Stochastic Partial Differential Equations

Wilhelm Stannat

Institut für Mathematik
Technische Universität Berlin

stannat@math.tu-berlin.de

October 15, 2020

Lecture held in the winter term 2020/21 at TU Berlin.
Please report typos and mistakes to: stannat@math.tu-berlin.de

Contents

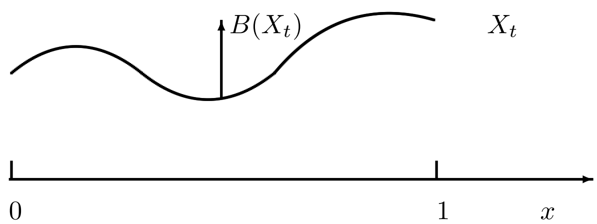
Kapitel	**Anzahl LE**
1. Introduction	1
2. Stochastic Analysis on Hilbert spaces	4
3. Stochastic differential equations on Hilbert spaces	2
4. Variational approach to SPDE	6
5. Stochastic reaction-diffusion systems	4
	$\sum = 16$

Lecture 1

Introduction: Stochastic partial differential equations

Some motivating examples:

Example 1: random motion of strings (Funaki, 1983)



where

- $X_t : [0,1] \to \mathbb{R}$ is a (random) function on $[0,1]$ ('string') and $X_t(x)$ denotes its value at position $x \in [0,1]$ (random variable (!))
- $B : \mathbb{R} \to \mathbb{R}$ models forcing term, $B(X_t(x))$ models direction of some exterior force acting on the string at position x
- time evolution of the moving string:

$$\partial_t X_t = \nu \underbrace{\partial^2_{xx} X_t}_{\text{diffusion}} + B(X_t) + \sigma \underbrace{\partial_t W_t}_{\text{noise}}$$

subject to boundary conditions:

- Neumann bdy cond: $\partial_x X_t(0) = \partial_x X_t(1) = 0$
- Dirichlet bdy cond: $X_t(0) = X_t(1) = 0$

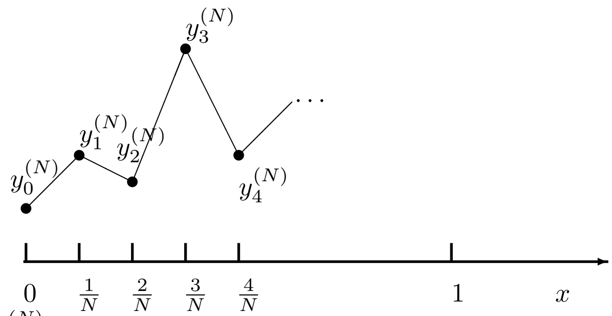
Here, W_t denotes a random function such that for all $x \in [0,1]$

$$t \mapsto W_t(x)$$

is a standard Brownian motion and for $x \neq y$ the Brownian motions $W_.(x)$ and $W_.(y)$ should be for simplicity independent. ('space-time white noise')

What is a rigorous mathematical description of this?

spatial discretization given $N \in \mathbb{N}$



The nodes $y_k^{(N)}(t)$, $k = 0, 1, \ldots, N$, are solutions of the following system of stochastic differential equations:

$$dy_k^{(N)}(t) = \nu \underbrace{\frac{1}{(\Delta x)^2}\left(y_{k+1}^{(N)}(t) - 2y_k^{(N)}(t) + y_{k-1}^{(N)}(t)\right)}_{\text{2nd order finite diff.}} dt$$
$$+ B(y_k^{(N)}(t))\, dt + \frac{1}{\sqrt{\Delta x}} dW_k(t),\ k = 0, 1, \ldots, N.$$

with $\Delta x = \frac{1}{N}$ ('spatial distance').

Why spatial scaling $\frac{1}{\sqrt{\Delta x}}$ in front of the Brownian motions ?

associated polygon:

$$X^{(N)}(t,x) := \frac{x_{k+1}^{(N)} - x}{x_{k+1}^{(N)} - x_k^{(N)}} y_k^{(N)}(t) + \frac{x - x_k^{(N)}}{x_{k+1}^{(N)} - x_k^{(N)}} y_{k+1}^{(N)}(t) \text{ if } x \in [x_k^{(N)}, x_{k+1}^{(N)}]$$
$$\in C([0,1]) \subset L^2([0,1])$$

with $x_k^{(N)} := \frac{k}{N}$.

Functional central limit theorem

Let $P^{(N)} := P \circ (X^{(N)})^{-1}$ on $C([0,1]; L^2([0,1]))$. Then $P^{(N)} \Rightarrow P^{(\infty)}$ weakly on $C([0,1]; L^2([0,1]))$ where the canonical process under $P^{(\infty)}$ solves the stochastic partial differential equation (spde)

$$dX_t = \left(\nu \partial_{xx}^2 X_t + B(X_t)\right) dt + dW_t \quad \in L^2([0,1]) \tag{1.1}$$

where W_t denotes the 'Brownian motion' on $L^2([0,1])$.

Formally: $\forall\ s,t \in [0,\infty)$, $\forall\ x,y \in [0,1]$

$$\mathbb{E}\left(W_s(x)W_t(y)\right) = s \wedge t\ \delta_0(x-y)$$

Alternative notation for the spde (1.1): Using $u(t,x) := X_t(x)$ and $W(t,x) = W_t(x)$, we can formally write

$$\partial_t u(t,x) = \nu \partial_{xx}^2 u(t,x) + B(u(t,x)) + \partial_t W(t,x) \quad \in L^2([0,1]) \tag{1.2}$$

Of course, the time-derivative of the Wiener process does not exist ion the classical sense and (1.2) has to be understood as usual in the sense of the associated integral equation.

Example 2: stochastic reaction-diffusion equations

A large class of widely used stochastic models for sptially distributeed random evolutions are given by stochastic reaction-diiffusion equations, e.g.

$$\begin{cases} \partial_t v(t,x) &= \nu \partial_{xx}^2 v(t,x) + bf(v(t,x)) + g(v(t,x))\partial_t W(t,x) \\ v(0,x) &= v_0(x) \end{cases} \tag{1.3}$$

where

- $x \in D \subset \mathbb{R}^d$
- $\lim_{|v|\to\infty} f(v) \cdot v = -\infty$, i.e., 'coercive'
- $(f(u) - f(v))(u - v) \leq |u - v|^2$, i.e., f one-sided Lipschitz
  - $f(v) = v(1-v)(v-a)$ - Nagumo
  - $f(v) = v(1-v)$ - Fisher resp. KPP (Kolm.-Petrovsky-Piskounov)
  - $f(v,w) = [v(1-v)(v-a) - w, \epsilon(v - \gamma w)]$ - FitzHugh-Nagumo
  - $f(v,w) = [a(b+1)v + v^2 w, w = bvv^2 w]$ - Brusselator

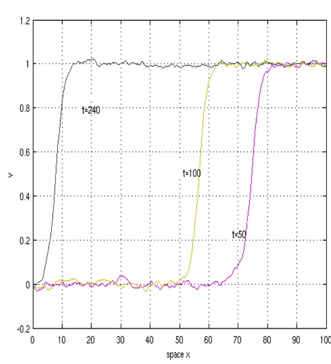
Motivation mathematical models for spatially extended random evolutions used in physical chemistry, material sciences, optics, biophysics, populatation biology

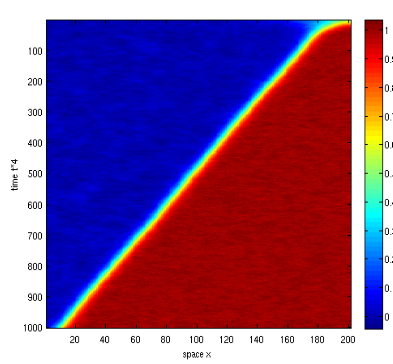
Dynamical features of (1.3) **of interest in the applications**

- bifurcations, i.e., qualitative different dynamical behaviour depending on the chosen set of parameters (e.g. excitable, oscillatory, ...)
- metastable and in general transient macroscopic structures (pulsees, wave fronts, solitons, ...)
- stochastic resonance, i.e., microscopic stochastic fluctuations can accumulate leading to change points on macroscopic scales (e.g. in climate models)
- superposition and interaction of physical/dynamical processes on various time scales and various spatial scales

1D-stochastic Nagumo equation

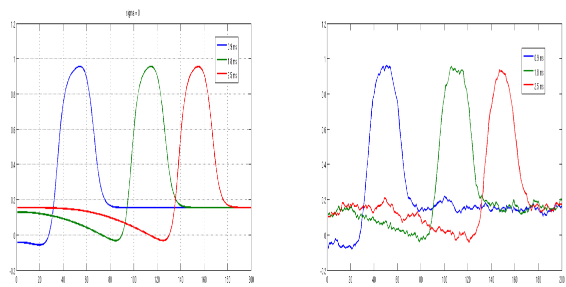
exhibits travelling waves moving with random speed

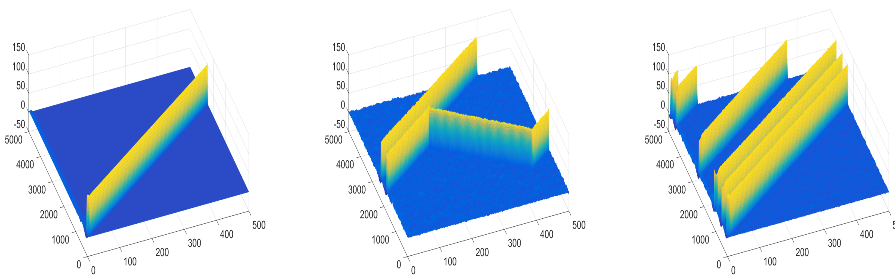




1D-stochastic FitzHugh-Nagumo system

exhibits travelling action potentials (AP's) moving with random speed. Other features include random spiking, multiple spikes, annihilation of spikes, ...





2D-Stochastic FitzHugh-Nagumo systems

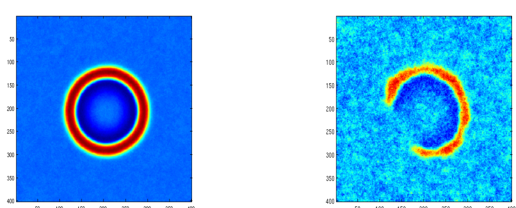
model for intracellular signalling pathways

$$\partial_t v(t,x,y) = \partial^2_{xx} v(t,x,y) + \partial^2_{yy} v(t,x,y) + (v(1-v)(v-a))(t,x,y) - w(t,x,y) + I(t,x,y) + \sigma \partial_t W(t,x,y)$$

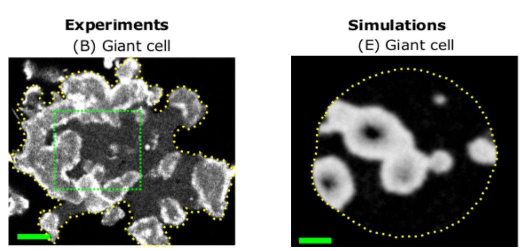
$$\partial_t w(t,x,y) = \epsilon(v - \gamma w)(t,x,y)$$

on the domain $[0,L] \times [0,L] \subset \mathbb{R}^2$

exhibits various spatial patterns



comparison with real experiments



Wave dynamics in the actin cytoskeleton (in *Dictyostelium discoideum*, see [5])

...

Lecture 2

Stochastic Analysis on Hilbert Spaces

2.1. Gaussian measures on Hilbert spaces

Preface We will use some basic facts of linear functional analysis, mostly concerning Hilbert spaces and linear operators on Hilbert spaces. These facts can be found in any introductory text book on functional analysis and essentially donot require any further understanding beyond Euclidean vector spaces, that are in fact finite-dimensional Hilbert spaces, and linear mappings on Euclidean vector spaces. The only essential difference is that in infinite dimensions linearity and continuity of linear mappings are no longer equivalent. In fact, linear mappings $L : U \to V$ between Hilbert spaces U and V (or any normed vector spaces) are continuous if and only if they are bounded, i.e. $L(B_1(0))$ is contained in a bounded subset or equivalently,

$$\sup_{\|h\|_U \le 1} \|Lh\|_V < \infty .$$

Definition 2.1. $(H, \langle \cdot, \cdot \rangle_H)$ is called a **(real) Hilbert space**, if:

(i) H is a vector space over $\mathbb{R}$,

(ii) $\langle \cdot, \cdot \rangle_H : H \times H \longrightarrow \mathbb{R}$ is a scalar product, i.e., a symmetric positive definite bilinear form. In particular, $|u|_H := \sqrt{\langle u, u \rangle_H}$ defines a norm on H.

(iii) $(H, |\cdot|_H)$ is a complete normed (real) vector space, hence a (real) Banach space.

Example 2.2. (i) $\mathbb{R}^N$, equipped with usual scalar product $\langle u, v \rangle := \sum_{i=1}^N u_i v_i$, but also with $\langle u, v \rangle_Q := \sum_{i,j=1}^N q_{ij} u_i v_j$, where $Q = (q_{ij})$ is any symmetric postivie definite $N \times N$-matrix.

(ii) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, then $L^2(\Omega, \mathcal{A}, \mu) =: L^2(\mu)$ with scalar product $\langle f, g \rangle := \int_\Omega fg d\mu$ is a real Hilbert space, but also with the scalar product

$$\langle f, g \rangle_Q := \int_\Omega \int_\Omega f(x) g(y) q(x,y) \mu(dx) \mu(dy)$$

where q is symmetric and positive definite in the sense that $\forall\, A_1, \dots, A_N \in \mathcal{A}$ the matrix $Q = (q_{ij})$, with elements $q_{ij} = \int_{A_i} \int_{A_j} q(x,y)\, \mu(dx)\mu(dy) \ge 0$, $1 \le i, j \le N$, is symmetric and positive definite.

(iii) The sequence space $l^2 = \left\{ (u_k)_{k \ge 1} \subset \mathbb{R} : \sum_{k=1}^\infty u_k^2 < \infty \right\}$ equipped with

$$\langle u, v \rangle := \sum_{k=1}^\infty u_k v_k .$$

Definition 2.3. $(H, \langle \cdot, \cdot \rangle_H)$ is called **separable**, if there exists a countable dense subset $H_0 \subset H$.

EXAMPLE 2.4. $L^2([0,1],dx)$ is separable, with countable dense subsets

$$H_0 = \left\{ \sum_{k=0}^{n} a_k x^k : n \geq 1, a_k \in \mathbb{Q} \right\} \quad = \text{all polynomials with rational coefficients}$$

or

$$H_0 = \left\{ \sum_{k=1}^{n} a_k \sin(2\pi kx) + \sum_{k=0}^{n} b_k \cos(2\pi kx) : n \geq 1, a_k, b_k \in \mathbb{Q} \right\} \quad = \text{all Fourier polynomials with rational coefficients}$$

In general: If Ω is a separable metric space and $\mathcal{A} = \mathcal{B}(\Omega)$ the Borel σ-algebra, then $L^2(\mu)$ is separable.

DEFINITION 2.5. A probability measure μ on $(H, \mathcal{B}(H))$ is called a **Gaussian measure**, if for all $h \in H$ the linear functional

$$l_h : g \mapsto \langle g, h \rangle_H, \quad H \longrightarrow \mathbb{R}$$

is normally distributed, i.e., there exist $m(h) \in \mathbb{R}$, $\sigma^2(h) \in \mathbb{R}_+$, such that for all $t \in \mathbb{R}$:

$$\int_H e^{itl_h} d\mu = e^{itm(h) - \frac{1}{2}t^2\sigma^2(h)}.$$

THEOREM 2.6. *A probability measure μ on $(H, \mathcal{B}(H))$ is a Gaussian measure if and only if*

$$\int_H e^{i\langle g,h \rangle_H} \mu(dg) = e^{i\langle m,h\rangle_H - \frac{1}{2}\langle Qh,h\rangle_H}$$

for all $h \in H$, where

a) *$m \in H$ is the mean vector,*
b) *$Q : H \longrightarrow H$ is linear, continuous and symmetric (i.e. $\langle Qg, h\rangle_H = \langle g, Qh\rangle_H$), positive semidefinite (i.e. $\langle Qg, g\rangle_H \geq 0$ $\forall g \in H$) and of finite trace, i.e.*

$$tr(Q) = \sum_{k=1}^{\infty} \langle Qe_k, e_k\rangle_H < \infty$$

for one (and hence for all) complete orthonormal system (CONS) (e_k) on H.

In this case we will write

$$\mu = \mathcal{N}(m, Q)$$

and Q is called the covariance operator of μ.

In particular:

(i) $\int_H \langle x, h\rangle_H \mu(dx) = \mathbb{E}_\mu(l_h) = m(h) = \langle m, h\rangle_H$

(ii)

$$\begin{aligned} \int_H (\langle x, h_1\rangle - \langle m, h_1\rangle)(\langle x, h_2\rangle - \langle m, h_2\rangle)\mu(dx) &= \mathbb{E}_\mu\left((l_{h_1} - \mathbb{E}_\mu l_{h_1})(l_{h_2} - \mathbb{E}_\mu l_{h_2})\right) \\ &= Cov_\mu(l_1, l_2) \\ &= \langle Qh_1, h_2\rangle_H \quad \forall h_1, h_2 \in H \end{aligned}$$

(iii) $\int_H \|x - m\|_H^2 \mu(dx) = tr(Q)$

EXAMPLE 2.7. (i) **Wiener-measure**

Let $(\beta(t))_{t\geq 0}$ be a 1-dimensional continuous **Brownian motion** on $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. a stochastic process satisfying

a) $\beta(0) = 0$ $\mathbb{P}$-a.s.,
b) for $0 \leq t_1 < t_2 < ... < t_n$ the increments

$$\beta(t_{i+1}) - \beta(t_i), \qquad 1 \leq i \leq n-1$$

are independent and $\mathcal{N}(0, t_{i+1} - t_i)$ distributed.
c) $t \mapsto \beta(t)$ is $\mathbb{P}$-a.s. continuous.

Let $T > 0$ be fixed and let

$$\beta. : \Omega \longrightarrow C([0,T]) \subseteq L^2([0,T]),\ \omega \mapsto (t \mapsto \beta(t,\omega))$$

be measurable w.r.t. the Borel σ-algebra on $L^2([0,T])$. Then the distribution

$$\mu(A) := \mathbb{P}(\beta(\cdot) \in A), \qquad A \in \mathcal{B}(L^2([0,T]))$$

is called the **Wiener measure**.

μ is a Gaussian measure (centered, i.e., the mean vector is zero) with covariance operator Q induced by the bilinear form

$$\begin{aligned}\langle Qg, h\rangle_{L^2([0,T])} &= \int_0^T Qg(t)h(t)\,dt \\ &= \int_0^T \int_0^T g(s)h(t)\, s \wedge t\, dsdt \\ &= \langle (-\Delta)^{-1} g, h\rangle_{L^2([0,T])}\end{aligned}$$

where Δ denotes the Laplace-Operator on $L^2([0,T])$ (i.e. $\Delta g = g''$), with Dirichlet boundary conditions in 0 ($g(0) = 0$) and Neumann boundary conditions in T ($g'(T) = 0$). With these boundary conditions, Δ becomes invertible and we have indeed that $Q = (-\Delta)^{-1}$.

PROOF.

$$\begin{aligned}\mathbb{E}_\mu(l_h) &= \int_\Omega l_h(\beta)d\mathbb{P} = \int_\Omega \langle h, \beta\rangle_{L^2([0,T])} d\mathbb{P} \\ &= \int_\Omega \int_0^T h(s)\beta(s)dsd\mathbb{P} = \int_0^T h(s) \int_\Omega \beta(s)d\mathbb{P}ds \\ &= \int_0^T h(s) \underbrace{\mathbb{E}_\mu(\beta(s))}_{=0} ds = 0, \text{ hence } m = 0.\end{aligned}$$

$$\operatorname{Cov}_\mu(l_g,l_h) = \underbrace{\int\limits_{L^2([0,T])} \langle x,g\rangle\langle x,h\rangle\mu(dx)}_{\text{since } m(g)=m(h)=0} = \mathbb{E}\left(\int\limits_0^T g(s)\beta(s)ds \int\limits_0^T h(s)\beta(s)ds\right)$$

$$= \int\limits_0^T\int\limits_0^T g(s)h(t)\underbrace{\mathbb{E}(\beta(s)\beta(t))}_{=s\wedge t} dsdt.$$

□

Remark It is instructive to identify Q from the bilinear form. Indeed,

$$\langle Qg,h\rangle_{L^2([0,T])} = \int\limits_0^T\int\limits_0^T g(s)h(t)\, s\wedge t\, dsdt = \int\limits_0^T\left(\int\limits_0^T g(s)\, s\wedge t\, ds\right)h(t)\, dt$$

implies that

$$Qg(t) = \int_0^T g(s)s\wedge t\, ds = \int_0^t g(s)s\, ds + t\int_t^T g(s)\, ds.$$

In particular, $Qg(0)=0$, $(Qg)'(t) = \int_t^T g(s)\, ds$, which is equal to 0 if $t=T$, and $(Qg)''(t) = -g(t)$.

(ii) **Brownian bridge**

The process

$$\beta^0(t) := \beta(t) - \frac{t}{T}\beta(T), \quad 0\le t\le T$$

satisfies $\beta^0(T)=0$ and is therefore called a **Brownian bridge** from 0 to 0 (in time T). The distribution

$$\mu^0(A) := \mathbb{P}(\beta^0(\cdot)\in A), \quad A\in\mathcal{B}(L^2([0,T]))$$

is called a **pinned Wiener measure**.

μ^0 is a Gaussian measure with mean vector 0 and covariance operator

$$\operatorname{Cov}_{\mu^0}(l_g,l_h) = \langle Q^0g,h\rangle = \int\limits_0^T\int\limits_0^T g(s)h(t)\left(\ s\wedge t\ -\frac{st}{T}\right)dsdt$$
$$= \langle(-\Delta_D)^{-1}g,h\rangle$$

where Δ_D now denotes the Dirichlet Laplacian on $L^2([0,T])$. In particular, $Q^0 = (-\Delta_D)^{-1}$.

Proof.

$$\operatorname{Cov}_{\mu^0}(l_g,l_h) = \mathbb{E}_{\mu^0}(l_g,l_h) = \int\limits_0^T\int\limits_0^T g(s)h(t)\mathbb{E}(\beta^0(s)\beta^0(t))dsdt$$

with

$$\begin{aligned}
\mathbb{E}(\beta^0(s)\beta^0(t)) &= \mathbb{E}\Big((\beta(s) - \frac{s}{T}\beta(T))(\beta(t) - \frac{t}{T}\beta(T))\Big) \\
&= \mathbb{E}(\beta(s)\beta(t)) - \frac{t}{T}\mathbb{E}(\beta(s)\beta(T)) - \frac{s}{T}\mathbb{E}(\beta(T)\beta(t)) + \frac{st}{T^2}\mathbb{E}(\beta^2(T)) \\
&= s \wedge t \ - \frac{st}{T} - \frac{st}{T} + \frac{st}{T} \\
&= s \wedge t \ - \frac{st}{T}.
\end{aligned}$$

□

In this case

$$Q^0 g(t) = \int_0^T g(s)\left(s \wedge t - \frac{st}{T}\right) ds = \int_0^t g(s) s\, ds + t\int_t^T g(s)\, ds - \frac{t}{T}\int_0^T g(s) s\, ds$$

indeed satisfies $Q^0 g(0) = Q^0 g(T) = 0$ and $(Qg)'' = -g$.

Sketch of proof of theorem 2.6.
$\Leftarrow$: obvious

$\Rightarrow$: First we have to show that $h \mapsto m(h) = \mathbb{E}_\mu l_h$ is linear and continuous.

linear:

$$\begin{aligned}
m(h+g) = \mathbb{E}_\mu(l_{h+g}) &\overset{\text{scalar prod.}}{\underset{\text{linear}}{=}} \mathbb{E}_\mu(l_h + l_g) \\
&\overset{\text{exp.}}{\underset{\text{linear}}{=}} \mathbb{E}_\mu(l_h) + \mathbb{E}_\mu(l_g) \\
&= m(h) + m(g)
\end{aligned}$$

continuous:
$m(h)$ is continuous (on H) if and only if $\sup_{\|h\|_H \le 1} m(h) < \infty$.

