion gives $1/2-\alpha$ Hölder continuity, up to a modification.

In particular, $X(0)=X(1)$ almost surely.

It is clear that $\mathbb{E}X(x)=0$ for any $x\in[0,1]$, and it remains to compute the covariance of $X(x)$ and $X(y)$ for $x,y\in[0,1]$: by independence of the standard gaussian random variables $(\gamma_k)_{k\in\mathbb{N}^*}$,

$$\begin{aligned}
\mathrm{Cov}(X(x),X(y)) &= \mathbb{E}X(x)X(y)\\
&= \sum_{k=1}^{+\infty}\frac{1}{k^2\pi^2}\sin(k\pi x)\sin(k\pi y)\\
&= \frac{1}{2}\sum_{k=1}^{+\infty}\frac{1}{k^2\pi^2}\big(\cos(2k\pi(x-y))-\cos(2k\pi(x+y))\big)\\
&= C(x-y)-C(x-y),
\end{aligned}$$

where for any $z\in\mathbb{R}$ we set $C(z)=\frac{1}{2}\sum_{k=1}^{+\infty}\frac{1}{k^2\pi^2}\big(1-\cos(2k\pi z)\big)$.

It is now an easy exercice about Fourier series to prove that C is even, 1-periodic and that $C(z)=z(1-z)$ for $z\in[0,1]$.

In particular, for $x\in[0,1]$, $\mathrm{Var}(X(x))=C(x)=x(1-x)$ (since $C(0)=0$).

To check the formula for the variance, one uses symmetry with respect to $1/2$ to restrict to the case $0\le x\le y\le 1/2$; then simple algebraic computations give the result. $\square$

This result is given in the very simple linear setting, and can be generalized. Using MCMC algorithms, this gives a natural way to sample according to a distribution of a diffusion on $[0,1]$ conditioned to be 0 at the boundary, if one is able to write a SPDE with unique invariant law this distribution.

The general study of long-time behavior, ergodicity and caracterization of the invariant laws of SPDEs, depending on the possible degeneracy of the noise, has a huge literature.

5. Numerical approximation to solutions of SPDEs

Once again, we restrict our attention to the following kind of SPDEs in $H=L^2(0,1)$

$$du(t)=Au(t)dt+F(u(t))dt+dW(t),$$

with the initial condition $u(0)=u_0\in H$, W a cylindrical Wiener process on H, A the Laplace operator in $(0,1)$, with homogeneous Dirichlet boundary conditions, and $F:H\to H$ a Lipschitz (and bounded, to simplify) function; in fact we assume F to be of class $\mathcal{C}^2$, with bounded first and second order derivatives (to get the optimal weak convergence rates below).

We first define time and space discretization numerical schemes, and give their respective orders of convergence, in the so-called strong and weak (probability) senses.

We then study in details a simpler case, where we discretize in space by projecting onto eigenspaces spanned by eigenfunctions of A.

5.1. **Discretization in space: finite element method.** Let $h>0$ be a mesh size, and $(V_h)_{h\in H}$ be a family of (finite dimensional) Finite Element Spaces. As a special case, consider piecewise linear approximation on $[0,1]$ and with the homogeneous Dirichlet boundary conditions. P_h denotes the H-orthogonal projector on V_h. We then define $F_h=P_h\circ F$.

The approximation A_h of A is defined in a variational way: for any $x_h\in V_h$

$$\begin{cases}A_hx_h\in V_h\\ <A_hx_h,y_h>=<Ax_h,y_h>.\end{cases}$$

On the finite dimensional space V_h, A_h is a linear, nonnegative, self-adjoint operator.

The approximated (continuous time) SPDE is:

$$du_h(t) = A_h u_h(t)dt + F_h(u_h(t))dt + P_h dW(t), \quad u_h(0) = P_h u_0.$$

It admits a unique mild solution in V_h. Notice that $(P_h W(t))_{t\geq 0}$ is simply a Wiener process in the finite dimensional space V_h.