First note that:

$$H_n := \left\{ h \in H : \int |l_h(x)|\mu(dx) \le n \right\} \subseteq H$$

for all $n \ge 1$ according to Fatou's lemma. Moreover

$$H = \bigcup_{n \ge 1} H_n.$$

Therefore Baire's category theorem implies that there exists some $n_0 \ge 1$ such that

$$\overset{\circ}{H}_{n_0} \ne \emptyset.$$

Hence we can find $\varepsilon > 0$ and some element $h_0 \in H$ such that $B_\varepsilon(h_0) \subseteq H_{n_0}$. But this implies that for all $h \in H$ with $\|h\|_H \le 1$:

$$h = \frac{2}{\varepsilon}\left(\frac{\varepsilon}{2}h\right) = \frac{2}{\varepsilon}\Big(\underbrace{\frac{\varepsilon}{2}h + h_0}_{\in B_\varepsilon(h_0)} - h_0\Big).$$

Finally this implies that

$$\begin{aligned}|m(h)| &= \left|\int\limits_H \langle x, h\rangle_H \mu(dx)\right| \\ &\leq \frac{2}{\varepsilon}\Big(\underbrace{\int\limits_H |l_{\frac{\varepsilon}{2}h+h_0}|\mu}_{\leq n_0} + \underbrace{\int\limits_H |l_{h_0}| d\mu}_{\leq n_0}\Big) \\ &\leq \frac{4}{\varepsilon} n_0 < \infty.\end{aligned}$$

Therefore $\sup\limits_{\|h\|_H \leq 1} m(h) \leq \frac{4}{\varepsilon} n_0 < \infty$ which implies that $m(h)$ is continuous. The Riesz representation theorem now implies that there exists $m \in H$ such that

$$\langle m, h\rangle_H = m(h).$$

Similarly one can show that Q is bilinear and continuous on $H \times H$. In particular, (again using Riesz representation theorem):

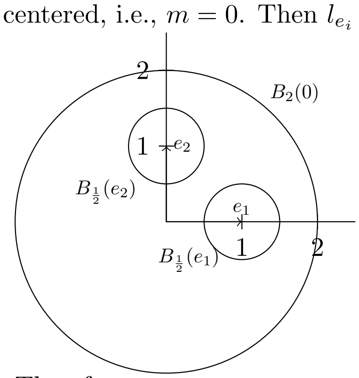
$$\forall h \in H\ \exists Q(h) : \langle Q(h), g\rangle = Q(h, g) \quad \forall g \in H$$

and therefore $h \mapsto Q(h)$ is linear and continuous.

Q ist positive definite, since

$$\langle Qh, h\rangle_H = \int\limits_H (l_h(x) - \langle x, m\rangle_H)^2 \mu(dx) \geq 0.$$

It remains to show that $\mathrm{tr}(Q) < \infty$. To this end let $(e_n)_n$ be a CONS and let μ be centered, i.e., $m = 0$. Then $l_{e_i} \sim \mathcal{N}(0, \langle Qe_i, e_i\rangle)$ and for given e_i we have that:



Therefore

$$\bigcup_{i\geq 1} B_{\frac{1}{2}}(e_i) \subseteq B_2(0).$$

Consequently, □

$$\infty > \mu(B_2(0)) \geq \sum_{i=1}^{\infty} \mu(B_{\frac{1}{2}}(e_i)) \sim \sum_{i=1}^{\infty} \langle Qe_i, e_i\rangle.$$

We will use the following functional analytic fact on trace-class operators without further proof:

PROPOSITION 2.8. *Let Q be linear, continuous, symmetric, positive semidefinite and let $tr(Q) < \infty$ (i.e. $tr(Q) = \sum_{i=1}^{\infty} \langle Qe_i, e_i \rangle < \infty$ where (e_i) is any CONS). Then there exists a CONS e_k, $k \geq 1$, of H, consisting of eigenvectors of Q, i.e.,*

$$Qe_k = \lambda_k e_k, \quad \lambda_k \geq 0.$$

0 is the only accumulation point of the sequence of eigenvalues $(\lambda_k)_{k\geq 1}$. Moreover, $tr(Q) = \sum\limits_{i=1}^{\infty} \lambda_k < \infty$, i.e. (λ_k) converges absolutely.

The following proposition provides the canonical representation of Gaussian measures on Hilbert spaces and also gives an instruction of how to simulate Hilbert space valued Gaussian random variables, resp. to draw samples from a Gaussian measure on some Hilbert space.

PROPOSITION 2.9 (canonical representation). *Let $m \in H$ and let Q and $(e_k)_{k\geq 1}$ be as in the previous Proposition 2.8. Let X be an H-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Then X is normally distributed with $X \sim \mathcal{N}(m, Q)$ if and only if*

$$X = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \beta_k e_k + m,$$

where $(\beta_k)_{k\geq 1}$ are independent, $\mathcal{N}(0,1)$ distributed random variables. The infinite series converges $\mathbb{P}$-a.s. and in L^p for all $p < \infty$.

PROOF. $\Leftarrow$**:** We have to show that $\sum\limits_{k=1}^{\infty} \sqrt{\lambda_k} \beta_k e_k + m$ converges on H. To this end let $m \leq n \in \mathbb{N}$. Then:

$$\begin{aligned}
\|\sum_{k=1}^{n} \sqrt{\lambda_k} \beta_k - \sum_{k=1}^{m} \sqrt{\lambda_k} \beta_k \|_H^2 &= \| \sum_{k=m+1}^{n} \sqrt{\lambda_k} \beta_k \|_H^2 \\
&= \langle \sum_{k=m+1}^{n} \sqrt{\lambda_k} \beta_k, \sum_{k=m+1}^{n} \sqrt{\lambda_k} \beta_k \rangle_H \\
&= \sum_{k,l=m+1}^{n} \sqrt{\lambda_k} \sqrt{\lambda_l} \beta_k \beta_l \underbrace{\langle e_k, e_l \rangle_H}_{=\delta_{kl}} \\
&= \sum_{k=m+1}^{n} \lambda_k \beta_k^2.
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbb{E}\left(\|\sum_{k=1}^{n} \sqrt{\lambda_k} \beta_k - \sum_{k=1}^{m} \sqrt{\lambda_k} \beta_k \|_H^2 \right) &= \mathbb{E}\left(\sum_{k=m+1}^{n} \lambda_k \beta_k^2 \right) \\
&= \sum_{k=m+1}^{n} \lambda_k \underbrace{\mathbb{E}\beta_k^2}_{=1} = \sum_{k=m+1}^{n} \lambda_k \overset{n,m\to\infty}{\longrightarrow} 0
\end{aligned}$$

Finally we have to verify that the infinite series has the correct distribution:

$$
\begin{aligned}
\mathbb{E}e^{i\langle h,X\rangle_H} &= \mathbb{E}e^{i\langle m,h\rangle_H + i\sum_{k=1}^{\infty}\sqrt{\lambda_k}\beta_k\langle e_k,h\rangle_H} \\
&= e^{i\langle m,h\rangle_H}\prod_{k=1}^{\infty}\underbrace{\mathbb{E}e^{i\sqrt{\lambda_l}\langle e_k,h\rangle_H\beta_k}}_{e^{-\frac{\lambda_k}{2}\langle e_k,h\rangle_H^2}} \\
&= e^{i\langle m,h\rangle_H}e^{-\frac{1}{2}\sum_{k=1}^{\infty}\lambda_k\langle e_k,h\rangle_H^2} \\
&= e^{i\langle m,h\rangle_H}e^{-\frac{1}{2}\sum_{k=1}^{\infty}\lambda_k\langle e_k,h\rangle_H\langle e_k,h\rangle_H} \\
&= e^{i\langle m,h\rangle_H}e^{-\frac{1}{2}\sum_{k=1}^{\infty}\langle\lambda_k e_k,h\rangle_H\langle e_k,h\rangle_H} \\
&= e^{i\langle m,h\rangle_H}e^{-\frac{1}{2}\sum_{k=1}^{\infty}\langle Qe_k,h\rangle_H\langle e_k,h\rangle_H} \\
&= e^{i\langle m,h\rangle_H}e^{-\frac{1}{2}\sum_{k=1}^{\infty}\langle e_k,Qh\rangle_H\langle e_k,h\rangle_H} \\
&\overset{\text{Parsevals'}}{\underset{\text{identity}}{=}} e^{i\langle m,h\rangle_H}e^{-\frac{1}{2}\langle Qh,h\rangle_H}
\end{aligned}
$$

□

Lecture 3

Stochastic Analysis on Hilbert Spaces

2.2. Wiener Processes

Let Q be as in the previous section and let $(U, \langle\cdot,\cdot\rangle_U)$ be any separable, real Hilbert space.

DEFINITION 2.10. A U-valued stochastic process $W(t)$, $t \in [0,T]$, on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **(standard) Q-Wiener process**, if

(i) $W(0) = 0$,
(ii) $t \mapsto W(t)$ is $\mathbb{P}$-a.s. continuous,
(iii) for $0 = t_0 < t_1 < \ldots < t_n \leq T$ the increments

$$W(t_i) - W(t_{i-1}), \quad i = 1, \ldots, n$$

are independent, $\mathcal{N}(0, (t_i - t_{i-1})Q)$-distributed.

PROPOSITION 2.11 (canonical representation of Q-Wiener processes). *Let $(e_k)_{k\geq 1}$ be a CONS, consisting of eigenfunctions of Q. Then $W(t)$, $t \in [0,T]$, is a Q-Wiener process, if and only if*

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k, \quad t \in [0,T], \tag{2.4}$$

where $\{\beta_k(t) : \lambda_k > 0\}$ are independent, 1-dimensional (continuous) Brownian motions. The sequence (2.4) *converges $\mathbb{P}$-a.s. on $C([0,T];U)$ and in $L^p(\Omega, \mathcal{F}, \mathbb{P}; C([0,T];U))$ for all $p < \infty$, i.e.*

$$\lim_{n\to\infty} \mathbb{E}\bigg(\sup_{t\in[0,T]} \|\sum_{k=1}^{n} \sqrt{\lambda_k}\beta_k(t)e_k - W(t)\|_U^p \bigg) = 0.$$

SKETCH OF PROOF:
We only present the proof for the case $p = 2$. The general case and the $\mathbb{P}$-a.s. convergence follows similarly from generalizations of the martingale convergence theorems to Hilbert-space valued martingales. To simplify notations let

$$W^n(t) := \sum_{k=1}^{n} \sqrt{\lambda_k}\beta_k(t)e_k.$$

Then

$$\|W^n(t) - W(t)\|_U^2 = \|\sum_{k=n+1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k\|_U^2 = \sum_{k=n+1}^{\infty} \lambda_k\beta_k^2(t).$$

Hence

$$\begin{aligned}\mathbb{E}\left(\sup_{t\in[0,T]}\|W^n(t)-W(t)\|_U^2\right) &= \mathbb{E}\left(\sup_{t\in[0,T]}\sum_{k=n+1}^{\infty}\lambda_k\beta_k^2(t)\right)\\ &\le \sum_{k=n+1}^{\infty}\lambda_k\,\underbrace{\mathbb{E}\left(\sup_{t\in[0,T]}\beta_k^2(t)\right)}_{\le C_T(\text{ Doob})}\\ &\le C_T\sum_{k=n+1}^{\infty}\lambda_k \overset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

Conversely: If $W(t)$ is a Q-Wiener process, define

$$\beta_k(t) := \begin{cases}\frac{1}{\sqrt{\lambda_k}}\langle W(t), e_k\rangle & \text{, if } \lambda_k>0\\ 0 & \text{, otherwise.}\end{cases}$$

It then follows for $s\le t$ and $k\ge 1$ that

$$\beta_k(t)-\beta_k(s) = \frac{1}{\sqrt{\lambda_k}}\underbrace{\langle \underbrace{W(t)-W(s)}_{\sim\mathcal{N}(0,(t-s)Q)}, e_k\rangle}_{\sim\mathcal{N}(0,(t-s)\lambda_k)},$$

so that $\beta_l(t), k\ge 1$ are 1-dimensional Wiener processes, hence 1-dimensional continuous Brownian motions. To see that they are indeed independent let $k\ne l$ and $s\le t$. Then:

$$\begin{aligned}\mathrm{Cov}(\beta_k(t),\beta_l(s)) &= \frac{1}{\sqrt{\lambda_k\lambda_l}}\mathbb{E}\left(\langle W(t),e_k\rangle\langle W(s),e_l\rangle\right)\\ &= \frac{1}{\sqrt{\lambda_k\lambda_l}}\left[\mathbb{E}\left(\langle W(t)-W(s),e_k\rangle\langle W(s),e_l\rangle\right)+\mathbb{E}\left(\langle W(s),e_k\rangle\langle W(s),e_l\rangle\right)\right]\\ &= \frac{1}{\sqrt{\lambda_k\lambda_l}}\left[\underbrace{\mathbb{E}\left(\langle W(t)-W(s),e_k\rangle\right)}_{=0}\underbrace{\mathbb{E}\left(\langle W(s),e_l\rangle\right)}_{=0}+s\langle Qe_k,e_l\rangle\right]\\ &= \frac{1}{\sqrt{\lambda_k\lambda_l}}s\lambda_k\underbrace{\langle e_k,e_l\rangle}_{=0}\\ &= 0\end{aligned}$$

□

Definition 2.12 (Q-Wiener process w.r.t. some filtration $(\mathcal{F}_t)$). An U-valued stochastic process $W(t)$, $t\in[0,T]$, is called Q-Wiener process w.r.t. a given filtration $(\mathcal{F}_t)$, if

(i) $W(t)$ is $\mathcal{F}_t$-measurable, i.e. W is $(\mathcal{F}_t)$-adapted,
(ii) $t\mapsto W(t)$ is $\mathbb{P}$-a.s. continuous,
(iii) $W(t)-W(s)$ is independent of $\mathcal{F}_s$ for all $0\le s<t\le T$ and $N(0,(t-s)Q)$-distributed.

Similar to the finite-dimensional case it holds that a Q-Wiener process $W(t)$, $t\in[0,T]$, always is a Q-Wiener process w.r.t. the minimal right-cotinuous filtration

$$\mathcal{F}_t^W = \bigcap_{s>t}\mathcal{F}_s^0, \quad \mathcal{F}_s^0 = \sigma(W(r): r\in[0,s])$$

generated by W.

2.3. Stochastic Integration on Hilbert spaces

The aim of this section is to define the stochastic-integral

$$\int_0^t \Phi(s)dW(s)$$

w.r.t. some H-valued Q-Wiener process.

2.3.1. Martingales in Banach spaces. Let $(E, \|\cdot\|)$ be a (real) Banach space, i.e. a complete normed ($\mathbb{R}$-) vector space. Let μ be a finite measure on a measure space $(\Omega, \mathcal{A})$.

2.3.1.1. *Bochner-Integral.* (E-valued μ-integration)

We want to define the Lebesgue integral

$$\int_\Omega f d\mu \in \mathbb{E}$$

for measurable real-valued functions $f : \Omega \longrightarrow \mathbb{R}$ to suitable sets of measurable mappings $f : \Omega \longrightarrow E$.

Similar to the case of the Lebesgue-integral we first consider elementary E-valued step functions

$$f = \sum_{k=1}^n x_k 1_{A_k} \tag{2.5}$$

with $x_k \in E$ and $A_k \in \mathcal{A}$. W.l.o.g. we may assume that A_k are pairwise disjoint.

For such integrands we then define the integral as follows:

$$\int_\Omega f\, d\mu := \sum_{k=1}^n x_k \mu(A_k) \in E,$$

and similar to the case of the Lebesgue-integral one can show that this definition is indepedent of the given representation of the elementary step function f.

It is then easy to see that the following **Bochner inequality** holds:

$$\left\| \int_\Omega f d\mu \right\|_E \leq \int_\Omega \|f\|_E d\mu.$$

Indeed,

$$\begin{aligned}
\left\|\int_\Omega f d\mu\right\|_E &= \left\| \sum_{k=1}^n x_k \mu(A_k) \right\|_E \leq \sum_{k=1}^n \|x_k\|_E \mu(A_k) \\
&= \sum_{k=1}^n \|x_k\|_E \int_\Omega 1_{A_k}\, d\mu = \int_\Omega \sum_{k=1}^n \|x_k\|_E 1_{A_k}\, d\mu \\
&= \int_\Omega \|f\|_E d\mu.
\end{aligned}$$

Using the Bocher inequality we can extend the integral $\int f d\mu$ to all integrands

$$
\begin{cases} f \quad : \Omega \longrightarrow E \text{ satisfying} \\ \quad \bullet\ f \text{ is } \textbf{strongly measurable}\text{, i.e., } f = \lim_{n\to\infty} f_n \ \mu - a.e. \\ \qquad\qquad f_n \text{ is of type (2.5)} \\ \quad \bullet\ \int \|f\|_E d\mu < \infty \end{cases} \tag{2.6}
$$

Remark For mappings $f : \Omega \longrightarrow E$ taking values in an infinite dimensional normed vector-space E, it is not true in general, whether $\mathcal{A}/\mathcal{B}(E)$-measurability implies that it can be approximated by pointwise limits elementary step functions. $\mathcal{A}/\mathcal{B}(E)$-measurability of a mapping f implies that f is **weakly measurable** in the sense that for all continuous linear functionals $\ell \in E'$ the real-valued mapping $\ell(f)$ is $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable. It is then a classical result in functional analysis, called **Pettis theorem**, that the following holds true:

THEOREM 2.13 (Pettis theorem). *A necessary and sufficient condition that $f : \Omega \longrightarrow E$ is (strongly) measurable is that it is weakly measurable and separably valued.*

Here, **separably valued** means that the range $f(\Omega)$ of the mapping f is a separable subset of E. It is obvious that if E is separable, e.g. $E = \mathbb{R}^d$, $L^p(\mathbb{R}^d)$, l^p, …, then every (weakly) measurable function $f : \Omega \longrightarrow E$ already is strongly measurable.

The set of all admissible integrands f will be denoted by $\mathcal{L}^1(\Omega, \mathcal{A}, \mu; E)$ or simply with $\mathcal{L}^1(\mu; E)$. Let $L^1(\mu; E)$ the set of all μ-equivalence classes in $\mathcal{L}^1(\mu; E)$, w.r.t. the equivalence relation $f \sim g$ if and only if $f = g$ μ-a.e. Similarly we define $\mathcal{L}^p(\mu; E)$ and $L^p(\mu; E)$.

Properties of the Bochner integral

(i) **Bochner inqeuality**

$$
\left\| \int_\Omega f d\mu \right\|_E \le \int_\Omega \|f\|_E d\mu
$$

(ii) **Linearity** Let $L : E \longrightarrow F$ be linear and continuous, where F is another Banach space. Then:

$$
L\left(\int_\Omega f\, d\mu \right) = \int_\Omega L \circ f\, d\mu.
$$

In particular: $\alpha, \beta \in \mathbb{R}$, $f = (f_1, f_2) \in E \times E$,

$$
L : E \times E \longrightarrow E, \quad (x_1, x_2) \mapsto \alpha x_1 + \beta x_2.
$$

Then

$$\alpha \int_{\Omega} f_1 d\mu + \beta \int_{\Omega} f_2 d\mu = L\left(\int_{\Omega} f d\mu\right)$$
$$= \int_{\Omega} L \circ f d\mu$$
$$= \int_{\Omega} \alpha f_1 + \beta f_2 d\mu.$$

(iii) **Fundamental theorem of calculus**
Let $f \in C^1([a,b];E)$, i.e., differentiable curve in E. Then

$$f(t) - f(s) = \int_s^t f'(r)\,dr$$

for all $a \leq s \leq t \leq b$.

2.3.1.2. *E-valued conditional expectations.* Form now on we assume that E is a separable Banach space. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We then have the following result on the existence of E-valued conditional expectations:

THEOREM 2.14. *Let $X : \Omega \longrightarrow E \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}; E)$ be a Bochner integrable E-valued random variable and $\mathcal{A}_0$ be a sub-σ-algebra of $\mathcal{A}$. Then there exists a $\mathbb{P}$-a.s. uniquely determined Bochner integrable E-valued r andom variable $X_0 \in \mathcal{L}^1(\Omega, \mathcal{A}_0, \mathbb{P}; E)$ such that*

$$\int_A X_0\, d\mathbb{P} = \int_A X\, d\mathbb{P} \qquad \forall\, A \in \mathcal{A}_0.$$

The $\mathcal{A}_0$-measurable random variable X_0 is called the conditional expectation of X given $\mathcal{A}_0$ and we write

$$X_0 =: \mathbb{E}\left(X \mid \mathcal{A}_0\right).$$

Moreover, Bochner's inequality holds

$$\left\|E\left(X \mid \mathcal{A}_0\right)\right\|_E \leq E\left(\|X\|_E \mid \mathcal{A}_0\right).$$

The proof of the theorem is first reduced to the case of elementary step functions $X = \sum_k x_k 1_{A_k}$ with A_k, $k \geq 1$, pairwise disjoint. In this case clearly,

$$E\left(X \mid \mathcal{A}_0\right) = \sum_k x_k E\left(1_{A_k} \mid \mathcal{A}_0\right),$$

and using the triangle inequality we obtain that

$$\left\|E\left(X \mid \mathcal{A}_0\right)\right\|_E = \left\|\sum_k x_k E\left(1_{A_k} \mid \mathcal{A}_0\right)\right\|_E$$
$$\leq \sum_k \|x_k\|_E E\left(1_{A_k} \mid \mathcal{A}_0\right) = E\left(\sum_k \|x_k\|_E 1_{A_k} \mid \mathcal{A}_0\right) = E\left(\|X\|_E \mid \mathcal{A}_0\right).$$

The general case can then be obtained considering pointwise limits of random variables of the above type.

EXAMPLE 2.15. Let $\mathcal{A}_0 := \sigma(A_k : k = 1, ..., n)$, A_k pairwise disjoint. Then

$$\mathbb{E}(X|\mathcal{A}_0) = \sum_{k=1}^{n} \frac{1}{\mu(A_k)} \int_{A_k} X \, d\mathbb{P} \, 1_{A_k}$$

(with the convention $\frac{1}{\mu(A_k)} \int_{A_k} X \, d\mathbb{P} := 0$, if $\mu(A_k) = 0$).

We can now extend the definition of a real-valued martingale $(M_t)_{t\geq 0}$ w.r.t. a given filtration $(\mathcal{A}_t)_{t\geq 0}$ to the general case of E-valued martingales.

DEFINITION 2.16. An E-valued stochastic process $(M_t)_{t\geq 0}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ is called an $(\mathcal{A}_t)$-martingale, if

(i) $\mathbb{E}(\|M_t\|) < \infty$ for all $t \geq 0$,
(ii) M_t is $\mathcal{A}_t$-measurable for all $t \geq 0$,
(iii) $\mathbb{E}(M_t|\mathcal{A}_s) = M_s$ for all $0 \leq s \leq t$.

Remark It is an immediate consequence of the definition that (M_t) is an $(\mathcal{A}_t)$-martingale if and only if for all $l \in E'$ the real-valued process $(l(M_t))$ is a real-valued $(\mathcal{A}_t)$-martingale.

EXAMPLE 2.17. (i) $X \in \mathcal{L}^1(\mathbb{P}; E)$, then $X_t = \mathbb{E}(X|\mathcal{A}_t)$ is a martingale.

(ii) Consider discrete time $t = 0, 1, 2, ...$ and $Y_1, Y_2, ... \in \mathcal{L}^1(\mathbb{P}; E)$, iid and $\mathcal{A}_t = \sigma(\{Y_1, ..., Y_t\})$. Then

$$X_t = \sum_{s=1}^{t} (Y_s - \mathbb{E}(Y_s))$$

is a martingale.