The following convergence result holds (in the case of piecewise linear interpolation; for different FEM approximations, the orders might be different):

Theorem 5.1. *For any $T \in \mathbb{R}^+$, and any $\alpha > 0$, there exists $C_\alpha(T) \in \mathbb{R}^+$ such that*

$$\mathbb{E}|u_h(T) - u(T)|_H^2 \leq C_\alpha(T) h^{1-\alpha} \quad \textit{(Strong convergence)}.$$

Moreover, for any $T \in \mathbb{R}^+$, any test function $\varphi \in \mathcal{C}_b^2(H, \mathbb{R})$ (bounded, of class $\mathcal{C}^2$, with bounded first and second order derivatives) and any $\alpha > 0$, there exists $C_\alpha(T, \varphi) \in \mathbb{R}^+$ such that

$$|\mathbb{E}\varphi(u_h(T)) - \mathbb{E}\varphi(u(T))| \leq C_\alpha(T) h^{1-\alpha} \quad \textit{(Weak convergence)}.$$

Notice that the second estimate is not implied by the first one, despite Lipschitz continuity of the test functions. This situation is rather typical for convergence of numerical schemes for stochastic equations.

For SPDEs in higher spatial dimension, with colored noise, similar results can be proved, using the same techniques, and with orders of convergence depending on the precise relations between A, the covariance of the noise, the scheme...

5.2. **Discretization in time: semi-implicit Euler scheme.** Let $\Delta t > 0$ be a time-step size, such that $N = \frac{T}{\Delta t} \in \mathbb{N}^*$.

We start from a method which is well-suited to approximate the deterministic equation (without the noise), check that it is well-defined in the stochastic setting, and then state (without proof) the strong and weak orders of convergence.

To approximate solutions of the equation without the noise, with no discretization in space, we have to use a scheme which is implicit, at least with respect to the linear part. For the nonlinear part, since the coefficient is Lipschitz and bounded, an explicit scheme is reasonable.

If we add the noise term, we thus desire to set:

$$u_{n+1} = u_n + \Delta t A u_{n+1} + \Delta t F(u_n) + W((n+1)\Delta t) - W(n\Delta t),$$

with the initial condition $u_0 = u(0)$.

There is *a priori* no reason for having $u_n \in H$ for all $1 \leq n \leq N$, since $W((n+1)\Delta t) - W(n\Delta t) \notin H$.

Nevertheless, since $S_{\Delta t} = (I - \Delta t A)^{-1}$ is Hilbert-Schmidt for any $\Delta t > 0$, we can define the recursion in H:

$$u_{n+1} = S_{\Delta t} u_n + \Delta t S_{\Delta t} F(u_n) + S_{\Delta t}\big(W((n+1)\Delta t) - W(n\Delta t)\big). \tag{10}$$

This definition gives rise to a discrete mild formula:

$$u_n = S_{\Delta t}^n u_0 + \Delta t \sum_{k=0}^{n-1} S_{\Delta t}^{n-k} F(u_k) + \sqrt{\Delta t} \sum_{k=0}^{n-1} S_{\Delta t}^{n-k} \chi_{k+1},$$

where we have introduced the notation $\chi_{n+1} = \frac{1}{\sqrt{\Delta t}}\big(W((n+1)\Delta t) - W(n\Delta t)\big)$: those random variables are independent and have the law of $W(1)$. They could be called *cylindrical Wiener increments*.

The convergence result is the following. The strong order 1/4 and the weak order 1/2 are expected, in link with the regularity of the continuous-time limit $u(.)$.

Theorem 5.2. *For any $T \in \mathbb{R}^+$, and any $\alpha > 0$, there exists $C_\alpha(T) \in \mathbb{R}^+$ such that*

$$\mathbb{E}|u_N - u(T)|_H^2 \leq C_\alpha(T)\Delta t^{1/2-\alpha} \quad \textit{(Strong convergence)}.$$

Moreover, for any $T \in \mathbb{R}^+$, any test function $\varphi \in \mathcal{C}_b^2(H,\mathbb{R})$ (bounded, of class $\mathcal{C}^2$, with bounded first and second order derivatives) and any $\alpha > 0$, there exists $C_\alpha(T,\varphi) \in \mathbb{R}^+$ such that

$$|\mathbb{E}\varphi(u_N) - \mathbb{E}\varphi(u(T))| \leq C_\alpha(T)\Delta t^{1/2-\alpha} \quad \textit{(Weak convergence)}.$$

5.3. **Full-discretization scheme.** A full time and space discretization scheme is easily defined from the methods described separately above; convergence rates are not surprising, one just combines the results.

5.4. **Galerkin approximation.** This short introduction to SPDEs ends with a simpler approximation, which is a useful tool for the analysis of the equations and for the proof of the convergence estimates for the discretization schemes introduced above. It also has a great advantage: we can also provide error estimates and prove them with essentially the same arguments, though in a much simpler technical setting.

For our choice of A, we know explicitely the diagonalization basis, such that $Ae_k = -\lambda_k e_k$:

$$\begin{cases} e_k(x) = \sqrt{2}\sin(k\pi x), \\ \lambda_k = k^2\pi^2. \end{cases}$$

We thus introduce $H_N = \operatorname{span}\{e_1, \ldots, e_N\}$ for any $N \in \mathbb{N}^*$, and P_N the associated orthogonal projection.

We define $F_N = P_N F$.

The discretization scheme is therefore the following:

$$du^N(t) = Au^N(t)dt + F_N(u^N(t))dt + P_N dW(t), \quad u^N(0) = P_N u_0. \tag{11}$$

The convergence result is:

Theorem 5.3. *For any $T \in \mathbb{R}^+$, and any $\alpha > 0$, there exists $C_\alpha(T) \in \mathbb{R}^+$ such that*

$$\mathbb{E}|u^N(T) - u(T)|_H^2 \leq C_\alpha(T)\frac{1}{\lambda_{N+1}^{1/2-\alpha}} \quad \textit{(Strong convergence)}.$$

Moreover, for any $T \in \mathbb{R}^+$, any test function $\varphi \in \mathcal{C}_b^2(H,\mathbb{R})$ (bounded, of class $\mathcal{C}^2$, with bounded first and second order derivatives) and any $\alpha > 0$, there exists $C_\alpha(T,\varphi) \in \mathbb{R}^+$ such that

$$|\mathbb{E}\varphi(u^N(T)) - \mathbb{E}\varphi(u(T))| \leq C_\alpha(T)\frac{1}{\lambda_{N+1}^{1/2-\alpha}} \quad \textit{(Weak convergence)}.$$

Once again the result is not surprising.

On the one hand, we know that $\mathbb{E}|u(T)|_\beta^2 < +\infty$ if and only if $\beta < 1/4$ (the stochastic part has this regularity, while the remainder is more regular), and the key estimate is the following:

$$\|(I - P_N)(-A)^{-\beta}\|_{\mathcal{L}(H)} \leq C_\beta \lambda_{N+1}^{-\beta}, \tag{12}$$

which is equivalent to the validity of $|(I - P_N)x|_H \leq C_\beta \lambda_{N+1}^{-\beta}|x|_\beta$ for any $x \in D(-A)^\beta$.

On the other hand, we see that the weak order of convergence is again twice the strong order one. In fact, this is the difficult point to be proved, with the adaptation of standard techniques.

Proof of strong convergence in Theorem 5.3 Thanks the mild formulation associated with (11), we decompose the error at time $t \in [0,T]$ into three parts:

$$u^N(t) - u(t) = e^{tA}(I-P_N)u_0 + \int_0^t e^{(t-s)A}(I-P_N)F(u_N(s))ds + \int_0^t e^{(t-s)A}(I-P_N)dW(s).$$

For the first two terms of the right-hand side above, we have the *deterministic* estimate

$$|e^{tA}(I-P_N)u_0|_H \leq \|(I-P_N)(-A)^{-1}\|_{\mathcal{L}(H)}\|e^{tA}(-A)\|_{\mathcal{L}(H)}|u_0|_H \leq \frac{C}{t\lambda_{N+1}}|u_0|_H,$$

with a singularity $1/t$ at time 0 (if we only assume $u_0 \in H$, without better regularity), and the *almost sure* one (for simplicity we have assumed F to be bounded)

$$\begin{aligned}
\Big|\int_0^t e^{(t-s)A}(I-P_N)F(u_N(s))ds\Big| &\leq \int_0^t \big|e^{(t-s)A}(I-P_N)F(u_N(s))\big|ds\\
&\leq \int_0^t \big\|e^{(t-s)A}(-A)^{1-\alpha}\big\|_{\mathcal{L}(H)}\big\|(-A)^{-1+\alpha}(I-P_N)\big\|_{\mathcal{L}(H)}\sup_H |F(.)|_H ds\\
&\leq \frac{C}{\lambda_{N+1}^{1-\alpha}}\int_0^t \frac{1}{(t-s)^{1-\alpha}}ds \sup_H |F(.)|_H,
\end{aligned}$$

and the singularity is integrable.