(iii) As a generalization of the previous example to continuous time consider a U-valued Q-Wiener process (W_t) w.r.t. $(\mathcal{A}_t)$ (where U is a separable real Hilbert space). Then (W_t) is an $\mathcal{L}^2$-integrable continuous martingale with

$$\mathbb{E}(\|W_t\|_U^2) = t \; \mathrm{tr}(Q).$$

Indeed, similar to the case of finite dimensional Brownian motion,

$$\begin{aligned}
\mathbb{E}(W_t|\mathcal{A}_s) &= \mathbb{E}(W_t - W_s + W_s|\mathcal{A}_s) \\
&= \underbrace{\mathbb{E}(W_t - W_s|\mathcal{A}_s)}_{=\mathbb{E}(W_t - W_s)=0} + \underbrace{\mathbb{E}(W_s|\mathcal{A}_s)}_{=W_s} \\
&= W_s,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(\|W_t\|_U^2) &= \mathbb{E}\left(\|\sum_{k=1}^{\infty} \sqrt{\lambda_k} \beta_k(t) e_k\|_U^2 \right) = \mathbb{E}\left(\sum_{k=1}^{\infty} \lambda_k \beta_k^2(t) \right) \\
&= \sum_{k=1}^{\infty} \lambda_k \mathbb{E}\left(\beta_k^2(t)\right) = t \sum_{k=1}^{\infty} \lambda_k \\
&= t \; \mathrm{tr}(Q).
\end{aligned}$$

Remark The notion of sub- and super-martingales cannot be immediately generalized to arbitrary Banach spaces, unless we assume that we are given some partial ordering "$\leq$" which is for example the case on L^p-spaces. We will not use this additional structure in the infinite-dimensional case and therefore do not study sub- and super-martingales in more detail.

Similar to the case of the real-valued conditional expectation we have:

THEOREM 2.18 (Doob's maximal inequality). *Let (M_t) be a right-continuous $(\mathcal{A}_t)$-martingale. Then:*

$$\left(\mathbb{E}\left(\sup_{0\leq t\leq T} \|M_t\|^p\right)\right)^{\frac{1}{p}} \leq \frac{p}{p-1} \left(\mathbb{E}\left(\|M_T\|^p\right)\right)^{\frac{1}{p}}, \quad \forall p > 1.$$

Main consequence: Let (M_t^n), $n \geq 1$, be an $\mathcal{L}^p$-Cauchy sequence of continuous martingales, i.e., $\lim\limits_{n,m\to\infty} \left(\mathbb{E}\left(\|M_T^n - M_T^m\|_E^p\right)\right)^{\frac{1}{p}} = 0$ for some $p > 1$. Then there exists a continuous martingale (M_t) (w.r.t. the same filtration), such that

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{0\leq t\leq T} \|M_t^n - M_t\|_E \geq \varepsilon\right) = 0 \quad \forall \varepsilon > 0\,,$$

hence $\lim\limits_{n\to\infty} M_t^n = M_t$ uniformly on $[0, T]$ in probability.

In other words:

$$\mathcal{M}_T^p := \{(M_t)_{t\in[0,T]} \text{ } E\text{-valued, cont. } (\mathcal{A}_t) - \text{martingale,}$$
$$\|M\|_{\mathcal{M}_T^p} := \sup_{0\leq t\leq T} \left(\mathbb{E}\left(\|M_t\|_E^p\right)\right)^{\frac{1}{p}} < \infty\}$$

is (again) a Banach space (if $p = 2$ a Hilbert space).

Lecture 4

Stochastic Analysis on Hilbert Spaces

2.3. Stochastic Integration on Hilbert spaces

2.3.2. Stochastic integral. Throughout this subsection let U and H be separable real Hilbert spaces and (W_t) be an U-valued Q-Wiener process w.r.t. $(\mathcal{F}_t)$. We want to construct the H-valued stochastic integral $\int_0^t \Phi(s)\,dW_s$ for appropriate integrands Φ in four steps similar to the real-valued case.

Step 1 Integration of elementary processes

$$\Phi(t) = \sum_{m=0}^{K-1} \Phi_m 1_{]t_m,t_{m+1}]}(t) \tag{2.7}$$

with

- $0 = t_0 < t_1 < \ldots < t_K = T$
- $\Phi_m : \Omega \longrightarrow L(U,H)$, $\mathcal{F}_{t_m}$-measurable and separable, where $L(U,H)$ denotes the (linear) space of all linear, continuous operators $U \longrightarrow H$.

We then define the stochastic integral

$$\int_0^t \Phi(s) dW_s := \sum_{m=0}^{K-1} \Phi_m (W_{t_{m+1}\wedge t} - W_{t_m \wedge t}), \quad 0 \le t \le T$$

which leads to a linear mapping

$$I : \mathcal{E} \longrightarrow \mathcal{M}_T^2$$

where $\mathcal{E}$ denotes the (linear) space of all elementary prozesses of type (2.7).

Step 2 Wiener-Itô isometry

We want to formulate the Wiener-Itô-isometry for H-valued stochastic integrals. To this end we first need to introduce a new class of linear operators:

Definition 2.19. Let $L \in L(U,H)$ be a continuous linear operator from U to H. L is called Hilbert-Schmidt operator, if

$$\|L\|_{L_2(U,H)}^2 := \sum_{k=1}^{\infty} \|Le_k\|_H^2 < \infty \tag{2.8}$$

for one (and hence all) CONS (e_k) of U. Let $L_2(U,H)$ be the (linear) space of all Hilbert-Schmidt operators L. Then $L_2(U,H)$ is again a real Hilbert space w.r.t. the scalar product

$$\langle A, B\rangle_{L_2(U,H)} := \sum_{k=1}^{\infty} \langle Ae_k, Be_k\rangle_H$$

(which is independent of the chosen CONS (e_k) (!)). The associated norm $\|\cdot\|_{L_2(U,H)}$ (as given in (2.8)) is called Hilbert-Schmidt norm.

Remark (separability of $L_2(U,H)$) A striking feature of the space $L_2(U,H)$ of Hilbert-Schmidt operators is the fact that it is again separable. Indeed, the set of finite rank operators

$$L_n u := \sum_{k=1}^n \langle u, e_k\rangle_U y_k\,, \quad y_k \in H, n \geq 1$$

is a separable dense subset of $L_2(U,H)$, since for any $L \in L_2(U,H)$ and L_n of the above type with $y_k := Le_k$ we have that $L_n e_k = Le_k$ for $k = 1,\dots,n$, and thus

$$\|L - L_n\|^2_{L_2(U,H)} = \sum_{k=n+1}^\infty \|Le_k\|^2 \to 0\,, n \to \infty.$$

In contrast to this, the space $L(U,H)$ is not separable, even not for separable Hilbert spaces U and H, unless U and H are both finite-dimensional.

Back to the stochastic intergal $\int_0^t \Phi(s)dW_s$. The following Wiener-Itô isometry holds:

$$\mathbb{E}\left(\left\|\int_0^t \Phi(s)dW_s\right\|_H^2\right) = \mathbb{E}\left(\int_0^t \|\Phi(s)\circ\sqrt{Q}\|^2_{L_2(U,H)}ds\right), \quad t \in [0,T] \tag{2.9}$$

In particular,

$$\left\|\int_0^\cdot \Phi(s)dW_s\right\|^2_{\mathcal{M}_T^2} = \mathbb{E}\left(\int_0^T \|\Phi(s)\circ\sqrt{Q}\|^2_{L_2(U,H)}ds\right).$$

Proof of the Wiener-Itô isometry.

W.l.o.g. $t = T$. To simplify notations let $\Delta_m := W(t_{m+1}) - W(t_m)$. Then

$$\begin{aligned}\mathbb{E}\left(\left\|\int_0^T \Phi(s)dW(s)\right\|_H^2\right) &= \mathbb{E}\left(\|\sum_{m=0}^{k-1}\Phi_m\Delta_m\|_H^2\right)\\ &= \sum_{m,n=0}^{k-1}\mathbb{E}\left(\langle\Phi_m\Delta_m, \Phi_n\Delta_n\rangle_H\right)\end{aligned} \tag{2.10}$$

W.l.o.g. let (e_k), $k \geq 1$, be a CONS of U consisting of eigenvectors of Q with corresponding eigenvalues λ_k, $k \geq 1$, and let $\beta_k(t) := \langle W(t), e_k\rangle_U$, $k \geq 1$, be the corresponding independent 1d-Brownian motions. For any m,n we then have that
(2.11)

$$\begin{aligned}&\mathbb{E}\left(\langle\Phi_m\Delta_m, \Phi_n\Delta_n\rangle_H\right)\\ &\quad= \sum_{k,l=1}^\infty \sqrt{\lambda_k}\sqrt{\lambda_l}\mathbb{E}\Big(\langle\Phi_m(e_k), \Phi_n(e_l)\rangle_H(\beta_k(t_{m+1}) - \beta_k(t_m))(\beta_l(t_{n+1}) - \beta_l(t_n))\Big)\end{aligned}$$

To further simplify notations, let $\Delta_n\beta_l := \beta_l(t_{n+1}) - \beta_l(t_n)$.

If $m \neq n$ or $k \neq l$ then

$$\mathbb{E}(\langle\Phi_m(e_k), \Phi_n(e_l)\rangle_H\Delta_m\beta_k\Delta_n\beta_l) = 0$$

since the Brownian increments $\Delta_m\beta_k$ and $\Delta_n\beta_l$ are independent of $\mathcal{F}_{m\wedge n}$, so that in case of $n \neq m$ or $k \neq l$

$$\mathbb{E}(\Delta_m\beta_k\Delta_n\beta_l \mid \mathcal{F}_{m\wedge n}) = \mathbb{E}(\beta_k(t_{m+1}) - \beta_k(t_m) \mid \mathcal{F}_{m\wedge n})\mathbb{E}(\beta_l(t_{n+1}) - \beta_l(t_n) \mid \mathcal{F}_{m\wedge n}) = 0$$

and therefore in this case

(2.12)
$$\mathbb{E}(\Phi_m\Delta_m\beta_k\Phi_n\Delta_n\beta_l) = \mathbb{E}(\Phi_m\Phi_n\mathbb{E}((\beta_k(t_{m+1}) - \beta_k(t_m))(\beta_l(t_{n+1}) - \beta_l(t_n)) \mid \mathcal{F}_{m\wedge n})) = 0.$$

In the remaining case $m = n$ and $k = l$ we have that

$$\mathbb{E}(\langle\Phi_m(e_k), \Phi_n(e_l)\rangle_H\Delta_m\beta_k\Delta_n\beta_l) = \mathbb{E}\left(\|\Phi_m(e_k)\|_H^2\right)(t_{m+1} - t_m). \tag{2.13}$$

Inserting both identities (2.12) and (2.13) into (2.11) yields the assertion, since

$$\begin{aligned}
\mathbb{E}\left(\left\|\int_0^T \Phi(s)dW(s)\right\|_H^2\right) &= \sum_{m=0}^{K-1}\sum_{k=1}^{\infty}\lambda_k\mathbb{E}\left(\|\Phi_m(e_k)\|_H^2\right)(t_{m+1} - t_m)\\
&= \sum_{m=0}^{K-1}\sum_{k=1}^{\infty}\mathbb{E}\left(\|\Phi_m(\sqrt{\lambda_k}e_k)\|_H^2\right)(t_{m+1} - t_m)\\
&= \sum_{m=0}^{K-1}\sum_{k=1}^{\infty}\mathbb{E}\left(\|\Phi_m(\sqrt{Q}e_k)\|_H^2\right)(t_{m+1} - t_m)\\
&= \sum_{m=0}^{K-1}\sum_{k=1}^{\infty}\mathbb{E}\left(\|(\Phi_m\circ\sqrt{Q})(e_k)\|_H^2\right)(t_{m+1} - t_m)\\
&= \mathbb{E}\left(\sum_{m=0}^{K-1}\|\Phi_m\circ\sqrt{Q}\|_{L_2(U,H)}^2(t_{m+1} - t_m)\right)\\
&= \mathbb{E}\left(\int_0^T\|\Phi(s)\circ\sqrt{Q}\|_{L_2(U,H)}^2ds\right).
\end{aligned}$$

Herein we used the fact that $Qe_k = \lambda_k e_k$ and therefore

$$\lambda_k\|\Phi_m e_k\|_H^2 = \left\|\Phi_m\left(\sqrt{\lambda_k}e_k\right)\right\|_H^2 = \left\|\Phi_m\circ\sqrt{Q}e_k\right\|_H^2.$$

□

The Wiener-Itô isometry now implies that

$$I : \mathcal{E} \longrightarrow \mathcal{M}_T^2$$

is the isometry

$$\|I(\Phi)\|_{\mathcal{M}_T^2}^2 = \|\Phi\|_T^2,$$

if we endow $\mathcal{E}$ with the seminorm

$$\|\Phi\|_T^2 := \mathbb{E}\left(\int_0^T\|\Phi\circ\sqrt{Q}\|_{L_2(U,H)}^2ds\right).$$

Step 3 We can now extend I uniquely to a linear isomery

$$\overline{I} : \overline{\mathcal{E}} \longrightarrow \mathcal{M}_T^2,$$

where $\overline{\mathcal{E}}$ denotes the abstract completion of $\mathcal{E}$ w.r.t. the seminorm $\|\cdot\|_T$. The 3rd step now in the construction/definition of the stochastic integral consists in the alternative caracterization of the abstract completion $\overline{\mathcal{E}}$ of admissible integrands. To this end we define

$$\begin{aligned}\mathcal{P}_T &:= \sigma\left(\{]s,t] \times F_s : 0 \leq s < t \leq T, F_s \in \mathcal{F}_s\} \cup \{\{0\} \times F_0 : F_0 \in \mathcal{F}_0\}\right)\\ &= \sigma\left(\{(H_s) : (H_s) \text{ left-cont. and } (\mathcal{F}_t) - \text{adapted}\}\right)\\ &= \text{predictable } \sigma - \text{algebra}.\end{aligned}$$

Remark Sometimes in the literature $\mathcal{P}_T$ is also called the previsible-σ-algebra.

With this definition we now have that

$$\overline{\mathcal{E}} = Ł^2(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; L_2^0)/_\sim$$

with $\Omega_T = \Omega \times [0,T]$ and $\mathbb{P}_T = \mathbb{P} \otimes dt$. Here, L_2^0 denotes the space of all continuous linear operators $T : U \to H$ such that $T \circ \sqrt{Q}$ is Hilbert-Schmidt and "$\sim$" denotes the equivalence relation

$$\Phi \sim \Psi \text{ if } \Phi \circ \sqrt{Q} = \Psi \circ \sqrt{Q} \; \mathbb{P}_T - a.e.$$

Step 4 Localization

We use the same localization as in the finite-dimensional stochastic integration theory. This allows us to extend the construction/definition of the stochastic integral

$$\int_0^t \Phi(s) dW(s)\, , t \in [0,T],$$

to all

$$\Phi : \Omega_T \longrightarrow L(U,H)\, , \text{adapted}, \; \mathcal{P}_T - \text{measurable},$$

satisfying

$$\mathbb{P}\left(\int_0^T \|\Phi(s) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds < \infty \right) = 1.$$

Properties of the stochastic integral

(i) **Linearity** If $L \in L(H, \tilde{H})$ then

$$L\left(\int_0^t \Phi(s) dW(s) \right) = \int_0^t L \circ \Phi(s) dW(s).$$

(ii) Let $f : \Omega_T \longrightarrow H$, be $(\mathcal{F}_t)$-adapted and continuous. Then the real-valued stochastic integral

$$\int_0^t \langle f(s), \Phi(s) dW(s) \rangle_H := \int_0^t \tilde{\Phi}_f(s) dW(s)$$

with

$$\tilde{\Phi}_f(s)(u) := \langle f(s), \Phi(s) u \rangle_H, u \in U,$$

is well-defined.

(iii) Let $p \geq 1$. Then there exists a universal constant c_p such that

$$\mathbb{E}\left(\left\|\int_0^t \Phi(s)\, dW(s)\right\|_H^{2p}\right) \leq c_p \mathbb{E}\left(\int_0^t \|\Phi(s)\|_{L_2^0}^2\, ds\right)^p .$$

In particular, the following inequality, called the Burkholder-Davis-Gundy inequality, holds

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left\|\int_0^t \Phi(s)\, dW(s)\right\|_H^{2p}\right) \leq c_p \mathbb{E}\left(\int_0^T \|\Phi(s)\|_{L_2^0}^2\, ds\right)^p$$

(iv) Quadratic variation: Let $M_t := \int_0^t \Phi(s)\, dW(s)$, then

$$\langle M\rangle_t := \int_0^t \|\Phi(s)\|_{L_2^0}^2\, ds\,, t \in [0,T]\,,$$

is the unique continuous increasing $(\mathcal{F}_t)_{t\geq 0}$-adapted process starting at zero such that

$$\|M_t\|_H^2 - \langle M\rangle_t\,, t\in[0,T]\,,$$

is an $(\mathcal{F}_t)_{t\in[0,T]}$-martingale. It can be shown that for any sequence of partitions $(\tau_n)_{n\geq 1}$ of $[0,T]$ with $\lim_{n\to\infty}|\tau_n| = 0$ it follows that

$$\lim_{n\to\infty}\sum_{t_i\in\tau_n, t_i\leq t}\|M_{t_{i+1}} - M_{t_i}\|_H^2 = \langle M\rangle_t \tag{2.14}$$

uniformly in t, in probability. More general, given any $h \in H$, the process

$$\int_0^t \|\sqrt{Q}\circ\Phi(s)^* h\|_U^2\, ds\,, t\in[0.T],$$

is the unique continuous $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process starting at zero such that

$$\langle M_t, h\rangle_H^2 - \int_0^t \|\sqrt{Q}\circ\Phi(s)^* h\|_U^2\, ds\,, t\in[0,T]\,,$$

is an $(\mathcal{F}_t)_{t\in[0,T]}$-martingale. In analogy with (2.14)

$$\lim_{n\to\infty}\sum_{t_i\in\tau_n, t_i\leq t}\big(\langle M_{t_{i+1}}, h\rangle_H - \langle M_{t_i}, h\rangle_H\big)^2 = \int_0^t \|\sqrt{Q}\circ\Phi(s)^* h\|_U^2\, ds$$

uniformly in t, in probability.

(v) Regularity of the stochastic integral: let $\alpha < \frac{1}{2}$ be given. Then for any $\Phi \in \mathcal{L}^2(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; L_2^0)$

$$M_t = \int_0^t \Phi(s)\, dW(s) \in W^{\alpha,2}([0,T];H)$$

where for a given Banach space E, the space $W^{\alpha,2}([0,T];E)$ consists of all functions $M \in L^2([0,T];E)$ satisfying

$$\int_0^T\int_0^T \frac{\|M_s - M_t\|_E^2}{|s-t|^{1+2\alpha}}\, ds\, dt < \infty\,.$$

Proof: The Wiener-Itô isometry implies that

$$\mathbb{E}\left(\int_0^T\int_0^T \frac{\|M_t - M_s\|_H^2}{|t-s|^{1+2\alpha}}\,ds\,dt\right) = \int_0^T\int_0^T \frac{\mathbb{E}\left(\int_{s\wedge t}^{s\vee t}\|\Phi(r)\|_{L_2^0}^2\,dr\right)}{|t-s|^{1+2\alpha}}\,ds\,dt$$
$$= 2\int_0^T\int_t^T \frac{\int_t^s \mathbb{E}\left(\|\Phi(r)\|_{L_2^0}^2\right)\,dr}{|t-s|^{1+2\alpha}}\,ds\,dt$$
$$\leq \ldots \leq C(\alpha)\int_0^T \mathbb{E}\left(\|\Phi(r)\|_{L_2^0}^2\right)\,dr\,.$$

2.3.3. Appendix: Stochastic integration w.r.t. cylindrical Wiener processes. The construction of stochastic integrals $\int_0^t \Phi(s)\,dW(s)$ can be extended to the case where the covariance operator Q is only bounded, but not necessarily of finite trace. To this end, one needs to extend the notion of a Q-Wiener process. To simplify the presentation, we restrict ourselves to the particular case $Q = I$.

The representation of the Q-Wiener process obtained in Proposition 2.11 leads in the case $Q = I$ to the infinite series

$$W(t) = \sum_{k=1}^\infty \beta_k(t)e_k\,,\,t\in[0,T]\,,$$

for independent one-dimensional Brownian motions β_k, $k\geq 1$. Note that this series does not converge in U, since

$$\left\|\sum_{k=1}^n \beta_k(t)e_k\right\|_U^2 = \sum_{k=1}^n \beta_k(t)^2$$

and thus

$$\mathbb{E}\left(\left\|\sum_{k=1}^n \beta_k(t)e_k\right\|_U^2\right) = \sum_{k=1}^n \mathbb{E}(\beta_k(t)^2) = n\cdot t\uparrow\infty$$

for $n\to\infty$. However, for any Hilbert space $(U_1, \langle\ ,\ \rangle_{U_1})$ for which there exists a Hilbert-Schmidt embedding $J: U\to U_1$ the infinite series converges in U_1, since

$$\mathbb{E}\left(\left\|J\left(\sum_{k=1}^n \beta_k(t)e_k\right)\right\|_{U_1}^2\right) = \sum_{k,l=1}^n \mathbb{E}\left(\langle\beta_k(t)J(e_k), \beta_l(t)J(e_l)\rangle_{U_1}\right)$$
$$= \sum_{k=1}^n \mathbb{E}\left(\beta_k(t)^2\right)\|J(e_k)\|_{U_1}^2 \text{ since } \mathbb{E}\left(\beta_k(t)\beta_l(t)\right) = \delta_{kl}t$$
$$= \sum_{k=1}^n t\,\|J(e_k)\|_{U_1}^2 \uparrow t\,\|J\|_{L_2(U,U_1)}^2\,.$$

REMARK 2.20. U_1 with the above properties always exists. For example, choose a sequence $(\alpha_k)_{k\geq 1}\in\ell^2$ with $\alpha_k\neq 0$ for all k, let $U_1 = U$ and define

$$J: U\to U_1\,,\,u\mapsto\sum_{k=1}^\infty \alpha_k\langle u, e_k\rangle_U e_k\,.$$

In the following we fix a sequence of independent one-dimensional Brownian motions β_k, $k \geq 1$, a CONS $(e_k)_{k\geq 1}$ of U and a Hilbert space U_1 for which there exists a Hilbert-Schmidt embedding $J : U \to U_1$. In particular, $Q_1 := J \circ J^* \in L(U_1)$ is symmetric, positive definite with finite trace and the infinite series

$$W_1(t) = \sum_{k=1}^{\infty} \beta_k(t) J(e_k)\,, t \in [0,T]\,,$$

converges in $\mathcal{M}_T^2(U_1)$ and defines a Q_1-Wiener process on U_1.

For a given predictable process Φ satisfying

$$\mathbb{P}\left(\int_0^T \|\Phi(s)\|^2_{L_2(U,H)}\,ds < \infty\right) = 1\,,$$

using

$$\|\Phi(s)\|^2_{L_2(U,H)} = \|\Phi(s) \circ J^{-1}\|^2_{L_2(\sqrt{Q_1}(U_1),H)}\,,$$

we conclude that the stochastic integral

$$\int_0^t \Phi(s) \circ J^{-1}\,dW_1(s)$$

w.r.t. the Q_1-Wiener process is well-defined. Finally, we set

$$\int_0^t \Phi(s)\,dW(s) := \int_0^t \Phi(s) \circ J^{-1}\,dW_1(s)\,.$$

The class of admissible integrands is given by

$$\begin{aligned}\mathcal{N}_W = \{\Phi : \Omega_T \to L_2(U,H) \mid \Phi \text{ predictable and} \\ \mathbb{P}\left(\int_0^T \|\Phi(s)\|^2_{L_2(U,H)}\,ds < \infty\right) = 1\}\,.\end{aligned}$$

Lecture 5

Stochastic Analysis on Hilbert Spaces

2.4. Strongly continuous semigroups

Throughout the whole section let $(E, \|\cdot\|)$ be a Banach space.

We will need some further basic functional analytic facts on linear operators A on the Banach space E. Mainly facts on the linear initial value problem

$$\begin{cases} \frac{d}{dt}u(t) &= Au(t)\,, t \geq 0 \\ u(0) &= u_0 \in D(A). \end{cases} \tag{2.15}$$

Here, $D(A)$ denotes the domain of A. If A is bounded (with full domain $D(A) = E$), a solution to (2.15) is given in terms of the operator exponential

$$u(t) := e^{tA}u_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k u_0, t \geq 0\,.$$

Since the most relevant operators are however unbounded, like operators involving the differential of a function, we need to find a suitable generalization. This is provided with the theory of semigroups.

Definition 2.21. A family $(T_t)_{t\geq 0} \subseteq L(E,E)$ of bounded continuous linear operators on E is called a strongly-continuous semigroup (resp. C_0-semigroup) (of linear operators) if

(i) $T_0 = \mathrm{Id}_E =: I$,
(ii) $T_t \circ T_s = T_{t+s}$ for all $s, t \geq 0$,
(iii) $t \mapsto T_t u$ is continuous for all $u \in E$.