Now for the noise term, we have the *mean-square* estimate (obtained by Itô isometry)

$$\begin{aligned}
\mathbb{E}\Big|\int_0^t e^{(t-s)A}(I-P_N)dW(s)\Big|_H^2 &= \int_0^t \big\|e^{(t-s)A}(I-P_N)\big\|_{\mathcal{L}_2(H)}^2 ds\\
&\leq \int_0^t \|e^{(t-s)A}(-A)^{1/2-\alpha/4}\|_{\mathcal{L}(H)}^2 ds\big\|(I-P_N)(-A)^{-1/4+\alpha/2}\|_{\mathcal{L}(H)}^2\big\|(-A)^{-1/4-\alpha/4}\|_{\mathcal{L}_2(H)}^2\\
&\leq \frac{C_\alpha(T)}{\lambda_{N+1}^{1/2-\alpha}}.
\end{aligned}$$

□

A few remarks after this proof can be made. First, there can be a singularity depending on the initial condition: it disappears if we assume $u_0 \in D(-A)^{1/4}$, and then the stronger result $\mathbb{E}\sup_{0\leq t\leq T}|u^N(t)-u(t)|_H^2 \leq C_\alpha(T)\frac{1}{\lambda_{N+1}^{1/2-\alpha}}$ also holds. Second, the order of convergence is imposed by the stochastic convolution, while other terms appearing in the mild formulation converge much faster (at a "deterministic order").

Sketch of proof of weak convergence in Theorem 5.3 We do not include all the (very technical) details, and rather insist on the main ideas.

First, we introduce $M > N$ and decompose the weak error as

$$\mathbb{E}\varphi(u^N(T)) - \mathbb{E}\varphi(u(T)) = \mathbb{E}\varphi(u^N(T)) - \mathbb{E}\varphi(u^M(T)) + \mathbb{E}\varphi(u^M(T)) - \mathbb{E}\varphi(u(T)).$$

We prove an estimate of the first term on the right-hand side, uniformly with respect to M:

$$|\mathbb{E}\varphi(u^N(T)) - \mathbb{E}\varphi(u(T))| \leq C_\alpha(T)\frac{1}{\lambda_{N+1}^{1/2-\alpha}},$$

and then let $M \to +\infty$, so that converges to 0.

Having discretized at level $M > N$, we are now in a finite-dimensional situation, which simplifies a lot of analytical arguments.

We introduce the solution $\mathcal{U}^M$ of the backward Kolmogorov equation associated with the M-dimensional diffusion process, and initial condition $\varphi^M := \varphi \circ P_M$: for $t \in (0, T]$, $u \in H_M$

$$\frac{\partial \mathcal{U}^M(t,u)}{\partial t} = \mathcal{L}_M \mathcal{U}^M(t,u), \quad \mathcal{U}^M(0,u) = \varphi^M(u), \tag{13}$$

where $\mathcal{L}_M$ is the generator: if $\Psi \in \mathcal{C}_c^2(H_M, \mathbb{R})$, for any $u \in H_M$

$$\mathcal{L}_M \Psi(u) =< AP_M u + F_M(u), D\Psi(u) > + \frac{1}{2}\mathrm{Tr}(D^2\Psi(u)), \tag{14}$$

where $D\Psi(u) \in \mathcal{L}(H_M, \mathbb{R})$ is identified with the gradient in H_M, and $D^2\Psi(u)$ with a linear operator in $\mathcal{L}(H_M, H_M)$.

Note that $\mathrm{Tr}(D^2\Psi(u)) = \Delta^M \Psi(u)$, $H^M = \mathrm{span}(e_1, \ldots, e_M)$ is identified with $\mathbb{R}^M$ and Δ^M is the usual Laplace operator. The above notation must be preferred if we think that $M \to +\infty$, and also if we had considered colored or multiplicative noise.

The reason for introducing $\mathcal{U}^M$ is the following (classical) calculation, using Itô formula:

$$\begin{aligned}
\mathbb{E}\varphi(u^N(T)) - \mathbb{E}\varphi(u^M(T)) &= \mathbb{E}\mathcal{U}^M(0, u^N(T)) - \mathcal{U}^M(T, P_M u_0) \\
&= \int_0^T \big(-\frac{\partial \mathcal{U}^M}{\partial t}(T-t, u^N(T)) + \mathcal{L}_N \mathcal{U}^M(T-t, u^N(T))\big) dt \\
&= \int_0^T (\mathcal{L}_N - \mathcal{L}_M)\mathcal{U}^M(T-t, u^N(t)) dt.
\end{aligned}$$

To obtain (the appropriate) rates of convergence, we need the following regularization result at the level of the Kolmogorov equation, with uniform bounds with respect to M (using the identifications made above):

Proposition 5.4. *For any $\alpha \in [0,1)$, there exists $C_\alpha \in \mathbb{R}^+$ such that for any $t > 0$ and $u \in H_M$*

$$|D\mathcal{U}^M(t,u)|_\alpha \leq C_\alpha t^{-\alpha}.$$

For any $\alpha, \beta \in [0,1)$, with $\alpha + \beta < 1$, there exists $C_{\alpha,\beta} \in \mathbb{R}^+$ such that for any $t > 0$ and $u \in H_M$

$$\|(-A)^\alpha D\mathcal{U}^M(t,u)(-A)^\beta\|_{\mathcal{L}(H)} \leq C_{\alpha,\beta} t^{-\alpha-\beta}.$$

We postpone a sketch of proof of this result after the end of the proof of Theorem 5.3.

If we now decompose

$$\begin{aligned}
\int_0^T (\mathcal{L}_N - \mathcal{L}_M)\mathcal{U}^M(T-t, u^N(t)) dt &= \int_0^T < (P_N - P_M)Au^N(t), D\mathcal{U}^M(T-t, u^N(t)) > dt \\
&+ \int_0^T < (P_N - P_M)F(u^N(t)), D\mathcal{U}^M(T-t, u^N(t)) > dt \\
&+ \frac{1}{2}\int_0^T \mathrm{Tr}\big((P_N - P_M)D^2\mathcal{U}^M(T-t, u^N(t))\big) dt,
\end{aligned}$$

we see that we are able to treat the last two terms on the right-hand side above, but not the first one (more precisely, we cannot easily obtain weak order 1/2, but only the order 1/4).

First, almost surely,

$$
\begin{aligned}
&\Big|\int_0^T < (P_N-P_M)F(u^N(t)), D\mathcal{U}^M(T-t,u^N(t)) > dt\Big| \\
&\qquad\qquad \leq \int_0^T \big| < (-A)^{-1+\alpha}(P_N-P_M)F(u^N(t)), (-A)^{1-\alpha}D\mathcal{U}^M(T-t,u^N(t)) > \big| dt \\
&\qquad\qquad \leq \int_0^T \|(-A)^{-1+\alpha}(P_N-P_M)\|_{\mathcal{L}(H)} \sup_H |F(.)|_H \frac{C_\alpha}{(T-t)^{1-\alpha}} dt \\
&\qquad\qquad \leq C_{\alpha,T}\lambda_{N+1}^{-1+\alpha}.
\end{aligned}
$$