Given a C_0-semigroup we can define a linear operator A with domain $D(A)$ as the derivative of $T_t u$ in $t = 0$, if this derivative exists in E:

$$Au := \lim_{t \searrow 0} \frac{1}{t}(T_t u - u) \in E$$

with domain

$$D(A) := \left\{ u \in E : \exists \lim_{t \searrow 0} \frac{1}{t}(T_t u - u) \in E \right\}.$$

The linear operator $(A, D(A))$ is called the **infinitesimal generator** of the C_0-semigroup.

Remark 2.22. (i) $(A, D(A))$ is densely defined, i.e., $D(A) \subseteq E$ is dense.
(ii) A is a closed linear operator, i.e., the graph

$$\Gamma(A) := \{(u, Au) \mid u \in D(A)\} \subset E \times E$$

is a closed subset in $E \times E$. This is equivalent to say that if a sequence $(u_n)_{n\geq 1} \subset D(A)$ converges to some element $u \in E$ and if in addition the sequence $(Au_n)_{n\geq 1}$ is E-Cauchy, then this implies that $u \in D(A)$ und $Au = \lim_{n\to\infty} Au_n$ in E.

The resolvent set $\rho(A)$ of a closed linear operator A is defined as

$$\rho(A) := \{\lambda \in \mathbb{C} : \lambda I - A : D(A) \longrightarrow E \text{ bijective}\}.$$

For $\lambda \in \rho(A)$ the operator

$$R_\lambda := (\lambda I - A)^{-1} : E \longrightarrow D(A)$$

is well-defined and continuous (!) due to the closed graph theorem. R_λ is called the resolvent of A. The complement $\mathbb{C}\backslash\rho(A)$ is called the **spectrum** of A.

THEOREM 2.23. *(Hille-Yosida theorem) Let $(A, D(A))$ be a closed (linear) Operator on E. Then th following statements are equivalent:*

(i) *A is the infinitesimal generator of a C_0-semigroup $(T_t)_{t\geq 0}$ with*

$$\|T_t\|_{L(E)} \leq Me^{\omega t}, \quad t \geq 0,$$

for given constants M, ω.

(ii) *$D(A) \subseteq E$ dense, $\rho(A) \supset]\omega, \infty[$ and*

$$\|R_\lambda^k\|_{L(E)} \leq \frac{M}{(\lambda - \omega)^k}, \quad k = 1, 2, ..., \lambda > \omega.$$

In this case

$$R_\lambda = \int_0^\infty e^{-\lambda t} T_t dt.$$

REMARK 2.24. A C_0-semigroup $(T_t)_{t\geq 0}$ is called a C_0-semigroup of contractions if $\|T_t\|_{L(E)} \leq 1$ for all $t \geq 0$. In this particular case the Hille-Yosida theorem simplifies considerably, because in this case the following statments are equivalent:

(i) A is the infinitesimal generator of C_0-semigroup of contractions $(T_t)_{t\geq 0}$.

(ii) $]0, \infty[\subseteq \rho(A)$ and

$$\|\lambda R_\lambda\|_{L(E)} \leq 1, \quad \forall \lambda > 0.$$

EXAMPLE 2.25. (i) **Laplace operator** on $E = L^2(R^d)$.

Let $W(t)$, $t \geq 0$, be the d-dimensional (continuous) Brownian motion. Then Itô's formula applied to some function $f \in C^2(\mathbb{R}^d)$ yields

$$f(W(t)) = f(W(0)) + \int_0^t \nabla f(W(s)) dW(s) + \frac{1}{2} \int_0^t \sum_{i=1}^d \underbrace{\partial^2_{x_i x_i} f(W(s))}_{=\Delta f(W(s))} ds.$$

If we define

$$T_t f(x) := \mathbb{E} f(W(t) + x) = \frac{1}{\sqrt{2\pi t}^d} \int_{\mathbb{R}^d} f(y) e^{-\frac{\|x-y\|^2}{2t}} dy$$

then $(T_t)_{t\geq 0}$ is a C_0-semigroup on $L^2(\mathbb{R}^d)$ and one can show that

$$Af := \frac{1}{2}\Delta f = \frac{1}{2} \sum_{i=1}^d \partial^2_{x_i x_i} f$$

is the infinitesimal generator, since

$$\begin{aligned}\frac{1}{h}(T_h(f(x)) - f(x)) &= \frac{1}{h}\mathbb{E}(f(W(h)+x) - f(x)) \\ &= \frac{1}{h}\mathbb{E}\Big(\int_0^h \nabla f(W(s)+x)dW(s) + \frac{1}{2}\int_0^h \Delta f(W(s)+x)ds\Big) \\ &= \underbrace{\frac{1}{h}\mathbb{E}\left(\int_0^h \nabla f(W(s)+x)dW(s)\right)}_{=0} + \frac{1}{2}\mathbb{E}\Big(\underbrace{\frac{1}{h}\int_0^h \Delta f(W(s)+x)ds}_{=\Delta f(W(h)+x)}\Big) \\ &\xrightarrow{h\searrow 0} \frac{1}{2}\mathbb{E}\left(\Delta f(W(0)+x)\right) = \frac{1}{2}\Delta f(x).\end{aligned}$$

The precise domain of the infinitesimal generator Δ is given as $D(\Delta) = H^{2,2}(\mathbb{R}^d)$, which is the (Hilbert-) space of all L^2-integrable functions having L^2-integrable weak derivatives up to second order.

(ii) **Laplace operator** on $E = L^2(D)$, $D \subseteq \mathbb{R}^d$ open. Define the first exit time

$$\tau_D = \inf\{t \geq 0 : W(t) + x \notin D\}.$$

Then the family $(T_t)_{t\geq 0}$ of linear operators

$$T_t f(x) := \mathbb{E}(f(W(t)+x)|\tau_D > t), \quad t \geq 0$$

is a C_0-semigroup.

Also in this case $\frac{1}{2}\Delta$ is the infinitesimal generator with domain $D(\Delta) \subseteq H_0^{1,2}(D) \cap H^{2,2}(D)$, where $H_0^{1,2}(D)$ denotes the (Hilbert-) space of two rimes weakly differentiable functions f with f and all partial derivatives up to second order L^2-integrable and f vanishing on the boundary ∂D of D.

Particular case $D =]0,1[$

In this case, the functions

$$e_k := \sqrt{2}\sin(\pi k x), \quad k = 1, 2, \ldots$$

are a CONS, consisting of eigenfunctions of Δ with corresponding eigenvalues $-\pi^2 k^2$. In particular,

$$T_t e_k(x) := e^{-\frac{t}{2}\pi^2 k^2} e_k(x), \quad k \in \mathbb{N}.$$

Hence, for general $f \in L^2(]0,1[)$ we then have the following **spectral representation** of the semigroup T_t

$$T_t f(x) = \sum_{k=1}^{\infty} \langle f, e_k\rangle_{L^2} e^{-\frac{t}{2}\pi^2 k^2} e_k(x).$$

In particular,

$$\frac{d}{dt}T_t(f)\Big|_{T=t} = -\pi^2 \sum_{k=1}^{\infty} k^2 \langle f, e_k\rangle_{L^2} e_k$$

and

$$D(\Delta) = \left\{ f \in L^2(]0,1[) : \sum_{k=1}^{\infty} k^4 \langle f, e_k \rangle^2 < \infty \right\}$$

Stochastic differential equations on Hilbert spaces

3.1. Mild, weak and strong solutions

Throughout the whole section we fix two separable (real) Hilbert spaces U, H and a Q-Wiener process $(W_t)_{t\geq 0}$ w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ on U. Consider the equation

$$\begin{cases} dX_t &= [AX_t + B(X_t)]\, dt + C(X_t)\, dW_t \in H \\ X_0 &= \xi \end{cases} \tag{3.1}$$

with

(A.1) $(A, D(A))$ generates a C_0-semigroup $(T_t)_{t\geq 0}$ on H
(A.2) $B : H \to H$ is $\mathcal{B}(H)$-measurable
(A.3) $C : H \to L_2(U_0, H)$ is strongly continuous, i.e.,

$$x \mapsto C(x)u\,, H \to H$$

is continuous for all $u \in U_0$.
(A.4) ξ is H-valued and $\mathcal{F}_0$-measurable.

Notions of solutions

- **mild solution:** An H-valued predictable process $(X_t)_{t\in[0,T]}$ satisfying

$$X_t = T_t\xi + \int_0^t T_{t-s}B(X_s)\, ds + \int_0^t T_{t-s}C(X_s)\, dW_s \qquad \mathbb{P}-a.s.\,,$$

for all $t \in [0,T]$, where all integrals have to be well-defined.
- **(analytically) strong solution:** A $D(A)$-valued predictable process $(X_t)_{t\in[0,T]}$ satisfying

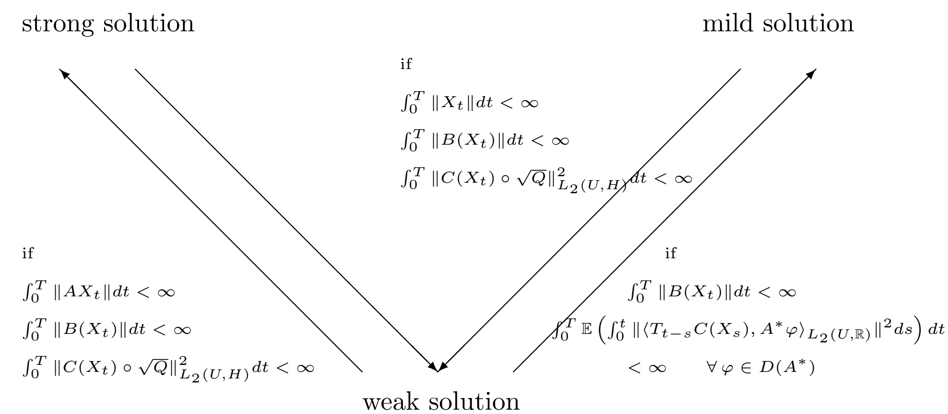
$$X_t = \xi + \int_0^t [AX_s + B(X_s)]\, ds + \int_0^t C(X_s)\, dW_s \qquad \mathbb{P}-a.s.\,,$$

for all $t \in [0,T]$, where all integrals have to be well-defined.
- **(analytically) weak solution:** An H-valued predictable process $(X_t)_{t\in[0,T]}$ satisfying

$$\begin{aligned}\langle X_t, \varphi\rangle_H = \langle \xi, \varphi\rangle_H + \int_0^t \langle X_s, A^*\varphi\rangle_H + \langle B(X_s), \varphi\rangle_H\, ds \\ + \int_0^t \langle \varphi, C(X_s)\, dW_s\rangle_H \qquad \mathbb{P}-a.s.\,,\end{aligned}$$

for all $t \in [0,T]$, $\varphi \in D(A^*)$. Here, $(A^*, D(A^*))$ is the dual operator of A and it is required that all integrals are well-defined.

<u>Interrelations between the different notions of solutions</u>



3.1.1. Existence and uniqueness of mild solutions. In this subsection we discuss existence and uniqueness of mild solutions of (3.1) under the following additional assumption:

$$\text{(H.1)} \qquad \begin{aligned} &\|B(t,x) - B(t,y)\|_H + \|(C(t,x) - C(t,y)) \circ \sqrt{Q}\|_{L_2(U,H)} \\ &\qquad \le L\|x-y\|_H \quad \forall x, y \in H, t \in [0,T], \end{aligned}$$

$$\text{(H.2)} \qquad \|B(t,x)\|_H^2 + \|C(t,x) \circ \sqrt{Q}\|^2_{L_2(U,H)} \le M\left(1 + \|x\|_H^2\right) \quad \forall x \in H, t \in [0,T].$$

Theorem 3.1. *Under hypotheses (A.1)-(A.4), (H.1) and (H.2) there exists a unique mild solution of* (3.1) *satisfying*

$$\mathbb{P}\left(\int_0^T \|X_s\|_H^2 \, ds < \infty\right) = 1 .$$

$(X_t)_{t\in[0,T]}$ *has a continuous modification and*

$$\mathbb{E}\left(\sup_{t\in[0,T]} \|X_t\|_H^p\right) \le c_{p,T}\left(1 + \mathbb{E}\left(\|\xi\|^p\right)\right) \qquad \forall p > 2 .$$

Proof. Application of Banach's fixed point theorem.

Fix $p \ge 2$ and define

$$\mathcal{H}_p := \left\{ Y : \Omega_T \longrightarrow H : Y \text{ predictable, } \|Y\|_p^p := \sup_{t\in[0,T]} \mathbb{E}\left(\|Y(t)\|_H^p\right)^{\frac{1}{p}} < \infty \right\}.$$

Then $(\mathcal{H}_p, \|\cdot\|_p)$ is a Banach space.

For $Y \in \mathcal{H}_p$ we define

$$\begin{aligned} \mathcal{K}(Y)(t) &:= T_t(\xi) + \int_0^t T_{t-s}(B(s,Y(s)))ds + \int_0^t T_{t-s}(C(s,Y(s)))dW(s) \\ &= T_t(\xi) + \mathcal{K}_1(Y)(t) + \mathcal{K}_2(Y)(t), \end{aligned}$$

with

$$\mathcal{K}_1(Y)(t) := \int_0^t T_{t-s}(B(s,Y(s)))ds$$

and

$$\mathcal{K}_2(Y)(t) := \int_0^t T_{t-s}(C(s,Y(s)))dW(s).$$

We will show that if $\mathbb{E}(\|\xi\|_H^p) < \infty$ this implies that $\mathcal{K}$ is a mapping

$$\mathcal{K} : \mathcal{H}_p \longrightarrow \mathcal{H}_p$$

and a strict contraction for T sufficiently small.

Concerning $\mathcal{K}_1$:

$$\begin{aligned}
\mathbb{E}(\|\mathcal{K}_1(Y)(t)\|_H^p) &= \mathbb{E}\left(\|\int_0^t T_{t-s}(B(s,Y(s)))ds\|_H^p \right) \\
&\overset{\text{Bochner}}{\leq} \mathbb{E}\left(\left(\int_0^t \|T_{t-s}(B(s,Y(s)))\|_H ds \right)^p \right) \\
&\overset{\text{Hölder}}{\leq} t^{p-1}\mathbb{E}\Bigg(\int_0^t \underbrace{\|T_{t-s}(B(s,Y(s)))\|_H^p}_{\leq M_1^p \underbrace{\|B(s,Y(s))\|_H^p}_{\overset{(H.2)}{\leq} 2^p M^p (1+\|Y(s)\|_H^p)}} ds \Bigg) \\
&\overset{t\leq T}{\leq} 2^p M_1^p M^p T^p (1 + \|Y\|_p^p)
\end{aligned}$$

For the stochastic part $\mathcal{K}_2$ we will apply Burkholder-Davis-Gundy inequality:

$$\begin{aligned}
\mathbb{E}(\|\mathcal{K}_2(Y)(t)\|_H^p) &= \mathbb{E}\left(\|\int_0^t T_{t-s}(C(s,Y(s)))dW(s)\|_H^p \right) \\
&\leq c_p \mathbb{E}\left(\left(\int_0^t \|T_{t-s}(C(s,Y(s))) \circ \sqrt{Q}\|_{L_2(U,H)}^2 ds \right)^{\frac{p}{2}} \right) \\
&\leq c_p M_1^p \mathbb{E}\left(\left(\int_0^t \|C(s,Y(s)) \circ \sqrt{Q}\|_{L_2(U,H)}^2 ds \right)^{\frac{p}{2}} \right) \\
&\underset{\substack{(H.2)\\ t\leq T}}{\overset{\text{Hölder}}{\leq}} c_p M_1^p 2^p M^p T^{\frac{p+1}{2}} (1 + \|Y\|_p^p)
\end{aligned}$$

These estimates show that $\mathcal{K}(Y) \in \mathcal{H}_p$. To see that $\mathcal{K}$ is a strict contraction for small T, observe that similar to the above estimates

$$\begin{aligned}
\mathbb{E}(\|\mathcal{K}(Y_1)(t) - \mathcal{K}(Y_2)(t)\|_H^p) &\leq 2^p \left[\mathbb{E}\left(\|(\mathcal{K}_1(Y_1) - \mathcal{K}_1(Y_2))(t)\|_H^p\right) + \mathbb{E}\left(\|(\mathcal{K}_2(Y_1) - \mathcal{K}_2(Y_2))(t)\|_H^p\right) \right] \\
&\overset{(H.1)}{\leq} \underbrace{c_p(T^p + T^{\frac{p+1}{2}})}_{<1 \text{ for } T \text{ small}} \|Y_1 - Y_2\|_p^p.
\end{aligned}$$

The existence of a continuous modification will follow from Proposition 3.2 below.

□

3.1.2. Existence of a continuous modification. First assume that $\mathbb{E}(\|\xi\|_H^{2p}) < \infty$ for some $p > 1$, so that in particular $\mathbb{E}\left(\int_0^T \|X_t\|^{2p} dt\right) < \infty$. Let $\Phi(t) := C(t, X_t)$, then

$$\mathbb{E}\left(\int_0^T \|\Phi(t) \circ \sqrt{Q}\|_{L_2(U,H)}^{2p} dt\right) \leq M^{2p} \mathbb{E}\left(\int_0^T (1 + \|X_t\|^2)^p dt\right) < \infty.$$

The following proposition now implies that the process

$$\int_0^t T_{t-s} C(s, X_s) dW_s, \quad t \in [0, T],$$

has a continuous modification.

The general case $\mathbb{E}(\|\xi\|^{2p}) = \infty$ follows from approximating ξ with $\xi_n = \xi 1_{\{\|\xi\|_H \leq n\}}$.

PROPOSITION 3.2. *Let $p > 1$ and $\Phi \in \mathcal{L}^{2p}(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; L_2)$, i.e.*

$$\mathbb{E}\left(\int_0^T \|\Phi(s) \circ \sqrt{Q}\|_{L_2(U,H)}^{2p} ds\right) < \infty.$$

Then there exists a constant C_T, such that

$$\mathbb{E}\left(\sup_{t \in [0,T]} \|\int_0^t T_{t-s}\Phi(s) dW_s\|_H^{2p}\right) \leq C_T \mathbb{E}\left(\int_0^T \|\Phi(s) \circ \sqrt{Q}\|_{L_2(U,H)}^{2p} ds\right). \tag{3.2}$$

Moreover, $W_{A,\Phi}(t) := \int_0^t T_{t-s}\Phi(s) dW_s$ has a continuous modification.

PROOF. We will use a technique, called the **factorization method**, based on the following integral identity:

$$\int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} ds = \frac{\pi}{\sin(\pi\alpha)} \quad \forall \alpha \in \left(0, \frac{1}{2}\right) \forall s < t.$$

Indeed, substituting $u = \frac{r-s}{t-s}$ yields

$$\int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} dr = \int_0^1 (1-u)^{\alpha-1} u^{-\alpha} du = \frac{\Gamma(\alpha)\Gamma(1-\alpha)}{\Gamma(1)} = \frac{\pi}{\sin(\pi\alpha)} =: \frac{1}{c_\alpha}.$$

We can thus rewrite the convolution

$$
\begin{aligned}
W_{A,\Phi}(t) &= \int_0^t T_{t-s}(\Phi(s))dW(s) \\
&= c_\alpha \int_0^t \int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} dr T_{t-s}(\Phi(s))dW(s) \\
&= c_\alpha \int_0^t \int_0^r (t-r)^{\alpha-1}(r-s)^{-\alpha} \underbrace{T_{t-s}(\Phi(s))}_{(T_{t-r}\circ T_{r-s})(\Phi(s))} dW(s)dr \\
&= c_\alpha \int_0^t (t-r)^{\alpha-1} T_{t-r}\underbrace{\left(\int_0^r (r-s)^{-\alpha} T_{r-s}(\Phi(s))dW(s)\right)}_{=:y(r)} dr.
\end{aligned} \tag{3.3}
$$

Here we used in the third equality that we can interchange the order of integration w.r.t. r and s ("stochastic Fubini theorem"), and in the fourth step we have used linearity of (T_t).

Hence for $p > \frac{1}{2\alpha} > 1$ it follows that

$$
\begin{aligned}
\|W_{A,\Phi}(t)\|_H^{2p} &= c_\alpha^{2p} \|\int_0^t (t-r)^{\alpha-1} T_{t-r}(y(r))dr\|_H^{2p} \\
&\le c_\alpha^{2p} \left(\int_0^t (t-r)^{\alpha-1}\|T_{t-r}(y(r))\|_H dr\right)^{2p} \\
&\le c_\alpha^{2p} \left(\int_0^t (t-r)^{\frac{2p}{2p-1}(\alpha-1)} dr\right)^{2p-1} \int_0^t \|T_{t-r}(y(r))\|_H^{2p} dr \\
&\le c_{\alpha,p,T} \int_0^t \|T_{t-r}(y(r))\|_H^{2p} dr
\end{aligned}
$$

Here we used Bochner's inequality in the first step, then Hölder's inequality in the second step (applied to $2p$ and $\frac{2p}{2p-1}$). In the last step we used the fact that

$$
p > \frac{1}{2\alpha} \Leftrightarrow \frac{2p}{2p-1}(\alpha-1) > \frac{2p}{2p-1}\left(\frac{1}{2p}-1\right) = -1.
$$

It follows that

$$
\sup_{t\in[0,T]} \|W_{A,\Phi}(t)\|_H^{2p} \le c_{\alpha,p,T} \int_0^T \|y(r)\|_H^{2p} dr. \tag{3.4}
$$

The Burkholder-Davis-Gundy inequality implies that

$$
\mathbb{E}\left(\|y(r)\|_H^{2p}\right) \le C_p \mathbb{E}\left(\left(\int_0^r (r-s)^{-2\alpha}\|\Phi(s)\circ\sqrt{Q}\|_{L_2}^2\, ds\right)^p\right),
$$

hence
(3.5)

$$\begin{aligned}\int_0^T \mathbb{E}\left(\|y(r)\|_H^{2p}\right) dr &\le C_p \mathbb{E}\left(\int_0^T \left(\int_0^r (r-s)^{-2\alpha}\|\Phi(s))\circ\sqrt{Q}\|_{L_2}^2\, ds\right)^p dr\right)\\ &\le C_{\alpha,p,T}\mathbb{E}\left(\int_0^T\int_0^r (r-s)^{-2\alpha}\|\Phi(s))\circ\sqrt{Q}\|_{L_2}^{2p}\, ds\, dr\right)\\ &\le \tilde{C}_{\alpha,p,T}\mathbb{E}\left(\int_0^T \|\Phi(s))\circ\sqrt{Q}\|_{L_2}^{2p}\, ds\right).\end{aligned}$$

In the second step we have used Jensens' inequality, applied to the probability measure

$$\frac{1}{Z}(r-s)^{-2\alpha}\, ds \quad \text{ on } [0,r], \text{ with } Z := \int_0^r (r-s)^{-2\alpha}\, ds = \frac{r^{1-2\alpha}}{1-2\alpha} \le \frac{T^{1-2\alpha}}{1-2\alpha},$$

since

$$\left(\int_0^r (r-s)^{-2\alpha}\|\Phi(s)\circ\sqrt{Q}\|_{L_2}^2\, ds\right)^p \le Z^{p-1}\int_0^r (r-s)^{-2\alpha}\|\Phi(s)\circ\sqrt{Q}\|_{L_2}^{2p}\, ds.$$

Combining with (3.4) implies

$$\mathbb{E}\left(\sup_{t\in[0,T]}\|W_{A,\Phi}(t)\|_H^{2p}\right) \le C_{\alpha,p,T}\mathbb{E}\left(\int_0^T \|\Phi(s)\circ\sqrt{Q}\|_{L_2}^{2p}\, ds\right).$$

We will show next that the mapping

$$F : y \mapsto \int_0^t (t-r)^{\alpha-1}T_{t-r}y(r)\, dr$$

maps $L^{2p}([0,T];H)$ into $C([0,T];H)$ which will imply together with (3.3) and (3.5) that $W_{A,\Phi}$ has a continuous modification.

Note that similar to the proof of (3.4)

$$\sup_{t\in[0,T]}\|F(y)(t)\|_H^{2p} \le c_{\alpha,T,p}\int_0^T \|y(t)\|_H^{2p}dt.$$

Since $C([0,T];H) \subset L^{2p}([0,T];H)$ is dense, it therefore suffices to show that $t \mapsto F(y)(t)$ is continuous, provided $t \mapsto y(t)$ is continuous. Indeed, we can approximate $y \in L^{2p}([0,T];H)$ by a sequence $(y_n) \subset C([0,T];H)$ converging to y in $L^{2p}([0,T];H)$. Then $F(y_n)$ converges uniformly to $F(y)$, which implies that $F(y)$ is continuous too.

So assume that $t \mapsto y(t)$ is continuous. We will then show that $F(y)$ is right-/left-continuous:

$$\begin{aligned}
&\|F(y(t+h)) - F(y(t))\|_H \\
&\quad = \left\| \int_0^{t+h} (t+h-r)^{\alpha-1} T_{t+h-r}(y(r))dr - \int_0^t (t-r)^{\alpha-1} T_{t-r}(y(r))dr \right\|_H \\
&\quad \leq \left\| \int_0^t (t+h-r)^{\alpha-1} T_{t-r}(T_h(y(r)))dr - \int_0^t (t-r)^{\alpha-1} T_{t-r}(y(r))dr \right\|_H \\
&\qquad\qquad + \left\| \int_t^{t+h} (t+h-r)^{\alpha-1} T_{t+h-r}(y(r))dr \right\|_H \\
&\quad \leq \int_0^t \|\left((t+h-r)^{\alpha-1} T_h - (t-r)^{\alpha-1}\right) T_{t-r}(y(r))\|_H dr \\
&\qquad\qquad + \int_t^{t+h} (t+h-r)^{\alpha-1} \|T_{t+h-r}(y(r))\|_H dr \\
&\quad \stackrel{h\searrow 0}{\longrightarrow} 0,
\end{aligned}$$

since $\left((t+h-r)^{\alpha-1}T_h - (t-r)^{\alpha-1}\right) T_{t-r}(y(r)) \longrightarrow 0$ in H and bounded in the H-norm by $2(t-r)^{\alpha-1}$ up to some uniform constant, and $(t+h-r)^{\alpha-1}T_{t+h-r}y(r)$ also bounded in the H-norm up to some uniform constant by $(t+h-r)^{\alpha-1}$.

The left-continuity is shown in a similar way, hence the assertion is proven. □

Lecture 7

Stochastic differential equations on Hilbert spaces

3.2. Stochastic differential equations with additive noise

In the particular case where the dispersion coefficient C does not depend on the solution, equation (3.1) is called a stochastic differential equation with additive noise. This case is very close to the deterministic analogue. Indeed, let

$$W_A(t) := \int_0^t T_{t-s} C \, dW_s \,, \qquad t \in [0,T] \,,$$

be the stochastic convolution and suppose that $(W_A(t))_{t\in[0,T]}$ has a version with continuous trajectories in H. Decomposing the mild solution

$$X_t = Y_t + W_A(t)$$

we formally obtain the following equation

$$dY_t = [AY_t + B(Y_t + W_A(t))] \, dt \,, Y_0 = \xi \tag{3.6}$$

for $(Y_t)_{t\in[0,T]}$. Equation (3.6) can be seen as a deterministic evolution equation with time-dependent random coefficients

$$B(\cdot + W_A(t)) \,.$$

In particular, if (3.6) has a unique mild solution $(Y_t(\omega))_{t\in[0,T]}$ for P-a.e. ω and the dependence of $(Y_t)_{t\in[0,T]}$ on ω is predictable, then $X_t = Y_t + W_A(t)$, $t \in [0,T]$, is a mild solution of (3.1).

2D-Stochastic Navier Stokes equations with additive noise

Let $D \subset \mathbb{R}^2$ be a bounded open domain with regular boundary ∂D. We consider the following stochastic Navier-Stokes equations

$$\left\{ \begin{array}{rcll} \partial_t u(t,x) - \nu \Delta u(t,x) + (u(t,x)\cdot\nabla) u(t,x) + \nabla p(t,x) & = & \dot{\xi}_t(x) & t \in [0,T], x \in D \\ \text{div } u(t,x) & = & 0 & t \in [0,T], x \in D \\ u(t,x) & = & 0 & t \in [0,T], x \in \partial D \\ u(0,x) & = & u_0(x) & x \in D \,, \end{array} \right. \tag{3.7}$$

where $(\xi_t)_{t\ge 0}$ is a (cylindrical) Wiener process. Here, $u : [0,T] \times D \to \mathbb{R}^2$ is the velocity field, $\nu > 0$ the viscosity and $p : [0,T] \times D \to \mathbb{R}$ denotes the pressure. We will consider the equation in similar function spaces as for the deterministic case:

$$D_0^\infty = \left\{ u \in C_0^\infty(D;\mathbb{R}^2) \,, \ \text{div } u = 0 \right\}$$

$$H = \text{ closure of } D_0^\infty \text{ in } L^2(D;\mathbb{R}^2) \text{ w.r.t. } \|u\|_H^2 := \int_D |u|^2 \, dx$$

$$V = \text{ closure of } D_0^\infty \text{ in } L^2(D;\mathbb{R}^2) \text{ w.r.t. } \|u\|_V^2 := \int_D |Du|^2 \, dx$$

Applying the Helmholtz projection $\Pi : L^2(D,\mathbb{R}^2) \to H$ one obtains the following abstract evolution equation

$$\begin{cases} du(t) & = [Au(t) + B(u(t),u(t)) + f(t)] \, dt + C \, dW_t \\ u(0) & = u_0 \end{cases} \tag{3.8}$$

on H, where

- $A = \Pi\Delta_D$ is the Stokes operator on H
- $B : V \times V \to V'$, ${}_{V'}\langle B(u,v), w\rangle_V = -\int_D w(x) \cdot (u(x) \cdot \nabla) v(x)\, dx$.

THEOREM 3.3. *If* $(W_A(t))_{t\in[0,T]}$ *has a version in* V *with continuous trajectories, then* (3.8) *has a unique mild solution.*

PROOF. In this case, equation (3.6) can be written as

$$dY_t = [\nu A Y_t + B(Y_t + W_A(t))]\, dt\,, Y_0 = u_0\,. \tag{3.9}$$

It is a classical result, that (3.9) has for ω with $t \mapsto W_A(t)(\omega)$, $[0,T] \to V$ continuous, a unique solution $Y_\cdot \in L^2([0,T];V)$, $\dot{Y}_\cdot \in L^1([0,T];V')$ satisfying also

$$\begin{aligned}\sup_{t\in[0,T]} \|Y_t\|_H^2 + \nu \int_0^T \|Y_t\|_V^2\, dt \\ \le \exp\left(\frac{2}{\nu}\int_0^T \|W_A(t)(\omega)\|_V^2\, dt\right) \\ \left(\|u_0\|_H^2 + \frac{1}{\nu}\int_0^T \|W_A(t)(\omega)\|_H^2 \|W_A(t)(\omega)\|_V^2\, dt\right).\end{aligned}$$

It can be also shown that the dependence of the unique solution $Y_\cdot$ on ω is predictable, so that $X_t = Y_t + W_A(t)$ is a mild solution of (3.8). □

REMARK 3.4. Regularity properties of the stochastic convolution W_A are well-studied (see the monographs [1,2]). The main difficulty with W_A is that it is not a martingale w.r.t. t. This does not contradict the properties of the stochastic integral, because for any $t > 0$ the process

$$W_A^{(t)}(s) = \int_0^s e^{(t-r)A} C\, dW_r\,, s \in [0,t]\,,$$

is a martingale up to time t.

References:

1 G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, 1992.
2 G. Da Prato, J. Zabczyk, Ergodicity for infinite dimensional systems, Cambridge University Press, 1996.

Lecture 8

CHAPTER 4

Variational approach to stochastic partial differential equations

4.1. Insertion from functional analysis: Weak topology

We start this chapter with a short insertion on basic functional analytic facts needed in the following. Basic facts concerning weak topology, Banach-Alaoglu's theorem, Riesz' representation theorem will be needed. In particular we will need the basic fact that weak convergece + norm convergence implies strong convergence in Hilbert spaces.

4.2. Variational approach to partial differential equations

Having recalled the required basic functional analytic facts, we are now ready for the introduction to the variational approach to stochastic partial differential equations. We first explain the basic concept in the deterministic case. To this end we will first need the concept of a Gelfand triple.

Let H be a separable, real Hilbert space and let $V \subseteq H$ be a second Hilbert space such that the embedding $V \hookrightarrow H$ is dense and continuous. We assume that

$$\|h\|_H \leq \|h\|_V$$

for all $h \in V$.

Identifying H with its dual space $H' = L(H, \mathbb{R})$ (using the Riesz-Isomorphism) we obtain a dense and continuous embedding

$$H \cong H' \hookrightarrow V'$$

via

$$\underbrace{h}_{\in H \cong H'} \quad \underbrace{\cong}_{\substack{\text{is the} \\ \text{embedding}}} \quad \underbrace{(v \mapsto \langle h, v\rangle_H, v \in V)}_{\in L(V,\mathbb{R}) = V'}$$

for all $h \in H$. This means, that an element h from $H \cong H'$ is identified with the linear functional

$$l_h(v) := \langle h, v\rangle_H, \quad v \in V.$$

Note that

$$\|l_h\|_{V'} := \sup_{\|v\|_V \leq 1} \underbrace{l_h(v)}_{\langle h,v\rangle_H} \leq \sup_{\|v\|_V \leq 1} \|h\|_H \|v\|_H \leq \sup_{\|v\|_V \leq 1} \|h\|_H \|v\|_V \leq \|h\|_H,$$

hence, indeed, the embedding $H \equiv H' \hookrightarrow V'$ is continuous.

Let $A \in L(V, V')$ be a bounded (hence continuous) linear operator $A : V \longrightarrow V'$, satisfying the following **coercivity condition**

$\exists \lambda, \alpha > 0$ such that

$$-\langle Au, u\rangle \ \geq \alpha \|u\|_V^2 - \lambda \|u\|_H^2 \quad \forall u \in V.$$

EXAMPLE 4.1. Let $D \subset \mathbb{R}^d$ be an open subset, $H = L^2(D), V = H^1(D)$, where

$$H^1(D) = \{u \in L^2(D) : \partial_{x_i} u \in L^2(D), i = 1, ..., d\}$$

is the Sobolev space of one times weakly differentiable functions f with f and all partial derivatives in $L^2(D)$.

Remark (weak derivatives - Sobolev spaces) $\partial_{x_i} u$ is understood in the weak sense, i.e., $u \in H^1(D)$ if and only if for all i there exists $u^{(i)} \in L^2(D)$ such that the following integration by parts formula

$$\int_D \partial_{x_i} \varphi(x) u(x)\, dx = -\int_D \varphi(x) u^{(i)}(x)\, dx \tag{4.1}$$

holds for all "test-functions" $\varphi \in C_c^1(D)$. Here, $C_c^1(D)$ denotes the space of one times differentiable functions having compact support in D. The function $u^{(i)}$ is uniquely determined via the integration by parts formula (4.1). Indeed, let $\tilde{u}^{(i)}$ be another weak derivative, satisfying the integration by parts formula (4.1), then

$$\int_D \varphi(x) \left(u^{(i)}(x) - \tilde{u}^{(i)}(x)\right) dx = -\int_D \partial_{x_i} \varphi(x) u(x) dx + \int_D \partial_{x_i} \varphi(x) u(x) dx = 0$$

for all $\varphi \in C_c^1(D)$ and thus $u^{(i)} = \tilde{u}^{(i)}$.

It is clear that for $u \in C_c^1(\mathbb{R}^d)$ the weak derivative $\partial_{x_i} u$ (in $L^2(\mathbb{R}^d)$) exists and coincides with the usual partial derivative, since

$$\int_D \partial_{x_i} \varphi(x) u(x)\, dx = -\int_D \varphi(x) \partial_{x_i}(x)\, dx$$

$H^1(D) = H^{1,2}(D)$ is again a separable real Hilbert space w.r.t. the inner product

$$\langle u, v \rangle := \int_D uv dx + \sum_{i=1}^d \int_D \partial_{x_i} u \partial_{x_i} v dx.$$

Denote with $H_0^1(D) \subset H^1(D)$ the closure of $C_c^1(D) \subset H^1(D)$. Then this closure is called Sobolev space (of order 1) with Dirichlet boundary conditions.

Illustration ($d = 1$). Let $D = (0, 1)$. $u \in H^1(D)$ implies that u is absolute continuous, i.e.,

$$u(x) - u(y) = \int_x^y u'(s) ds$$

with Radon-Nikodym derivative $u' \in L^2(D)$. In particular, u has a continuous version $\tilde{u}$. Moreover, $u \in H_0^1(D)$ if and only if $u \in H^1(D)$ and $\tilde{u}(0) = \tilde{u}(1) = 0$.

Having introduced the Gelfand triple

$$H^1(D) \hookrightarrow L^2(D) = L^2(D)' \hookrightarrow H^1(D)'$$

we can now realize the Laplace operator

$$\Delta := \sum_{i=1}^d \partial_{x_i x_i}^2$$

as a continuous (!) linear operator $\Delta : H^1(D) \longrightarrow H^1(D)'$ as follows:

Given some $u \in H^1(D)$ let $\Delta u \in H^1(D)'$ be the unique linear functional defined by

$$v \mapsto {}_{H^1(D)'}\langle \Delta u, v\rangle_{H^1(D)} = -\sum_{i=1}^{d} \int\limits_D \partial_{x_i} u\, \partial_{x_i} v\, dx, \quad v \in H^1(D).$$

Let us now consider the initial value problem

$$\begin{cases} \frac{du}{dt}(t) &= Au(t) + f(t),\ t \geq 0 \\ u(0) &= u_0. \end{cases} \tag{4.2}$$

THEOREM 4.2. *Let $A \in L(V, V')$ be coercive, $u_0 \in H$ and $f \in L^2([0,T]; V')$. Then equation (4.2) has a unique solution*

$$u \in C([0,T]; H) \cap L^2([0,T]; V).$$

We need the following lemma for the proof:

LEMMA 4.3. *Let $u \in L^2([0,T]; V)$ be such that $t \mapsto u(t)$ is absolute continuous with values in V' and suppose that $\frac{du}{dt} \in L^2([0,T]; V')$. Then $u \in C([0,T]; H)$ and*

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|_H^2 = {}_{V'}\langle \frac{du}{dt}(t), u(t)\rangle_V \quad a.e..$$

PROOF. First note that $u(t) - u(s) = \int\limits_s^t u'(r)\, dr$ implies for all $v \in L^2([0,T]; V)$ that

$$\begin{aligned} {}_{V'}\langle u(t) - u(s), v\rangle_V &= \int\limits_s^t \langle u'(r), v\rangle dr \overset{\text{C.S.}}{\leq} \int\limits_s^t \|u'(r)\|_{V'} \|v\|_V\, dr \\ &\overset{\text{C.S.}}{\leq} \sqrt{t-s} \left(\int\limits_s^t \|u'(r)\|_{V'}^2 dr \right)^{\frac{1}{2}} \|v\|_V \overset{s \to t}{\longrightarrow} 0. \end{aligned}$$

Applying the above inequality to $v = u(t) - u(s)$ we conclude that $\frac{1}{t-s}\|u(t) - u(s)\|_H^2$ is bounded for all $s < t$. In particular, $\frac{1}{h}\|u(t+h) - u(t)\|_H^2$ is bounded in t, h, so that

$$\begin{aligned} \frac{1}{2h}(\|u(t+h)\|_H^2 - \|u(t)\|_H^2) &= \frac{1}{2h} \int\limits_t^{t+h} {}_{V'}\langle u'(r), u(t+h) + u(t)\rangle_V\, dr \\ &\underset{\substack{\text{Lebesgue} \\ \text{diff.} \\ \text{lemma}}}{\overset{h \searrow 0}{\longrightarrow}} {}_{V'}\langle u'(t), u(t)\rangle_V. \end{aligned}$$

□

PROOF OF THEOREM 4.2. **Uniqueness** Let $u, v \in L^2([0,T]; V)$ be two solutions of (4.2). Then $u - v$ solves the parabolic equation

$$\frac{d(u-v)}{dt}(t) = A(u(t) - v(t))$$

with initial condition $u(0) - v(0) = 0$. Lemma (4.3) now implies that

$$\frac{1}{2}\frac{d}{dt}\|u(t) - v(t)\|_H^2 = {}_{V'}\langle A(u(t) - v(t)), u(t) - v(t)\rangle_V \overset{\text{coercive}}{\leq} \lambda \|u(t) - v(t)\|_H^2.$$

Consequently,

$$\begin{aligned}\frac{d}{dt}\left(e^{-2\lambda t}\|u(t)-v(t)\|_H^2\right) &= e^{-2\lambda t}\left(-2\lambda\|u(t)-v(t)\|_H^2 + 2\langle A(u(t)-v(t)), u(t)-v(t)\rangle\right)\\ &\le e^{-2\lambda t}\left(-2\lambda\|u(t)-v(t)\|_H^2 + 2\lambda\|u(t)-v(t)\|_H^2\right) = 0.\end{aligned}$$

Since $u(0)-v(0)=0$ and $e^{-2\lambda t}$ is strictly positive it follows that

$$\|u(t)-v(t)\|_H = 0$$

and therefore $u(t)-v(t)=0$ proving the uniqueness part.

Existence
We will apply an approximation procedure called **Galerkin-method**. To this end let $(e_k)_k$ be a CONS of H consisting of elements in V. It is always possible to choose such a system, since V is dense in H. Let

$$V_n := \operatorname{span}\{e_1, ..., e_n\}$$

be the n-dimensional subspace of V (and H too) spanned by the first n basis vectors. It follows from the theory of (finite) systems of linear ordinary differential equations that for all $n \in \mathbb{N}$ there exists $u_n \in C([0,T]; V_n)$ such that

$$\begin{cases} \frac{d}{dt}\langle u_n(t), e_k\rangle &= \langle Au_n(t), e_k\rangle + \langle f(t), e_k\rangle,\ t \ge 0\\ \langle u_n(0), e_k\rangle &= \langle u_0, e_k\rangle \end{cases} \tag{4.3}$$

for $k = 1, \dots, n$.

The next step consist of taking the limit $n \to \infty$. An essential ingredient for this is to derive the following a priori estimate

$$\sup_{n\ge 1}\left(\sup_{t\in[0,T]}\|u_n(t)\|_H^2 + \int_0^T \|u_n(t)\|_V^2 dt\right) < \infty \tag{4.4}$$

which implies that the approximating sequence $(u_n)_{n\ge 1}$ will be bounded in the appropriate function spaces. For the proof of the a priori estimate note that (4.3) implies

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u_n(t)\|_H^2 &= \frac{1}{2}\frac{d}{dt}\sum_{k=1}^n \langle u_n(t), e_k\rangle^2\\ &= \sum_{k=1}^n \langle Au_n(t) + f(t), e_k\rangle\langle u_n(t), e_k\rangle\\ &= \langle Au_n(t) + f(t), u_n(t)\rangle\\ &\le -\alpha\|u_n(t)\|_V^2 + \lambda\|u_n(t)\|_H^2 + \|f(t)\|_{V'}\|u_n(t)\|_V\\ &\le -\alpha\|u_n(t)\|_V^2 + \lambda\|u_n(t)\|_H^2 + \frac{1}{\sqrt{\alpha}}\|f(t)\|_{V'}\sqrt{\alpha}\|u_n(t)\|_V\\ &\overset{\text{Young-}}{\underset{\text{ineq.}}{\le}} -\frac{\alpha}{2}\|u_n(t)\|_V^2 + \lambda\|u_n(t)\|_H^2 + \frac{1}{2\alpha}\|f(t)\|_{V'}^2.\end{aligned}$$

Integrating both sides of the inequality w.r.t. t yields that

$$\begin{aligned}
&\|u_n(t)\|_H^2 + \alpha \int_0^t \|u_n(s)\|_V^2 ds \\
&\qquad \leq \|u_0\|_H^2 + 2\lambda \int_0^t \|u_n(s)\|_H^2 ds + \frac{1}{\alpha} \int_0^t \|f(s)\|_{V'} ds \\
&\qquad \leq \|u_0\|_H^2 + \frac{1}{\alpha} \int_0^t \|f(s)\|_{V'} ds + 2\lambda \int_0^t \left(\|u_n(s)\|_H^2 + \alpha \int_0^s \|u_n(r)\|_V^2 dr \right) ds
\end{aligned}$$

Finally, Gronwall's lemma implies the a priori estimate (4.4).
Using this boundedness, we can extract a subsequence, again denoted with u_n, converging weakly to some $u \in L^2([0,T];V)$. Since $Au_n \longrightarrow Au$ weakly in $L^2([0,T];V')$ we obtain from (4.3) that

$$\frac{d}{dt}\langle u(t), e_k\rangle = \langle Au(t) + f(t), e_k\rangle, \qquad k = 1, 2, \ldots$$

and therefore

$$\frac{d}{dt} u(t) = AU(t) + f(t)$$

which proves the assertion. □

4.2.1. Generalization to the nonlinear case. We assume that $A : V \longrightarrow V'$ is still continuous and coercive, but not necessarily linear. The a priori estimate (4.4) in the existence part in the proof of Theorem 4.2 remains true, so that we can still extract a weakly convergent subsequence $u_n \to u$ in $L^2([0,T];V)$ satisfying $A(u_n) \to g$ in $L^2([0,T];V')$ for some unknown element g. In contrast to the linear case we cannot immediately conclude that $A(u) = g$. In order to do this, there are the following two strategies:

1. Monotonicity (Dissipativity)
We assume the following monotonicity condition

$$\langle A(u) - A(v), u - v\rangle \leq \lambda \|u - v\|_H^2 \quad \text{for some } \lambda \in \mathbb{R}, \tag{4.5}$$

together with the boundedness assumption of the type

$$\|A(u)\|_{V'} \leq M(1 + \|u\|_V),$$

and a certain continuity assumption (see below).
For simplicity assume that (4.5) holds with $\lambda = 0$, otherwise we can pass w.l.o.g. from the operator $A(\cdot)$ to the operator $A(\cdot) - \lambda I$. Then

$$\int_0^T \langle A(u_n) - A(v), u_n - v\rangle(t) dt \leq 0 \quad \forall v \in L^2([0,T];V).$$

The left hand side can be divided up into four terms. Three of these terms,

$$\int_0^T \langle A(u_n), v\rangle(t) dt, \ \int_0^T \langle A(v), u_n\rangle(t) dt \text{ and } \int_0^T \langle A(v), v\rangle(t) dt$$

converge towards the expected limit. The only remaining difficult term is

$$\int_0^T \langle A(u_n), u_n\rangle(t)dt = \frac{1}{2}\left(\|u_n(T)\|_H^2 - \sum_{k=1}^n \langle u_0, e_k\rangle^2\right) - \int_0^T \langle f, u_n\rangle(t)dt.$$

Again, the second and third term converge towards the expected limit, the first term not necessarily. But, since u_n, $n \geq 1$, is bounded, we can extract a subsequence such that both, $u_n \to u$ weakly in $L^2([0,T];V)$ und $u_n(T) \to u(T)$ weakly in H. Similar to the linear case we then have

$$\frac{d}{dt}\langle u(t), e_k\rangle \;=\; \langle g(t) + f(t), e_k\rangle, \quad k = 1, 2, \ldots$$

and therefore $\frac{d}{dt}u(t) = g(t) - f(t) \in V'$ which implies, using Lemma 4.3, that

$$\begin{aligned}\int_0^T \langle g(t), u(t)\rangle dt &= \frac{1}{2}\left(\|u(T)\|_H^2 - \|u_0\|_H^2\right) - \int_0^T \langle f(t), u(t)\rangle dt \\ &\leq \limsup_{n\to\infty} \frac{1}{2}\left(\|u_n(T)\|_H^2 - \sum_{k=1}^n \langle u_0, e_k\rangle^2\right) - \int_0^T \langle f(t), u_n(t)\rangle dt \\ &\leq \limsup_{n\to\infty} \int_0^T \langle A(u_n(t)), u_n(t)\rangle dt.\end{aligned}$$

In fact we have for all $v \in L^2([0,T];V)$

$$\int_0^T \langle g(t) - A(v(t)), u(t) - v(t)\rangle dt \leq \limsup_{n\to\infty} \int_0^T \langle A(u_n(t)) - A(v(t)), u_n(t) - v(t)\rangle dt \leq 0.$$

Choosing $v(t) = u(t) + \theta w(t)$ for any $w \in L^2([0,T];V)$ and $\theta > 0$, implies that

$$\int_0^T \langle g(t) - A((u + \theta w)(t)), w(t)\rangle dt \leq 0.$$

Taking the limit $\theta \searrow 0$ and assuming that $\theta \mapsto A((u + \theta w)(t))$ is continuous, we arrive at

$$\int_0^T \langle g(t) - A(u(t)), w(t)\rangle dt \leq 0 \quad \forall w \in L^2([0,T];V).$$

Since w is arbitrary we can conclude from this $g = A(u)$.