Second,

$$
\begin{aligned}
&\Big|\int_0^T \mathrm{Tr}\big((P_N-P_M)D^2\mathcal{U}^M(T-t,u^N(t))\big)dt\Big| \\
&\leq \int_0^T \Big|\mathrm{Tr}\big((-A)^{-1/2+\alpha}(P_N-P_M)(-A)^{-1/2-\alpha/2}(-A)^{1/2-\alpha/4}D^2\mathcal{U}^M(T-t,u^N(t))(-A)^{1/2-\alpha/4}\big)\Big| dt \\
&\leq \int_0^T \|(-A)^{-1/2+\alpha}(P_N-P_M)\|_{\mathcal{L}(H)} \mathrm{Tr}\big((-A)^{-1/2-\alpha/2}\big)\|(-A)^{1/2-\alpha/4}D^2\mathcal{U}^M(T-t,u^N(t))(-A)^{1/2-\alpha/4}\|_{\mathcal{L}(H)} dt \\
&\leq C_\alpha \lambda_{N+1}^{-1/2+\alpha}\int_0^T \frac{C}{(T-t)^{1-\alpha/2}}dt.
\end{aligned}
$$

It remains to control

$$
\Big|\mathbb{E}\int_0^T < (P_N-P_M)Au^N(t), D\mathcal{U}^M(T-t,u^N(t)) > dt\Big|;
$$

of course $u^N(t) \in H_N \subset\subset D(A)$, but we want an estimate which is uniform with respect to M. We can then only use $u^N(t) \in D(-A)^{1/4-\alpha}$ for $\alpha > 0$, and arguments used for the other terms fail to give the right expected order of convergence (we only recover the strong order, while we want to improve this order).

There is non trivial argument to be explained, based on a Malliavin integration by parts.

First, we observe that in the mild expression of $u^N(t)$, only the stochastic convolution has low spatial regularity, and that we only have to treat

$$
\Big|\mathbb{E}\int_0^T < (P_N-P_M)A\int_0^t e^{(t-s)A}P_N dW(s), D\mathcal{U}^M(T-t,u^N(t)) > dt\Big|.
$$

The main tool is the following identity

$$
\begin{aligned}
&\mathbb{E}\int_0^T < (P_N-P_M)A\int_0^t e^{(t-s)A}P_N dW(s), D\mathcal{U}^M(T-t,u^N(t)) > dt \\
&\qquad\qquad = \mathbb{E}\int_0^T\int_0^t \mathrm{Tr}\big(e^{(t-s)A}(P_N-P_M)AD^2\mathcal{U}^M(T-t,u^N(t))\mathcal{D}_s u^N(t)\big)dsdt,
\end{aligned}
$$

where $\mathcal{D}_s u^N(t) \in \mathcal{L}(H_M)$ is the Malliavin derivative at time t, such that $\mathcal{D}_s^h u^N(t) := \mathcal{D}_s u^N(t)h$, for $h \in H_M$, is solution of

$$
\frac{d\mathcal{D}_s^h u^N(t)}{dt} = A\mathcal{D}_s^h u^N(t) + P_M DF(u^N(t)).\mathcal{D}_s^h u^N(t) \quad \mathcal{D}_s^h u^N(s) = h.
$$

We easily obtain

$$
\|\mathcal{D}_s u^N(t)\|_{\mathcal{L}(H)} \leq C.
$$

We can now conclude:

$$
\begin{aligned}
&\Big|\mathbb{E}\int_0^T\int_0^t \mathrm{Tr}\big(e^{(t-s)A}(P_N-P_M)AD^2\mathcal{U}^M(T-t,u^N(t))\mathcal{D}_s u^N(t)\big)dsdt\Big|\\
&\qquad\leq \int_0^T\int_0^t \mathbb{E}\big|\mathrm{Tr}\big(e^{(t-s)A}(P_N-P_M)AD^2\mathcal{U}^M(T-t,u^N(t))\mathcal{D}_s u^N(t)\big)\big|dsdt\\
&\qquad\leq \int_0^T\int_0^t \mathrm{Tr}((-A)^{-1/2-\alpha/2})\|(P_N-P_M)(-A)^{-1/2+\alpha}\|_{\mathcal{L}(H)}\|\mathcal{D}_s u^N(t)(-A)^{1-\alpha/4}e^{(t-s)A}\|_{\mathcal{L}(H)}\\
&\qquad\qquad\qquad\qquad\|(-A)^{1-\alpha/4}D^2\mathcal{U}^M(T-t,u^N(t))\|_{\mathcal{L}(H)}dsdt\\
&\qquad\leq C\lambda^{-1/2+\alpha}\int_0^T\int_0^t \frac{1}{(T-t)^{1-\alpha/4}(t-s)^{1-\alpha/4}}dsdt.
\end{aligned}
$$

□

We now give a sketch of proof of Proposition 5.4, which has been a useful tool to deal with the Kolmogorov equation, uniformly with respect to dimension M, and to obtain the weak order of convergence.