This last argument is called the **"monotonicity trick"**.

2. Compactness method

The basic additional assumption in this method is that the embedding $V \hookrightarrow H$ is compact (and not only continuous). This means that that every bounded subset $A \subset V$ is precompact in H. This is not in general true for infinite dimensional spaces, since the Bolzano-Weierstrass theorem only holds in finite dimensional vector spaces.

LEMMA 4.4. *Let the embedding $V \hookrightarrow H$ be compact. Then*

$$X := L^2([0,T];V) \cap H^{1,2}([0,T];V') \hookrightarrow L^2([0,T];H)$$

is compact too. In particular, any sequence (u_n) bounded in X contains a subsequence strongly convergent in $L^2([0,T];H)$.

PROOF. For $\varepsilon > 0$ define

$$J_\varepsilon g(s) := \frac{1}{2\varepsilon}\int_{s-\varepsilon}^{s+\varepsilon} g(t)\,dt = \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} g(s+t)\,dt.$$

(We extend any integrable mapping $g : [0,T] \to V$ by 0 to $g : \mathbb{R} \to V$.) Clearly, J_ε defines a bounded linear operator from $L^2([0,T];V)$ into $C([0,T];V)$ (resp. from $L^2([0,T];H)$ into $C([0,T];H)$), since

$$\begin{aligned}\|J_\varepsilon g(s)\|_V &= \frac{1}{2\varepsilon}\left\|\int_{-\varepsilon}^{\varepsilon} g(s+t)\,dt\right\|_V \le \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon}\|g(s+t)\|_V\,dt \\ &\le \frac{1}{\sqrt{2\varepsilon}}\left(\int_{-\varepsilon}^{\varepsilon}\|g(s+t)\|_V^2\,dt\right)^{\frac12} \le \frac{1}{\sqrt{2\varepsilon}}\|g\|_{L^2([0,T];V)}.\end{aligned}$$

We will next show that for $\varepsilon > 0$ the mapping $J_\varepsilon : X \to C([0,T];H)$ is compact using Arzela-Ascoli's theorem. To this end first note that for $g \in X$

$$\frac{d}{ds}J_\varepsilon g(s) = \underbrace{\frac{1}{2\varepsilon}(g(s+\varepsilon) - g(s-\varepsilon))}_{\in L^2([0,T];V)} = \underbrace{\frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon}\frac{d}{ds}g(s+t)\,dt}_{\in L^2([0,T];V')},$$

hence

$$\begin{aligned}\|J_\varepsilon g(t_2) - J_\varepsilon g(t_1)\|_H &= \left\|\int_{t_1}^{t_2}\frac{d}{ds}J_\varepsilon g(s)\,ds\right\|_H \\ &\le \frac{1}{2\varepsilon}\int_{t_1}^{t_2}\|g(s+\varepsilon) - g(s-\varepsilon g)\|_H\,ds \\ &\le \frac{1}{2\varepsilon}\sqrt{t_2-t_1}\left(\int_{t_1}^{t_2}\|g(s+\varepsilon) - g(s-\varepsilon g)\|_H^2\,ds\right)^{\frac12} \\ &\le \frac{\sqrt{t_2-t_1}}{\varepsilon}\|g\|_X.\end{aligned}$$

Suppose now that $A \subset X$ is bounded. Then $A \subset C([0,T];H)$ is bounded and uniformly equicontinuous, since

$$\sup_{|t_1-t_2|\le\delta}\sup_{g\in A}\|g(t_1) - g(t_2)\|_H \overset{\delta\to 0}{\longrightarrow} 0.$$

Now Arzela-Ascoli's theorem can be applied to conclude that A is relatively compact in $C([0,T],H)$ and thus also in $L^2([0,T];H)$.

We will prove next that

$$\|J_\varepsilon g - g\|_{L^2([0,T];V')} \le \frac{2\sqrt{\varepsilon T}}{3}\|g\|_X. \tag{4.6}$$

This will prove the assertion, since for $g_n \in X$ converging weakly to some $g \in X$, is follows that $J_\varepsilon g_n \to J_\varepsilon g$ strongly in $L^2([0,T];H)$ for all $\varepsilon > 0$ and thus $g_n \to g$ strongly in $L^2([0,T];H)$. Indeed: given $\tilde{\varepsilon} > 0$, there exists $C_{\tilde{\varepsilon}}$ such that

$$\|u\|_{L^2([0,T];H)} \le \tilde{\varepsilon}\|u\|_{L^2([0,T];V)} + C_{\tilde{\varepsilon}}\|u\|_{L^2([0,T];V')}$$

for all $u \in L^2([0,T];V)$. Therefore,

$$\begin{aligned}
&\|g_n - g\|_{L^2([0,T];H)} \\
&\quad \le \|g_n - J_\varepsilon g_n\|_{L^2([0,T];H)} + \|J_\varepsilon(g_n - g)\|_{L^2([0,T];H)} + \|J_\varepsilon g - g\|_{L^2([0,T];H)} \\
&\quad \le \tilde{\varepsilon}\|g_n - J_\varepsilon g_n\|_{L^2([0,T];V)} + C_{\tilde{\varepsilon}}\|g_n - J_\varepsilon g_n\|_{L^2([0,T];V')} + \|J_\varepsilon(g_n - g)\|_{L^2([0,T];H)} \\
&\qquad + \tilde{\varepsilon}\|J_\varepsilon g - g\|_{L^2([0,T];V)} + C_{\tilde{\varepsilon}}\|J_\varepsilon g - g\|_{L^2([0,T];V')}.
\end{aligned}$$

We can now choose $\varepsilon > 0$ such that

$$\sup_n C_{\tilde{\varepsilon}}\|J_\varepsilon g_n - g_n\|_{L^2([0,T];V')} < \tilde{\varepsilon}$$

so that

$$\limsup_{n\to\infty} \|g_n - g\|_{L^2([0,T];H)} \le 2\tilde{\varepsilon}\sup_n \|g_n - J_\varepsilon g_n\|_{L^2([0,T];V)} + 2\tilde{\varepsilon}$$

which implies the assertion, since $\tilde{\varepsilon}$ was arbitrary.

It remains to prove (4.6). But this follows from

$$\begin{aligned}
\|J_\varepsilon g(s) - g(s)\|_{V'} &= \left\|\frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} g(s+t) - g(s)\,dt\right\|_{V'} = \left\|\frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon}\int_0^t \frac{d}{du}g(s+u)\,du\,dt\right\|_{V'} \\
&\le \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon}\sqrt{|t|}\left(\int_0^t \left\|\frac{d}{du}g(s+u)\right\|_{V'}^2 du\right)^{\frac{1}{2}} dt \\
&\le \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon}\sqrt{|t|}\,dt\|g\|_X = \frac{2\sqrt{\varepsilon}}{3}\|g\|_X
\end{aligned}$$

and therefore

$$\|J_\varepsilon g - g\|_{L^2([0,T];V')} \le \sqrt{T}\sup_{s\in[0,T]}\|J_\varepsilon g(s) - g(s)\|_{V'} \le \frac{2\sqrt{\varepsilon T}}{3}\|g\|_X.$$

□

Remark With this lemma we can now prove Theorem 4.2 for continuous nonlinear $A : V \mapsto V'$ satisfying the coercivity assumption.

Lecture 9

Variational approach to stochastic partial differential equations

4.3. The variational approach to stochastic partial differential equations

Let $(W_t)_t$ be an U-valued Q-Wiener process, where U is a separable real Hilbert space. Let V, H be as in the previous section. The following set of assumptions will set up the basic framework for our analysis.

Monotone-coercive SPDEs

Let $A : V \longrightarrow V'$ and $B : V \longrightarrow L(U, H)$ satisfy the following conditions:

(V.1) $\exists \alpha > 0, \lambda, \nu$ such that

$$2\langle A(u), u\rangle + \|B(u) \circ \sqrt{Q}\|^2_{L_2(U,H)} \leq -\alpha\|u\|^2_V + \lambda\|u\|^2_H + \nu \quad \forall u \in V.$$

(V.2) $\exists \lambda > 0$ such that

$$2\langle A(u) - A(v), u - v\rangle + \|(B(u) - B(v)) \circ \sqrt{Q}\|^2_{L_2(U,H)} \leq \lambda\|u - v\|^2_H \quad \forall u, v \in V.$$

(V.3) $\exists M > 0$ such that

$$\|A(u)\|_{V'} \leq M\,(1 + \|u\|_V) \quad \forall u \in V.$$

(V.4) $\forall u, v, w \in V$ such that the mapping

$$\lambda \mapsto \langle A(v + \lambda u), w\rangle$$

is continuous.

Given A and B as above, we want to solve the following stochastic evolution equation

$$du(t) = A(u(t))dt + B(u(t))dW(t), \quad u(0) = u_0. \tag{4.7}$$

The following theorem is the main result of the variational approach to SPDE:

Theorem 4.5. *Under the assumptions (V.1) - (V.4) there exists for all $u_0 \in H$ a uniquely determined adapted process $u(t)$, $t \geq 0$, such that $u \in L^2([0,T]; V) \cap C([0,T]; H)$ and satisfying the equation*

$$\langle u(t), v\rangle = \langle u_0, v\rangle + \int_0^t \langle A(u(s)), v\rangle ds + \int_0^t \langle v, B(u(s))dW(s)\rangle.$$

For the proof we will need the following version of Itô's lemma:

Lemma 4.6. *Let $u_0 \in H$, u (resp. v) be an adapted process with paths in $L^2([0,T]; V)$ (resp. $L^2([0,T]; V')$) and let M be a continuous H-valued local martingale, such that*

$$u(t) = u_0 + \int_0^t v(s)ds + M_t.$$

Then:

(i) $u \in C([0,T];H)$ *a.s.,*

(ii) $\|u(t)\|_H^2 = \|u_0\|_H^2 + 2\int\limits_0^t \langle v(s), u(s)\rangle ds + 2\int\limits_0^t \langle u(s), dM_s\rangle + \langle M\rangle_t$ *for all* $t \in [0,T]$ *a.s.*

PROOF. For all $k \in \mathbb{N}$ we have that

$$u_k(t) := \langle u(t), e_k\rangle = \langle u_0, e_k\rangle + \int\limits_0^t \langle v(s), e_k\rangle ds + \underbrace{\langle M_t, e_k\rangle}_{=:M_t^k}$$

is a continuous real-valued local semimartingale, such that Itô's-formula now implies that

$$du_k^2(t) = 2\langle v(t), e_k\rangle u_k(t)dt + 2u_k(t)dM_t^k + d\langle M^k\rangle_t.$$

Summing up w.r.t. k yields

$$\begin{aligned}
\|u(t)\|_H^2 &= \sum_{k=1}^\infty u_k^2(t) \\
&= \underbrace{\sum_{k=1}^\infty \langle u_0, e_k\rangle^2}_{=\|u_0\|_H^2} + 2\int\limits_0^t \underbrace{\sum_{k=1}^\infty \langle v(s), e_k\rangle u_k(s)\, ds}_{=\langle v(s), u(s)\rangle} \\
&\quad + 2\int\limits_0^t \underbrace{\sum_{k=1}^\infty u_k(s) dM_s^k}_{=\langle u(s), dM_s\rangle} + \underbrace{\sum_{k=1}^\infty \langle M^k\rangle_t}_{=\langle M\rangle_t}
\end{aligned}$$

This proves (ii). For the proof of (i) note that $u \in C([0,T];V')$ (since $u_0 \in H \subset V'$ is continuous, $\int_0^t v(s)\, ds \in V'$ is continuous and $M_t \in H \subset V'$ continuous) and therefore it remains to show that $t \mapsto u(t)$ is continuous in H w.r.t. the weak topology, since (ii) implies that $t \mapsto \|u(t)\|_H$ is continuous.
To see the weak continuity, consider first some $h \in V$. Then

$$\langle u(t), h\rangle = \langle u_0, h\rangle + \int\limits_0^t \langle v(s), h\rangle ds + \langle M_t, h\rangle$$

is obviously continuous. For arbitrary $h \in H$ choose an approximating sequence $(h_m)_m \subset V$, converging towards h in H. Then

$$\sup_{t\in[0,T]} |\langle u(t), h - h_m\rangle| \le \underbrace{\sup_{t\in[0,T]} \|u(t)\|_H}_{<\infty} \|h - h_m\|_H \overset{m\nearrow\infty}{\longrightarrow} 0.$$

Indeed, $\sup\limits_{t\in[0,T]} \|u(t)\|_H < \infty$ since $u \in L^\infty([0,T];H)$ due to part (ii) pf the lemma.
It follows that $\langle u(t), h_m\rangle \to \langle u(t), h\rangle$ uniformly w.r.t. t, hence

$$t \mapsto \langle u(t), h\rangle$$

is continuous too. □

We also need the following version of Itô's formula:

LEMMA 4.7. *Let $\Phi : H \longrightarrow \mathbb{R}$ be two time continuously Frechet differentiable with $\Phi'(u) \in V$ for all $u \in V$ and such that $u \mapsto \Phi'(u)$, $V \longrightarrow V$ is continuous w.r.t. the weak topology and $\|\Phi'(u)\|_V \leq M\,(1 + \|u\|_V)$. Then the following Itô-formula holds:*

$$\Phi(u(t)) = \Phi(u_0) + \int_0^t \langle v(s), \Phi'(u(s))\rangle ds + \int_0^t \langle \Phi(u(s)), dM_s\rangle + \frac{1}{2}\int_0^t tr\,(\Phi''(u(s))Q(s))\,ds.$$

Here, $Q(s)$ is the integrated covariance of M, i.e., the nonnegative semidefinite and symmetric linear operator defined by

$$\int_0^t \langle Q(s)h, g\rangle_H ds = \langle\langle M, h\rangle_\cdot, \langle M, g\rangle_\cdot\rangle_t.$$

The proof of the Lemma is an exercise.

PROOF OF THEOREM 4.5.

Uniqueness Let $u, v \in L^2([0,T]; V) \cap C([0,T]; H)$ be two adapted solutions of (4.7). Define the stopping times

$$\tau_n := \inf\{t \in [0,T] : \|u(t)\|_H^2 > n \text{ or } \|v(t)\|_H^2 \text{ or } \int_0^t \|u(s)\|_V^2 \vee \|v(s)\|_V^2 ds > n\}.$$

Then $\tau_n \overset{n\nearrow\infty}{\longrightarrow} T$ a.s. and Lemma 4.6, applied to $u - v$, yields:

$$\begin{aligned}
\|u - v\|_H^2 &= 2\int_0^t \langle A(u(s)) - A(v(s)), u(s) - v(s)\rangle ds \\
&\quad + 2\int_0^t \langle u(s) - v(s), (B(u(s)) - B(v(s)))dW(s)\rangle \\
&\quad + \int_0^t \|(B(u(s)) - B(v(s))) \circ \sqrt{Q}\|_{L_2(U,H)}^2 ds \\
&\overset{(V.2)}{\leq} \underbrace{\int_0^t \langle u(s) - v(s), (B(u(s)) - B(v(s))dW(s))\rangle}_{\text{martingale up to } \tau_n \text{ for all } n\in\mathbb{N}} + \lambda\int_0^t \|u(s) - v(s)\|_H^2 ds.
\end{aligned}$$

Taking expectations yields

$$\mathbb{E}(\|u(t\wedge\tau_n) - v(t\wedge\tau_n)\|_H^2) \leq \lambda\int_0^t \mathbb{E}(\|u(s\wedge\tau_n) - v(s\wedge\tau_n)\|_H^2)ds,$$

so that Gronwall's lemma now implies that

$$\|u(t\wedge\tau_n) - v(t\wedge\tau_n)\|_H = 0 \quad \text{a.s.}$$

Taking the limit $n \to \infty$ we obtain that

$$u(t) - v(t) \quad \text{a.s.}$$

which proves uniqueness.

The proof of **existence** is carried out in the following two lemmata. □

Again, let $V_n = \text{span}\{e_1, ..., e_n\}$.

LEMMA 4.8. *For all n there exists an adapted process $u_n \in C([0,T];V_n)$, such that*

(4.8)
$$\langle u_n(t), e_k\rangle = \langle u_n(t), e_k\rangle = \langle u_0, e_k\rangle + \int_0^t \langle A(u_n(s)), e_k\rangle ds + \int_0^t B_k(u_n(s))dW^n(s)$$

where

- $B_k(u)h := \langle B(u)h, e_k\rangle$,
- $W^n(t) := \sum_{l=1}^{n} \langle W(t), f_l\rangle f_l$, *with (f_l) being a CONS of U.*

(W.l.o.g. let (f_l) be eigenvectors of Q with corresponding eigenvalues $\lambda_l \geq 0$.)

LEMMA 4.9. *The following a priori estimate holds:*

$$\sup_n \mathbb{E}\left(\sup_{t\in[0,T]} \|u_n(t)\|_H^2 + \int_0^T \|u_n(s)\|_V^2 ds\right) < \infty.$$

We will postpone the proof of both lemmata and first complete the proof of the existence part of Theorem 4.5.

Proof of existence part of Theorem 4.5

Lemma 4.9 implies, that the sequence $(u_n)_n$ is bounded in $L^2(\Omega; C([0,T];H)) \cap L^2(\Omega_T;V)$. This implies, using assumption (V.2):

(i) $(A(u_n)_n$ is bounded in $L^2(\Omega_T;V') = L^2(\Omega;L^2([0,T];V'))$,
(ii) $(B(u_n)\circ\sqrt{Q})_n$ is bounded in $L^2(\Omega_T;L_2(U,H)) = L^2(\Omega;L^2([0,T];L_2(U,H)))$.

Hence there exists a subsequence again denoted with $(u_n)_n$, such that

(i) $u_n \longrightarrow u$ weakly in $L^2(\Omega;L^2([0,T];V))$,
(ii) $A(u_n) \longrightarrow g$ weakly in $L^2(\Omega;L^2([0,T];V'))$,
(iii) $B(u_n)\circ\sqrt{Q} \longrightarrow \xi\circ\sqrt{Q}$ weakly in $L^2(\Omega;L^2([0,T];L_2(U,H)))$.

We can take the limit $n \to \infty$ in (4.8) and deduce from this

$$\langle u(t), e_k\rangle = \langle u_0, e_k\rangle + \int_0^t \langle g(s), e_k\rangle ds + \int_0^t \langle e_k, \xi\circ\sqrt{Q}\rangle dW(s)$$

It remains to prove that

LEMMA 4.10. $g = A(u)$ *and* $\xi\circ\sqrt{Q} = B(u)\circ\sqrt{Q}$.

PROOF OF LEMMA 4.8. We can rewrite the integral equation (4.8) as a finite-dimensional stochastic differential equation

$$d\tilde{u}_n(t) = \widetilde{A}(\tilde{u}_n(t))dt + \widetilde{B}(\tilde{u}_n(t))d\widetilde{W}^n(t) \tag{4.9}$$

where

- $u_n(t) = i_n(\tilde{u}_n(t)$ and
$$i_n : \mathbb{R}^n \longrightarrow V_n \subset V,\ \tilde{u} \mapsto \sum_{k=1}^{n} \tilde{u}_k e_k$$
denotes the natural embedding,
- $\widetilde{A}(\tilde{u}) = \pi_n A(i_n \tilde{u})$ and
$$\pi_n : V \longrightarrow \mathbb{R}^n,\ u \mapsto (\langle u, e_k\rangle)_{1\le k\le n}$$
denotes the natural orthogonal projection,
- $\widetilde{B}(\tilde{u}) = (B_k(i_n\tilde{u}))_{k=1,\dots,n}$,
- $\widetilde{W}^n(t) = (\langle W(t), f_l\rangle)_{l=1,\dots,n}$ for some CONS $(f_l)_l$ of U.

Clearly, $\widetilde{A}$ and $\widetilde{B}$ are continuous and

$$\begin{aligned}
2\langle \widetilde{A}(\tilde{u}_1) - \widetilde{A}(\tilde{u}_2), \tilde{u}_1 - \tilde{u}_2\rangle_{\mathbb{R}^n} &+ \|(\widetilde{B}(\tilde{u}_1) - \widetilde{B}(\tilde{u}_2)) \circ \sqrt{Q}\|^2_{\mathbb{R}^n} \\
&= 2\langle A(i_n\tilde{u}_1) - A(i_n\tilde{u}_2), i_n\tilde{u}_1 - i_n\tilde{u}_2\rangle + \|(B(i_n\tilde{u}_1) - B(i_n\tilde{u}_2)) \circ \sqrt{Q}\|^2_{L_2(U,H)} \\
&\le \lambda\|i_n(\tilde{u}_1 - \tilde{u}_2)\|^2_H \\
&= \lambda\|\tilde{u}_1 - \tilde{u}_2\|^2_{\mathbb{R}^n}
\end{aligned}$$

and similarly,

$$\begin{aligned}
2\langle \widetilde{A}(\tilde{u}), \tilde{u}\rangle_{\mathbb{R}^n} + \|\widetilde{B}(\tilde{u}) \circ \sqrt{Q}\|^2_{\mathbb{R}^n} &\le -\alpha\|i_n\tilde{u}\|^2_V + \lambda\|i_n\tilde{u}\|^2_H + \nu \\
&\le \lambda\|i_n\tilde{u}\|^2_H + \nu.
\end{aligned}$$

Now existence and uniqueness of the (probabilistically) strong solution to (4.9) (and therefore also to (4.8)) follows from standard results on solutios to stochastic differential equations (see, e.g. Wahrscheinlichkeitstheorie III). □

Proof of Lemma 4.9.

Applying Itô's formula to(4.8) yields

$$\begin{aligned}
\langle u_n(t), e_k\rangle^2_H = \langle u_0, e_k\rangle^2_H &+ 2\int_0^t \langle A(u_n(s)), e_k\rangle\langle u_n(s), e_k\rangle ds \\
&+ 2\int_0^t \langle u_n(s), e_k\rangle B_k(u_n(s)) dW^n(s) \\
&+ \sum_{l=1}^{n} \int_0^t B_k(u_n(s))(\sqrt{Q} f_l)^2 ds.
\end{aligned}$$

Summing up w.r.t. $k = 1, \dots, n$ yields

$$
\begin{aligned}
\|u_n(t)\|_H^2 = \underbrace{\sum_{k=1}^n \Big(\langle u_0, e_k\rangle^2}_{\le \|u_0\|_H^2} &+ 2\int_0^t \langle A(u_n(s)), e_k\rangle \langle u_n(s), e_k\rangle ds \\
&\underbrace{+ 2\int_0^t \langle u_n(s), e_k\rangle B_k(u_n(s)) dW^n(s)}_{2\int_0^t \langle u_n(s), B(u_n(s)) dW^n(s)\rangle} \\
&\underbrace{+ \sum_{l=1}^n \int_0^t B_k(u_n(s))(\sqrt{Q} f_l)^2 ds}_{\le \|B(u_n(s))\circ\sqrt{Q}\|^2_{L_2(U,H)}} \Big).
\end{aligned}
\tag{4.10}
$$

Indeed, recall that $B_k(u)h := \langle B(u)h, e_k\rangle$, $k = 1, 2, \dots, n$ and in the following write

$$B_{k,l}(u) := \langle B(u) f_l, e_k\rangle, k, l = 1, 2, \dots, n.$$

Then

$$B_k(u) dW^n(s) = \sum_{l=1}^n B_{k,l}(u) d \underbrace{\langle W(s), f_l\rangle}_{=:\beta_s^l}$$

and (β_s^l) are 1-dimensional Brownian motions with covariance

$$d\langle \beta_s^{l_1}, \beta_s^{l_2}\rangle = \langle Q f_{l_1}, f_{l_2}\rangle_U \, ds = \|\sqrt{Q} f_{l_1}\|^2 \delta_{l_1,l_2} \, ds,$$

since f_l are eigenvectors of Q. Consequently, the Itô corrections term for $\langle u_n(t), e_k\rangle^2$ has the form

$$
\begin{aligned}
&\sum_{l_1,l_2} \int_0^t B_{k,l_1}(u_n(s)) B_{k,l_2}(u_n(s)) \|\sqrt{Q} f_{l_1}\|^2 \delta_{l_1,l_2} \, ds \\
&= \sum_{l=1}^n \int_0^t B_{k,l}(u_n(s))^2 \|\sqrt{Q} f_l\|^2 \, ds.
\end{aligned}
$$

Summing up w.r.t. $k = 1, 2, \dots, n$ yields the Itô correction term for $\|u_n(t)\|^2$:

$$
\begin{aligned}
&\sum_{k=1}^n \sum_{l=1}^n \int_0^t B_{k,l}(u_n(s))^2 \|\sqrt{Q} f_l\|^2 \, ds \\
&= \sum_{l=1}^n \int_0^t \underbrace{\|B^n(u_n(s)) f_l\|_H^2 \|\sqrt{Q} f_l\|^2}_{=\|B^n(u_n(s))\circ\sqrt{Q} f_l\|_H^2} \, ds \\
&\le \int_0^t \|B^n(u_n(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} \, ds \\
&\le \int_0^t \|B(u_n(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} \, ds.
\end{aligned}
$$

Taking expectations and using (V.1) we obtain that

$$\mathbb{E}(\|u_n(t)\|_H^2) \le \|u_0\|_H^2 - \alpha \int_0^t \mathbb{E}(\|u_n(s)\|_V^2) ds + \lambda \int_0^t \mathbb{E}(\|u_n(s)\|_H^2) ds + \nu t$$

$$\le \|u_0\|_H^2 + \nu T + \lambda \int_0^t \mathbb{E}(\|u_n(s)\|_H^2) ds.$$

Gronwall's lemma now implies that

$$\mathbb{E}(\|u_n(t)\|_H^2) \le (\|u_0\|_H^2 + \nu T) e^{\lambda t} \le (\|u_0\|_H^2 + \nu T) e^{\lambda T}.$$

It follows that

$$\alpha \int_0^T \mathbb{E}(\|u_n(s)\|_V^2) ds \le \|u_0\|_H^2 + \nu T + \lambda \int_0^T \mathbb{E}(\|u_n(s)\|_H^2) ds$$

$$\le \|u_0\|_H^2 + c_{T,\|u_0\|_H,\lambda,\nu}.$$

The upper bound on the right hand side is independent of n which proves the boundedness of the second term.

In order to control the first term $\mathbb{E}\left(\sup_{t\in[0,T]} \|u_n(t)\|_H^2\right)$ we have to derive an upper bound of

$$\sup_{t\in[0,T]} e^{-\lambda t}\|u_n(t)\|_H^2 \le \|u_0\|_H^2 + \sup_{t\in[0,T]} e^{-\lambda t}\nu t + \sup_{t\in[0,T]} \int_0^t e^{-\lambda s}\langle u_n(s), B(u_n(s)) dW^n(s)\rangle.$$

But, the Burkholder-Davis-Gundy inequality implies

$$\mathbb{E}\left(\sup_{t\in[0,T]} \int_0^t e^{-\lambda s}\langle u_n(s), B(u_n(s)) dW^n(s)\rangle\right)$$

$$\le \text{const. } \mathbb{E}\left(\left(\int_0^T \sum_{l=1}^n e^{-2\lambda s}\langle u_n(s), B(u_n(s)) \circ \sqrt{Q} f_l\rangle^2 ds\right)^{\frac{1}{2}}\right)$$

$$\overset{\text{C.S.}}{\le} \text{const. } \mathbb{E}\left(\sup_{t\in[0,T]} e^{-\frac{\lambda t}{2}}\|u_n(t)\|_H \left(\int_0^T e^{\lambda(T-s)}\|B(u_n(s)) \circ \sqrt{Q}\|_{L_2(U,H)}^2 ds\right)^{\frac{1}{2}}\right)$$

$$\overset{\text{Young}}{\le} \frac{1}{2}\mathbb{E}\left(\sup_{t\in[0,T]} e^{-\lambda t}\|u_n(t)\|_H^2\right) + \frac{\text{const.}^2}{2} \underbrace{\mathbb{E}\left(\int_0^T e^{\lambda(T-s)}\|B(u_n(s)) \circ \sqrt{Q}\|_{L_2(U,H)}^2 ds\right)}_{\overset{(V.3)}{\le} \mathbb{E}(1+\int_0^T \|u_n(s)\|_V^2 ds}$$

Hence

$$\mathbb{E}\left(\sup_{t\in[0,T]} \|u_n(t)\|_H^2\right) \le \text{const.}$$

where the constant is independent of n! This proves the boundedness of the first term, hence the assertion. □

Proof of Lemma 4.10.

We apply again the "monotonicity-trick". To this end assume w.l.o.g. that (V.2) holds with $\lambda = 0$ (otherwise consider $A(u) - \lambda u$). Then

$$\int_0^T 2\langle A(u_n(s) - A(v(s)), u_n(s) - v(s)\rangle + \|(B(u_n(s)) - B(v(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds \leq 0$$

for all $v \in L^2([0,T];V)$. Weak convergence implies that

- $\int_0^T \langle A(u_n), v\rangle ds \longrightarrow \int_0^T \langle g, v\rangle ds$ for all $v \in L^2([0,T];V)$,
- $\int_0^T \langle A(v, u_n\rangle ds \longrightarrow \int_0^T \langle A(v), u\rangle ds$ for all $v \in L^2([0,T];V)$,
- $\int_0^T \langle B(u_n) \circ \sqrt{Q}, B(v) \circ \sqrt{Q}\rangle_{L_2(U,H)} ds \to \int_0^T \langle \xi \circ \sqrt{Q}, B(v) \circ \sqrt{Q}\rangle_{L_2(U,H)} ds$ for all $v \in L^2([0,T];V)$

weakly in $L^2(\Omega)$. Suppose we have in addition that

(4.11)
$$\mathbb{E}\left(\int_0^T 2\langle g(s), u(s)\rangle + \|\xi \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right) \leq \liminf_{n\to\infty} \mathbb{E}\left(\int_0^T 2\langle A(u_n(s)), e_n(s)\rangle ds + \int_0^T \|B(u_n(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right).$$

It then follows for all $v \in L^2(\Omega; L^2([0,T];V))$ that

$$\mathbb{E}\left(\int_0^T 2\langle g(s) - A(v(s)), u(s) - v(s)\rangle + \|(\xi - B(v(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right) \leq 0,$$

since

$$\mathbb{E}\left(\int_0^T 2\langle g(s) - A(v(s)), u(s) - v(s)\rangle + \|(\xi - B(v(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right) \leq \liminf_{n\to\infty} \mathbb{E}\left(\int_0^T 2\langle A(u_n) - A(v), u_n - v\rangle ds + \int_0^T \|(B(u_n) - B(u)) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right) \overset{(V.2)}{\leq} 0.$$

Choosing first $v = u$ we obtain that

$$\mathbb{E}\left(\int_0^T \|\xi \circ \sqrt{Q} - B(u) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right) \leq 0.$$

Since $\|\cdot\|_{L_2(U,H)} \geq 0$ this implies

$$\xi \circ \sqrt{Q} = B(u) \circ \sqrt{Q} \ \mathbb{P} - a.s.$$

But then

$$\mathbb{E}\left(\int_0^T \langle g(s) - A(v(s)), (u-v)(s)\rangle \, ds\right) \le 0$$

for all $v \in L^2(\Omega; L^2([0,T]; V))$. Choosing this time $v = u - \theta w$, with $\theta > 0$ and $w \in L^2(\Omega; L^2([0,T]; V))$ arbitrary, yields

$$\theta\mathbb{E}\left(\int_0^T \langle g(s) - A(u(s) - \theta w(s)), w(s)\rangle ds\right) \le 0.$$

Since A is hemi-continuous (due to (V.4)) we have that $\theta \mapsto \langle A(u-\theta w), w\rangle$ is continuous. Hence taking the limit $\theta \searrow 0$ implies that

$$\mathbb{E}\left(\int_0^T \langle g(s) - A(u(s)), w(s)\rangle ds\right) \le 0.$$

Since w was arbitrary, we conclude that $g(s) = A(u(s))$ a.s.

It remains to establish inequality (4.10). To this end note that (4.9) implies that

$$\mathbb{E}\left(\int_0^T 2\langle A(u_n(s)), u_n(s)\rangle + \|B(u_n(s)) \circ \sqrt{Q}\|^2_{L_2(U,H)} ds\right)$$
$$\ge \mathbb{E}\left(\|u_n(T)\|_H^2\right) - \|u_0\|_H^2.$$

Hence (4.10) follows from the inequality

$$\liminf_{n\to\infty} \mathbb{E}\left(\|u_n(T)\|_H^2\right) \ge \mathbb{E}\left(\|u(T)\|_H^2\right),$$

which is a consequence of the weak convergence of $u_n(T) \longrightarrow u(T)$ in $L^2(\Omega; H)$ along some subsequence. □

Example 4.11 (stochastic heat equation with multiplicative noise). Consider the stochastic bilinear heat equation

$$\partial_t u(t,x) = \frac{1}{2}\partial_{xx} u(t,x) + \theta \partial_x u(t,x)\partial_t W_t, \qquad u(0,x) = u_0(x)$$

with $\theta \in \mathbb{R}$. In this case

(1) $A(u) = \frac{1}{2}u_{xx} : H^1(\mathbb{R}) \subset L^2(\mathbb{R}) \simeq L^2(\mathbb{R})' \subset (H^1(\mathbb{R}))'$ and

$${}_{(H^1)'}\langle A(u), u\rangle_{H^1} = \frac{1}{2}\int u_{xx} u dx = -\frac{1}{2}\int u_x u_x dx = -\frac{1}{2}\int u_x^2 dx = -\frac{1}{2}\|u\|_{H^1}^2 + \frac{1}{2}\|u\|_H^2$$

(2) $B(u)h = \partial_x u h$ with

$$\|B(u) \circ \sqrt{Q}\|^2_{L_2(U,H)} = \|\partial_x u \circ \sqrt{Q}\|^2_{L_2(U,H)} \overset{U=\mathbb{R}}{=} \|\partial_x u\|^2_{L^2(\mathbb{R})},$$

(3) (W_t) is a 1-dimensional Wiener process (hence standard Brownian motion), i.e. $U = \mathbb{R}$.

In this case the coercivity assumption (V.1)

$$2\langle A(u), u\rangle + \|B(u) \circ \sqrt{Q}\|^2_{L_2(U,H)} = (\theta^2 - 1)\|\partial_x u\|^2_{L^2(R)}$$

is satisfied if and only if $|\theta| < 1$.

In the case $\theta = 1$ we have that

$$u(t,x) := u_0(x + W_t)$$

is an explicit solution, in the case $\theta = -1$ we can consider $u(t,x) = u_0(x - W_t)$. Indeed, an applications of Ito's formula yields for $u_0 \in C^2(\mathbb{R})$

$$\begin{aligned} u_0(x + W_t) &= u_0(x) + \int_0^t \partial_x u_0(x + W_s) dW_s + \frac{1}{2} \int_0^t \partial_{xx} u_0(x + W_s) ds \\ &= u_0(x) + \int_0^t \partial_x u(s,x) dW_s + \frac{1}{2} \int_0^t u_{xx}(s,x) ds \end{aligned}$$

It is clear that $u(t,x) = u_0(x + W_t)$ has the same spatial regularity as u_0, hence there is no regularizing effect in the case and the equation becomes hyperbolic.

In the case $\theta = 0$ the spde reduces to the deterministic heat equation with explicit solution

$$u(t,x) = \frac{1}{\sqrt{2\pi t}} \int e^{-\frac{(x-y)^2}{2t}} u_0(y) dy$$

In the case $|\theta| > 1$ we obtain that

$$2\langle A(u), u\rangle + \|B(u) \circ \sqrt{Q}\|^2 = -(1 - \theta^2)\|\partial_x u\|_{L^2}^2$$

which corresponds to a backward heat equation which is no longer well-posed.

Can one detect these effects in numerical simulations?

Lecture 10

Variational approach to stochastic partial differential equations

4.4. Compactness method for solving SPDEs in the variational setting

We keep the assumptions (V.1) and (V.3), i.e.

(V.1) $\exists\, \alpha > 0,\ \lambda, \nu$ such that

$$2\langle A(u), u\rangle + \|B(u) \circ \sqrt{Q}\|_{L_2(U,H)} \le -\alpha \|u\|_V^2 + \lambda \|u\|_H^2 + \nu$$

(V.3) $\exists M$ such that

$$\|A(u)\|_{V'} \le M(1 + \|u\|_V).$$

This time we assume in addition

(V.5) $\exists M, \delta > 0$, such that

$$\|B(u) \circ \sqrt{Q}\|_{L_2(U,H)} \le M(1 + \|u\|_V^{1-\delta})$$

corresponding to sublinear growth since $\delta > 0$.

(V.6) $V \hookrightarrow H$ ist compact, i.e., bounded subsets of V are precompact in H.

(V.7) $u \mapsto A(u) : V_{\text{weak}} \cap H \longrightarrow V'_{\text{weak}}$ is continuous,
$u \mapsto B(u) : V_{\text{weak}} \cap H \longrightarrow L_2(U, H)$ is continuous.

Definition 4.12. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is called a solution of the martingale problem associated with the SPDE (4.7)

$$du(t) = A(u(t))dt + B(u(t))dW(t), \quad u(0) = u_0$$

if

(i) $\mathbb{P}(u(0) = u_0) = 1$,

(ii) the stochastic process

$$M_t := u(t) - u(0) - \int_0^t A(u(s))ds$$

is a continuous H-valued $\mathbb{P}$-martingale with associated increasing process

$$\langle M\rangle_t = \int_0^t B(u(s)) \circ Q \circ B(u(s))^* ds.$$

The second condition in the above definition is equivalent to the following condition:

(ii') Let (e_n) be a CONS of H consisting of elements $e_k \in V$. Then for all $i \ge 1$, $\varphi \in C_b^2(\mathbb{R})$, $0 \le s \le t$ and any Φ_s continuous, bounded and $\mathcal{F}_s$-measurable:

$$\mathbb{E}\left[(M_t^{i,\varphi} - M_s^{i,\varphi})\Phi_s\right] = 0,$$

where

$$
\begin{aligned}
M_t^{i,\varphi} := \varphi(\langle u(t), e_i\rangle_H) - \varphi(\langle u_0, e_i\rangle) \\
- \left(\int_0^t \varphi'(\langle u(s), e_i\rangle)\langle A(u(s)), e_i\rangle ds \right. \\
\left. + \frac{1}{2}\int_0^t \varphi''(\langle u(s), e_i\rangle\langle (B\circ Q\circ B^*)(u(s))e_i, e_i\rangle ds \right).
\end{aligned}
$$

Theorem 4.13. *Under the assumptions (V.1), (V.3) and (V.5)-(V.8) there exist a solution $\mathbb{P}$ to the martingale problem associated with* (4.7).

Proof. We consider the same Galerkin approximation as in the proof of 4.5: $V_n := \text{span}\{e_1, ..., e_n\}$ for some CONS (e_n) of H consisting of elements $e_k \in V$. Let Π_n be the orthogonal projection in H onto V_n. Then there exists for all $n \geq 1$ a solution $\mathbb{P}_n$ on $(\Omega, \mathcal{F})$, such that

$(0)_n$ $\text{supp}(\mathbb{P}_n) \subset C([0,T]; V_n)$

$(i)_n$ $\mathbb{P}_n(u(0) = \Pi_n u_0) = 1$

$(ii)_n$ $\forall \varphi \in C_b^2(\mathbb{R})$ and Φ_s continuous, bounded and $\mathcal{F}_s$-measurable

$$\mathbb{E}_n\left[(M_t^{i,\varphi} - M_s^{i,\varphi})\Phi_s\right] = 0.$$

$\mathbb{P}_n$ is obtained as the law of the corresponding finite-dimensional stochastic differential equation.

Suppose now the following lemma holds:

Lemma 4.14. *The sequence of probability measures $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is tight on Ω.*

Continuing the proof of Theorem 4.13

Lemma 4.14 implies that we can extract a subsequence, again denoted with $(\mathbb{P}_n)_n$, such that $\mathbb{P}_n \overset{w}{\Rightarrow} \mathbb{P}$ weakly on $C([0,T];V) \subset \underbrace{C([0,T];H)}_{=\Omega}$. The limiting probability measure $\mathbb{P}$ satisfies condition (i) and the mapping

$$\Omega \longrightarrow \mathbb{R}, \quad \omega \mapsto (M_t^{i,\varphi}(\omega) - M_s^{i,\varphi}(\omega))\Phi_s(\omega)$$

is continuous on Ω.

The coercivity assumption (V.1) together with Lemma 4.9 implies that

$$\sup_{n\in\mathbb{N}} \mathbb{E}_n\left(\sup_{t\in[0,T]} \|u(t)\|^2 + \int_0^T \|u(t)\|_V^2 ds \right) < \infty.$$

In particular, we obtain for some $p > 1$ (depending on δ in (V.5))

$$\sup_{n\in\mathbb{N}} \mathbb{E}_n\left(|M_t^{i,\varphi} - M_s^{i,\varphi}|^p\right) < \infty,$$

since,

$$M_t^{i,\varphi} - M_s^{i,\varphi} = \underbrace{\varphi(\langle u(t), e_i\rangle) - \varphi(\langle u(s), e_i\rangle)}_{\text{bounded}}$$
$$+ \int_s^t \underbrace{\varphi'(\langle u(r), e_i\rangle)}_{\text{bounded}} \underbrace{\langle A(u(r)), e_i\rangle}_{\overset{(V.3)}{\leq} M(1+\|u(r)\|_V)\|e_i\|_V} dr$$
$$+ \frac{1}{2}\int_s^t \underbrace{\varphi''(\langle u(r), e_i\rangle)}_{\text{bounded}} \underbrace{\langle (B\circ Q\circ B^*)(u(r))e_i, e_i\rangle}_{\overset{(V.5)}{\leq} M\left(1+\|u(r)\|_V^{2(1-\delta)}\|e_i\|_V^2\right)} dr,$$

so that

$$\left|M_t^{i,\varphi} - M_s^{i,\varphi}\right| \underset{\delta\leq\frac{1}{2}}{\overset{\text{w.l.o.g.}}{\leq}} M_{s,t,\varphi,\|e_i\|_V,T}\left(1+\int_0^T \|u(s)\|_V^2 ds\right)^{1-\delta}$$

and we can choose $p = \frac{1}{1-\delta} > 1$.

It follows that

$$\lim_{n\to\infty} \mathbb{E}_n\left((M_t^{i,\varphi} - M_s^{i,\varphi})\Phi_s\right) = \mathbb{E}\left((M_t^{i,\varphi} - M_s^{i,\varphi})\Phi_s\right) = 0,$$

which implies (ii). Indeed, for all $K > 0$,

$$\Psi_K := (-K) \vee \left(M_t^{i,\varphi} - M_s^{i,\varphi}\right) \wedge K$$

is bounded and continuous. Furthermore,

$$\begin{aligned}
&\left|\mathbb{E}_n[(\Psi_K - (M_t^{i,\varphi} - M_s^{i,\varphi}))\Phi_s]\right| \\
&\qquad \leq \|\Phi_s\|_\infty \mathbb{E}_n\left(|M_t^{i,\varphi} - M_s^{i,\varphi}| 1_{\{|M_t^{i,\varphi} - M_s^{i,\varphi}|\geq K\}}\right) \\
&\qquad \leq \|\Phi_s\|_\infty \frac{1}{K^{p-1}} \mathbb{E}_n\left(|M_t^{i,\varphi} - M_s^{i,\varphi}|^p\right) \\
&\qquad \underset{\text{for } K\nearrow\infty}{\overset{\text{unif. in } n}{\longrightarrow}} 0.
\end{aligned}$$

Consequently for all K:

$$\begin{aligned}
&\lim_{n\to\infty} \left|\mathbb{E}_n[(M_t^{i,\varphi} - M_s^{i,\varphi})\Phi_s] - \mathbb{E}[(M_t^{i,\varphi} - M_s^{i,\varphi})\Phi_s]\right| \\
&\quad \leq \|\Phi_s\|_\infty \limsup_{n\to\infty} \left(\mathbb{E}_n(|(M_t^{i,\varphi} - M_s^{i,\varphi}) - \Psi_K|) + \mathbb{E}(|(M_t^{i,\varphi} - M_s^{i,\varphi}) - \Psi_K|)\right) \\
&\overset{K\nearrow\infty}{\longrightarrow} 0.
\end{aligned}$$

□

It remains to prove Lemma 4.14.

Proof of Lemma 4.14. Let

- τ_1 be the weak topology on $L^2([0,T];V)$,
- τ_2 be the uniform topology on $C([0,T];V')$,
- τ_3 be the strong topology on $L^2([0,T];H)$.

We will show that $(\mathbb{P}_n)_n$ is tight w.r.t. all three topologies. It will then follow that there exists some subsequence, again denoted with $(\mathbb{P}_n)_n$, and a probability measure $\mathbb{P}$ on $C([0,T];V') \cap L^2([0,T];V)$ such that

$$\mathbb{E}_n(\Phi(u)) = \lim_{n\to\infty} \int \Phi \, d\mathbb{P}_n = \int \Phi \, d\mathbb{P}$$

for all bounded Φ that are continuous w.r.t. τ_i, for some $i = 1, 2, 3$.

For given l and k let

$$K := \{u \in C([0,T];V') \cap L^2([0,T];V) \mid \sup_{t\in[0,T]} \|u(t)\|_H \le l, \int_0^T \|u(t)\|_V^2 dt \le k\}.$$

The a priori moment estimates on the Galerkin approximation imply that $\mathbb{P}_n(K^c)$ can be made arbitrarily small for large l and k uniformly in $n \in \mathbb{N}$.

1. τ_1-tightness This tightness follows from boundedness of $K \subset L^2([0,T];V)$ and the fact that bounded subsets in separable Hilbert spaces (like $L^2([0,T];V)$) are relatively weakly compact (Banach-Alaoglu).

2. τ_2-tightness We have to show that for all $h \in V$ with $\|h\|_V = 1$, the set of functions

$$\{t \mapsto \langle u(t), h\rangle : u \in K\}$$

is a compact subset of $C([0,T])$.