Proof of Proposition 5.4 We write, for $\mathrm{u},h\in H_M$, $D\mathcal{U}^M(t,\mathrm{u}).h=\mathbb{E}[D\varphi(u^M(t,\mathrm{u})).\eta^{h,M}(t,\mathrm{u})]$, where

$$
\frac{d\eta^{h,M}(t,\mathrm{u})}{dt}=A\eta^{h,M}(t,\mathrm{u})+P_M DF(u^M(t,\mathrm{u})).\eta^{h,M}(t,\mathrm{u}),\quad \eta^{h,M}(0,\mathrm{u})=h,
$$

and $u^M(0,\mathrm{u})=\mathrm{u}$. The claim follows using the mild formulation for $\eta^{h,M}(t,\mathrm{u})$, and Gronwall Lemma:

$$
|\eta^{h,M}(t,\mathrm{u})|_H\leq Ct^{-\alpha}|h|_{-\alpha}+K\int_0^t|\eta^{h,M}(s,\mathrm{u})|_H ds.
$$

The proof for the second-order derivative is similar:

$$
D^2\mathcal{U}^M(t,\mathrm{u}).(h_1,h_2)=\mathbb{E}[D\varphi(u^M(t,\mathrm{u})).\xi^{h_1,h_2,M}(t,\mathrm{u})]+\mathbb{E}[D^2\varphi(u^M(t,\mathrm{u})).(\eta^{h_1,M}(t,\mathrm{u}),\eta^{h_2,M}(t,\mathrm{u}))],
$$

with

$$
\begin{aligned}
\frac{d\xi^{h_1,h_2,M}(t,\mathrm{u})}{dt}=&A\xi^{h,M}(t,\mathrm{u})+P_M D^2F(u^M(t,\mathrm{u})).(\eta^{h_1,M}(t,\mathrm{u}),\eta^{h_2,M}(t,\mathrm{u}))\\
&+P_M DF(u^M(t,\mathrm{u})).\xi^{h,M}(t,\mathrm{u})\\
\xi^{h,M}(0,\mathrm{u})=&0.
\end{aligned}
$$

We then easily get (mild formulation and Gronwall Lemma) $|\xi^{h_1,h_2,M}(t,\mathrm{u})|_H\leq C|h_1|_{-\alpha}|h_2|_{-\beta}$ if $\alpha+\beta<1$. □

References

[1] C.-E. Bréhier. Strong and weak orders in averaging for SPDEs. *Stochastic Process. Appl.*, 122(7):2553–2593, 2012.

[2] R. Dalang, D. Khoshnevisan, C. Mueller, D. Nualart, and Y. Xiao. *A minicourse on stochastic partial differential equations*, volume 1962 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Held at the University of Utah, Salt Lake City, UT, May 8–19, 2006, Edited by Khoshnevisan and Firas Rassoul-Agha.

[3] R.C. Dalang and L. Quer-Sardanyons. Stochastic integrals for spde's: a comparison. *http://arxiv.org/abs/1001.0856*, 2010.

[4] J. Van Neerven, M. Veraar, and L. Weis. Stochastic integration in banach spaces - a survey. *arxiv*, 2013.

[5] E. Pardoux. Stochastic partial differential equations. *Lectures given in Fudan University, Shangai*, 2007.

[6] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44. Cambridge University Press, In Encyclopedia of Mathematics and Its Applications, 1992.

[7] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[8] J. B. Walsh. *An introduction to stochastic partial differential equations*, volume Lect. Notes Math. 1180, 265-437. Ecole d'Eté de Probabilités de Saint-Flour XIV - 1984, 1986.

Université Paris-Est, CERMICS (ENPC), 6-8-10 Avenue Blaise Pascal, Cité Descartes , F-77455 Marne-la-Vallée, France

INRIA Paris-Rocquencourt-Matherials, Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay, France

E-mail address: brehierc@cermics.enpc.fr,charles-edouard.brehier@inria.fr