But this follow from Arzela-Ascoli's theorem, since

(i) $\sup_{\substack{u\in K\\ \|h\|_V\le 1}} |\langle u(t), h\rangle_H| < \infty$

(ii) $\sup_{\substack{u\in K\\ \|h\|_V\le 1}} \sup_{|t-s|\le\delta} |\langle u(t), h\rangle - \langle u(s), h\rangle| \longrightarrow 0$ for $\delta \searrow 0$ since, w.r.t. $\mathbb{P}_n$,

$$\begin{aligned} |\langle u(t), h\rangle - \langle u(s), h\rangle| &= \left| \int_s^t \langle A(u(r), h\rangle dr + \int_s^t \langle h, B^n(u(r)) dW^n(r)\rangle \right| \\ &\le M\left(1 + \int_s^t \|u(r)\|_V dr\right) + \underbrace{\left| \int_s^t \langle h, B^n(u(r)) dW^n(r)\rangle \right|}_{(*)} \end{aligned}$$

In order to control $(\star)$, consider

$$
\begin{aligned}
\mathbb{E}_n&\left(\sup_{r\in[s,t]}\left|\int_s^r \langle h, B^n(u(r))dW^n(r)\rangle\right|^p\right)\\
&= \mathbb{E}_n\left(\sup_{r\in[s,t]}\left|\langle h, \int_s^r B^n(u(r))dW^n(r)\rangle\right|^p\right)\\
\overset{\text{Burkholder-Davis-Gundy}}&{\leq} c_p\mathbb{E}_n\left(\left(\int_s^t \|h\|_H^2 \underbrace{\|B(u(r))\circ\sqrt{Q}\|^2_{L_2(U,H)}}_{\leq M(1+\|u(r)\|_V^{1-\delta})^2} dr\right)^{\frac{p}{2}}\right)\\
&\leq 2^{\frac{p}{2}} c_p M^p \|h\|_H^p \mathbb{E}_n\left(\left((t-s)+\int_s^t \|u_r\|_V^{2(1-\delta)}\,dr\right)^{\frac{p}{2}}\right)\\
&\leq 2^{\frac{p}{2}+1} c_p M^p \|h\|_H^p\left((t-s)^{\frac{p}{2}} + (t-s)^{\frac{p}{2}\delta}\mathbb{E}_n\left(\int_s^t \|u_r\|_V^2\,dr\right)^{\frac{p}{2}(1-\delta)}\right)\\
&\leq 2^{\frac{p}{2}+1} c_p M^p \|h\|_H^p\left((t-s)^{\frac{p}{2}} + \sup_n\left(\mathbb{E}_n\left(\int_s^t \|u_r\|_V^2\,dr\right)\right)^{\frac{2}{p(1-\delta)}}(t-s)^{\frac{p}{2}\delta}\right)\\
&\overset{|t-s|\searrow 0}{\longrightarrow} 0 \text{ uniformly in } \|h\|_V = 1.
\end{aligned}
$$

3. τ_3**-tightness** This will follow from the next lemma. □

LEMMA 4.15. *Given a sequence* $(u_n)_n$*, bounded in* $L^2([0,T];V)\cap L^\infty([0,T];H)$ *and equicontinuous as* V'*-valued functions, such that* $u_n(0)\longrightarrow u_0$ *strongly in* H*, one can extract a subsequence strongly convergent in* $L^2([0,T];H)$*.*

PROOF. We will use the following fact: $\forall\varepsilon>0\ \exists c_\varepsilon$ such that for all $u\in V$:

$$\|u\|_H \leq \varepsilon\|u\|_V + c_\varepsilon\|u\|_{V'}.$$

Indeed, otherwise we could find $\varepsilon>0$ and a sequence $(u_n)_n$ such that

$$\|u_n\|_H \geq \varepsilon\|u_n\|_V + n\|u_n\|_{V'}.$$

But then, using $\tilde{u}_n := \frac{u_n}{\|u_n\|_H}$, and

$$1 = \|\tilde{u}_n\|_H \geq \varepsilon\|\tilde{u}_n\|_V + n\|\tilde{u}_n\|_{V'},$$

we can extract a weakly convergent subsequence $\tilde{u}_n \to \tilde{u}\in V$ for some $\tilde{u}\in V$, hence strongly convergent in H. Since

$$\|\tilde{u}\|_{V'} = \lim_{n\to\infty}\|\tilde{u}_n\|_{V'} \leq \frac{1}{n} \overset{n\to\infty}{\longrightarrow} 0,$$

we obtain that $\tilde{u}=0$, which gives a contradiction, since $\tilde{u}\neq 0$.

We know that there exists a subsequence $(u_n)_n$ converging to some $u\in C([0,T];V')$, hence also in $L^2([0,T];V')$. Clearly, $u\in L^2([0,T];V)$, hence for all $\varepsilon>0$ there exists some c_ε such that

$$\begin{aligned}\limsup_{n\to\infty}\int_0^T \|u_n(t)-u(t)\|_H^2 dt &\le \limsup_{n\to\infty}\varepsilon\int_0^T \|u_n(t)-u(t)\|_V^2 dt \\ &\quad + \underbrace{\limsup_{n\to\infty} c_\varepsilon \int_0^T \|u_n(t)-u(t)\|_{V'}^2 dt}_{=0} \\ &\le \varepsilon T \sup_{t\in[0,T]} \|u_n(t)-u(t)\|_V^2.\end{aligned}$$

$\square$

Lecture 11

Variational approach to stochastic partial differential equations

4.5. Generalization of Theorem 4.5

We again consider the Gelfand triple $V \hookrightarrow H \cong H' \hookrightarrow V'$ for separable real Hilbert spaces. Let us now formulate generalized conditions given in the monograph [Prevot/Röckner: A concise course on SPDE, Springer] such that the main result, Theorem 4.5, on existence and uniqueness of a variational solution to (4.7)

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

still holds true.

Again, let $(W(t))$ be some Q-Wiener process on a separable real Hilbert space U and suppose that

- $A = A(t, u, \omega) : [0, T] \times V \times \Omega \longrightarrow V'$
- $B = B(t, u, \omega) : [0, T] \times V \times \Omega \longrightarrow L(U, H)$

are progressively measurable. In the following we will suppress the ω-dependence, just writing $A(t, u)$ und $B(t, u)$.

The generalizations of (V.1)-(V.4) are then given as follows:

(V.1) $\exists\, \alpha > 0, p \in]1, \infty[, \lambda \in \mathbb{R}$ und an adapted process $f \in L^1([0, T \times \Omega, dt \otimes \mathbb{P})$, such that

$$2\langle A(t, u), u\rangle + \|B(t, u) \circ \sqrt{Q}\|^2_{L_2(U,H)} \leq -\alpha\|u\|^p_V + \lambda\|u\|^2_H + f(t),$$

for all $u \in V$.

(V.2) $\exists\, \lambda \in \mathbb{R}$, such that

$$2\langle A(t, u_1) - A(t, u_2), u_1 - u_2\rangle + \|(B(t, u_1) - B(t, u_2) \circ \sqrt{Q}\|^2_{L_{2(U,H)}} \leq \lambda\|u_1 - u_2\|^2_H,$$

for all $u_1, u_2 \in V$.

(V.3) $\exists M > 0$ and an adapted process $g \in L^{\frac{p}{p-1}}([0, T] \times \Omega, dt \otimes \mathbb{P})$, with the same p as in (V.1), such that

$$\|A(t, u)\|_{V'} \leq g(t) + M\|u\|^{p-1}_V$$

for all $u \in V$.

(V.4) $\forall u, v, w \in V$ and $\forall \omega \in \Omega$ the mapping

$$\lambda \mapsto \langle A(t, u + \lambda v, \omega), w\rangle$$

is continuous.

We then have the following analogous result:

Theorem 4.16. *Under the generalized assumptions (V.1)-(V.4) there exists for all $u_0 \in H$ a unique H-valued contiuous adapted process $u(t), t \in [0, T]$, such that $u \in L^p([0, T] \times \Omega; V) \cap L^2([0, T] \times \Omega; H)$ satisfying the equation*

$$\langle u(t), v\rangle = \langle u_0, v\rangle + \int_0^t \langle A(s, u(s)), v\rangle ds + \int_0^t \langle v, B(s, u(s))dW(s)\rangle.$$

The proof is similar to the proof of Theorem 4.5. It requires the following generalization of Itô's formula:

LEMMA 4.17. *Let $u_0 \in H$ and u (resp. v) be an adapted process with $u \in L^p([0,T]\times\Omega;V)$ (resp. $v \in L^{\frac{p}{p-1}}([0,T]\times\Omega;V')$) and let M be a continuous local martingale with $\mathbb{E}(\langle M\rangle_T)<\infty$ (hence L^2-bounded) such that*

$$u(t) = u_0 + \int_0^t v(s)ds + M_t.$$

Then:

(i) $u \in C([0,T];H)$ *a.s. and* $\mathbb{E}\left(\sup_{t\in[0,T]}\|u(t)\|_H^2\right)<\infty$

(ii) $\|u(t)\|_H^2 = \|u_0\|_H^2 + 2\int_0^t {}_{V'}\langle v(s),u(s)\rangle_V ds + 2\int_0^t \langle u(s), dM_s\rangle + \langle M\rangle_t$

Remark Note that

$$\int_0^t |\langle v(s),u(s)\rangle| ds \le \int_0^t \|v(s)\|_{V'}\|u(s)\|_V ds$$
$$\overset{\text{Hölder}}{\le} \left(\int_0^t \|v(s)\|_{V'}^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \left(\int_0^t \|u(s)\|_V^p ds\right)^{\frac{1}{p}} < \infty$$

is integrable, hence $\int_0^t \langle v(s), u(s)ds\rangle$ well-defined.

The novelty in this Lemma is the additional integrability $\mathbb{E}\left(\sup_{t\in[0,T]}\|u(t)\|_H^2\right)<\infty$, that can be proven as follows:

First note that

$$\|u(t)\|_H^2 = \|u(s)\|_H^2 + 2\int_s^t \langle v(r),u(r)\rangle ds + 2\langle u(s), M(t)-M(s)\rangle$$
$$+\|M(t)-M(s)\|_H^2 - \|u(t)-u(s)-(M(t)-M(s))\|_H^2$$

which implies for $s \le t$

$$\|u(t)\|_H^2 \le 2\|u(s)\|_H^2 + 2\left(\int_0^T \|u(t)\|_V^p ds\right)^{\frac{1}{p}}\left(\int_0^T \|v(t)\|_{V'}^{\frac{p}{p-1}} dt\right)^{\frac{p-1}{p}} + 2\|M(t)-M(s)\|_H^2.$$

Hence

$$\mathbb{E}\left(\sup_{t\in[0,T]}\|u(t)\|_H^2\right)\le 2\|u_0\|_H^2+2\mathbb{E}\left(\int_0^T\|u(t)\|_V^p dt\right)^{\frac{1}{p}}\mathbb{E}\left(\int_0^T\|v(t)\|_{V'}^{\frac{p}{p-1}}dt\right)^{\frac{p-1}{p}}$$

$$+\underbrace{2\mathbb{E}\left(\sup_{t\in[0,T]}\|M(t)\|_H^2\right)}_{\overset{\text{max.}}{\underset{\text{inequ.}}{\le}}\underbrace{2\left(\frac{2}{2-1}\right)^2}_{=8}\mathbb{E}\left(\|M(T)\|_H^2\right)}<\infty.$$

Proof of Theorem 4.16. The proof of uniqueness is exactly the same as in Theorem 4.5. For the proof of existence we need the following a prior estimate on Galerkin approximations: □

Lemma 4.18.

$$\sup_{n\in\mathbb{N}}\mathbb{E}\left(\sup_{t\in[0,T]}\|u_n(t)\|_H^2+\int_0^T\|u_n(s)\|_V^p ds\right)<\infty.$$

Proof. Similar to the proof of 4.9, we apply Itô's formula to obtain

$$\begin{aligned}\langle u_n(t),e_k\rangle^2=\langle u_0,e_k\rangle^2&+\int_0^t\langle A(s,u_n(s)),e_k\rangle\langle u_n(s),e_k\rangle ds\\&+2\int_0^t\langle u_n(s),e_k\rangle B_k(s,u_n(s))dW^n(s)\\&+\sum_{l=1}^n B_k(s,u_n(s))\circ(\sqrt{Q}f_l)^2 ds\end{aligned}$$

and thus

$$\begin{aligned}e^{-\lambda t}\|u_n(t)\|\le\|u_0\|_H^2&+2\int_0^t e^{-\lambda s}\left(\langle A(s,u_n(s)),u_n(s)\rangle-\lambda\|u_n(s)\|_H^2\right)ds\\&+\int_0^t e^{-\lambda s}\langle u_n(s),B(s,u_n(s))dW^n(s)\rangle\\&+\int_0^t e^{-\lambda s}\|B(s,u_n(s))\circ\sqrt{Q}\|_{L_2(U,H)}^2 ds\\\le\|u_0\|_H^2-\alpha\int_0^t e^{-\lambda s}\|u_n(s)\|_V^p ds&+\int_0^t e^{-\lambda s}f(s)ds\\&+2\int_0^t e^{-\lambda s}\langle u_n(s),B(s,u_n(s))W^n(s)\rangle.\end{aligned}$$

Taking expectations we first obtain that

$$e^{-\lambda t}\mathbb{E}\left(\|u_n(s)\|_H^2\right) + \alpha \int_0^t e^{-\lambda s}\mathbb{E}\left(\|u_n(s)\|_V^p\right) ds \leq \|u_0\|_H^2 + \underbrace{\mathbb{E}\left(\int_0^t e^{-\lambda s} f(s) ds\right)}_{<\infty},$$

hence

$$\mathbb{E}\left(\int_0^T \|u_n(s)\|_V^p ds\right) \leq \frac{e^{\lambda T}}{\alpha}\left(\|u_0\|_H^2 + \mathbb{E}\left(\int_0^T |f(s)| ds\right)\right) < \infty$$

uniformly in n. This proves the upper bound for the second term. For the first term we

$$\sup_{s\in[0,t]} e^{\lambda s}\|u_n(s)\|_H^2 \leq \|u_0\|_H^2 + \int_0^t e^{-\lambda s}|f(s)|ds + 2\sup_{s\in[0,t]}\int_0^s e^{-\lambda s}\langle u_n(r), B(r, u_n(r)) dW^n(r)\rangle$$

and using the Burkholder-Davis-Gundy inequality

$$\begin{aligned}\mathbb{E}\left(\sup_{s\in[0,t]} e^{-\lambda s}\|u_n(s)\|_H^2\right) \leq \|u_0\|_H^2 &+ \mathbb{E}\left(\int_0^t e^{-\lambda s}|f(s)|ds\right)\\ &+ \frac{1}{2}\mathbb{E}\left(\sup_{s\in[0,t]} e^{-\lambda s}\|u_n(s)\|_H^2\right)\\ &+ \frac{\text{const.}^2}{2}\underbrace{\mathbb{E}\left(\int_0^t e^{-\lambda s}\|B(s,u_n(s))\circ\sqrt{Q}\|_{L_2(U,H)}^2 ds\right)}_{\overset{(V.3)}{\leq} \int_0^t e^{-\lambda s} 0(\lambda\|u_n(s)\|_H^2 + f(s) + g(s) + M\|u_n(s)\|_V^p) ds}\end{aligned}$$

hence

$$\begin{aligned}\mathbb{E}\left(\sup_{s\in[0,t]} e^{-\lambda s}\|u_n(s)\|_H^2\right) \leq 2\|u_0\|_H^2 &+ 2\mathbb{E}\left(\int_0^T e^{-\lambda s}(2|f(s)| + |g(s)|) ds\right)\\ &+ 2M\mathbb{E}\left(\int_0^T \|u_n(s)\|_V^p ds\right)\\ &+ \lambda\int_0^t \mathbb{E}\left(\sup_{r\in[0,s]} e^{-\lambda r}\|u_n(r)\|_H^2\right) ds,\end{aligned}$$

which yields the upper bound for the first term, using Gronwall's lemma. □

EXAMPLE 4.19. Consider the following SPDE

$$\partial_t u(t,x) = (\partial_{xx}u(t,x) + f(t,u(t,x)))dt + \sigma(t,x,u(t,x))\frac{\partial W}{\partial t}(t,x). \tag{4.12}$$

(4.12) is called a stochastic reaction-diffusion equation since the right hand side is a superposition of diffusion generated by the Laplace operator and a (nonlinear) reaction term $f(t,u)$. We consider (4.12) on the unit interval for $x \in [0,1]$ with

Neumann boundary conditions for the Laplace operator. Furthermore we assume that σ are f continuous in (t,x,u) and satisfy:

(N.1) $f(t,x,u)u \leq -m|u|^p + \lambda u^2 + f_0(t,x)$ for all $u \in \mathbb{R}$,

(N.2) For $u, u_1, u_2 \in \mathbb{R}$ and some $\lambda_0 \in \mathbb{R}$:
- *(i)* $(f(t,x,u_1) - f(t,x,u_2))(u_1 - u_2) \leq \lambda_0 (u_1 - u_2)^2$,
- *(ii)* $(\sigma(t,x,u_1) - \sigma(t,x,u_2))^2 \leq \lambda_0 (u_1 - u_2)^2$,
- *(iii)* $\sigma^2(t,x,u) \leq \lambda_0 u^2$

(N.3) $|f(t,x,u)| \leq g(t,x) + M|u|^{p-1}$ for some $g \in L^1([0,T] \times [0,T])$, $M > 0$.

Condition (N.3) implies, that the operator $A(t,\cdot) : V \longrightarrow V'$ (with $V = H^{1,2}(0,1)$) is well-defined, since $\|u\|_\infty \leq \sqrt{2}\|u\|_V$:

$$u \in V \Rightarrow \left| u(x) - \int_0^1 u(y)dy \right| \leq \int_0^1 |u(x) - u(y)dy \leq \|\partial_x u\|_{L^2}$$
$$\Rightarrow |u(x)| \leq \|\partial_x ur\|_{L^2} + \|u\|_{L^2} \leq \sqrt{2}\|u\|_V.$$

This implies for the operator $A(t,u) = \partial_{xx} u + f(t,x,u)$ that

$$|\langle A(t,u), v\rangle| = \Big| \underbrace{\int_0^1 \partial_{xx} u\, v\, dx}_{= -\int_0^1 \partial_x u \partial_x v\, dx} + \int_0^1 f(t,x,u) v\, dx \Big|$$
$$\leq \|u\|_V \|v\|_V + \int_0^1 g(t,x)\, dx + M \underbrace{\int_0^1 |u(x)|^{p-1} |v(x)| dx}_{\leq \|u\|_V^{p-1} \|v\|_V}$$

so that

$$\|A(t,u)\|_{V'} \leq c_p + \int_0^1 g(t,x)\, dx + \widetilde{M} \|u\|_V^{p-1}$$

which implies that condition (V.3) holds.

Let $(W(t))_{t\geq 0}$ be a Q-Wiener process on H. Then using (N.2):

$$\begin{aligned}
\|(\sigma(t,x,u_1) - \sigma(t,x,u_2) \circ \sqrt{Q}\|^2_{L_2(U,H)} &= \sum_{k\geq 1} \|(\sigma(t,x,u_1) - \sigma(t,x,u_2))^2 \sqrt{Q} e_k\|^2_H \\
&\leq \sum_{k\geq 1} \lambda_0 \int (u_1 - u_2)^2(x) (\sqrt{Q} e_k)^2 dx \\
&\leq \lambda_0 \|\sqrt{Q}\|^2_{L_2(U,H)} \|u_1 - u_2\|_V.
\end{aligned}$$

Hence for $\lambda_0 \|\sqrt{Q}\|^2_{L_2(U,H)} < 1$:

$$\begin{aligned}
&2\langle A(t,u_1) - A(t,u_2), u_1 - u_2\rangle + \|(\sigma(t,u_1) - \sigma(t,u_2)) \circ \sqrt{Q}\|^2_{L_2(U,H)} \\
&\qquad \leq -2 \int \partial_x (u_1 - u_2)^2(x) dx + 2\lambda_0 \|u_1 - u_2\|^2_{L^2} + \lambda_0 \|\sqrt{Q}\|^2_{L_2(U,H)} \|u_1 - u_2\|^2_V \\
&\qquad \leq (\lambda_0 \|\sqrt{Q}\|^2_{L_2(U,H)} - 2) \|u_1 - u_2\|^2_V + 2(\lambda_0 + 1) \|u_1 - u_2\|^2_H
\end{aligned}$$

which implies that (V.2) holds with $\lambda = 2(\lambda_0 + 1)$.

Similarly,

$$
\begin{aligned}
&2\langle A(t,u),u\rangle + \|\sigma(t,u)\circ\sqrt{Q}\|^2_{L_2(U,H)} \\
&\qquad\le -2\|\partial_x u\|^2_{L^2} + 2\lambda_0\|\sqrt{Q}\|^2_{L_2(U,H)}\|u\|^2_V - m\int |u|^p dx + \lambda\int |u|^p dx + \int f_0(t,x)dx \\
&\qquad\le 2\left(\lambda_0\|\sqrt{Q}\|^2_{L_2(U,H)} - 1\right)\|u\|^2_V - m\int |u(x)|^p dx + (\lambda+2)\int u^2 dx + \int f_0(t,x)dx
\end{aligned}
$$

which implies (V.1).

Condition (V.4) is satisfied due to the joint continuity of f and σ in (t,x,u).
We can now apply Theorem 4.16 to obtain a unique variational solution of the stochastic reaction-diffusion equation

$$\partial_t u(t,x) = (\partial_{xx}u(t,x) + f(t,u(t,x)))dt + \sigma(t,x,u(t,x))\frac{\partial W}{\partial t}(t,x).$$

CHAPTER 5

Lecture 14

Stochastic reaction-diffusion equations

5.2. SPDEs and associated Markov processes

Consider the general stochastic evolution equation

$$\begin{cases} dX(t) &= [AX(t) + B(X(t))]\,dt + C(X(t))dW(t), \\ X(0) &= \xi \in H \end{cases} \tag{5.1}$$

on some separable real Hilbert space H. Here, $(W(t))_{t\geq 0}$ is a cylindrical Wiener process on a possibly different second separable real Hilbert space U (hence covariance operator Id_U), defined on some underlying probability space $(\Omega, \mathcal{F}, P)$, and adapted to some filtration $(\mathcal{F}_t)_{t\geq 0}$ that is assumed to be right-continuous and complete in the sense that $\mathcal{F}_0$ contains all P-null sets.

5.2.1. Transition semigroup and Markov property. Suppose that for all initial conditions $\xi \in H$ (5.1) has a unique mild or variational solution $X(t,\xi)$ and that the family of transition probabilities

$$x \mapsto p_t(x, A) := P\left(X(t,x) \in A\right),\, H \to \mathbb{R}$$

is measurable for all $t \geq 0$ and $A \in \mathcal{B}(H)$. In this case the family of transition probabilities defines a stochastic kernel on $(H, \mathcal{B}(H))$ with associated integral operators

$$P_t F(x) := E(F(X(t,x))),\; t \geq 0\,.$$

operating on the space $\mathcal{B}_b(H)$ of bounded Borel measurable functions $F : H \to \mathbb{R}$. The unique solution of (5.1) then satisfies the following (simple) Markov property

$$E\left(F(X(s+t,\xi) \mid \mathcal{F}_s\right) = P_t F(X(s,\xi)) \qquad P-a.s.$$

for any s, $t \geq 0$, which implies the Chapman-Kolmogorov equation $P_{s+t} = P_s \circ P_t$, i.e., the semigroup property for the family $(P_t)_{t\geq 0}$.

Definition 5.1. $(P_t)_{t\geq 0}$ is said to have the **Feller property** if $P_t(C_b(H)) \subset C_b(H)$ for all $t \geq 0$. Here, $C_b(H)$ denotes the space of bounded continuous functions $F : H \to \mathbb{R}$. If in addition $P_t(\mathcal{B}_b(H)) \subset C_b(H)$ for all $t > 0$, $(P_t)_{t\geq 0}$ is said to have the **strong Feller property**.

The (strong) Feller property of a semigroup $(P_t)_{t\geq 0}$ of Markovian integral operators has been identified in the classical theory of Markov processes on a locally compact state space as a generic property that implies the existence of an associated (strong) Markov process (see [**1**]). A large part of the potential theory, developed for (strong) Markov processes on locally compact state spaces, can be carried over to the case where the transition semigroup associated with (5.1) is (strong) Feller.

A sufficient condition for the Feller property is the continuous dependence in probability of the mild solution $X(t,x)$ w.r.t. its initial condition, i.e., $\lim_{n\to\infty} \|x_n - x\|_H = 0$ implies

$$\lim_{n\to\infty} \|X(t,x_n) - X(t,x)\|_H = 0 \qquad \text{in probability}$$

which holds under the following estimate

$$E\left(\|X(t,x)-X(t,y)\|_H^2\right) \le C_t\|x-y\|_H^2 \tag{5.2}$$

for any two initial conditions $x, y \in H$, using Chebychev's inequality. Estimate (5.2) is implied by the assumption (H.1) made in Chapter 3 (resp. (V.2) in Chapter 4) for the main existence theorem of a milde (resp. variational) solution. Various generalizations can be found in the literature, see for example [**3**].

5.2.2. Kolmogorov operator. As emphasized in the previous subsection the Markov property of the solution of (5.1) implies the semigroup property for the associated transition semigroup $(p_t)_{t\ge 0}$ (resp. for the corresponding integral operators $(P_t)_{t\ge 0}$). The function P_tF should formally satisfy the (forward) Kolmogorov equation

$$\frac{d}{dt}P_tF(x) = LP_tF(x) \tag{5.3}$$

where

$$\begin{aligned} LF(x) = \frac{1}{2}\ &\mathrm{tr}(C(x)C(x)^*F''(x)) + \langle x, A^*F'(x)\rangle_H \\ &+ \langle B(x), F'(x)\rangle_H \end{aligned} \tag{5.4}$$

is called the Kolmogorov operator associated with (5.1) in honor of Kolmogorov who was the first to formulate and to solve the forward and backward partial differential equations satisfied by the transition probabilities of one-dimensional diffusion processes (see [**12**]). Note that L is a differential operator of second order and its precise domain of definition consists of all twice Frechet-differentiable functions $F : H \to \mathbb{R}$ satisfying $F'(x) \in D(A^*)$ for all $x \in H$. An appropriate subspace of smooth test-functions is provided by the space

$$\begin{aligned} \mathcal{F}C_b^2(D(A^*)) = \{F(x) = \varphi\left(\langle x, e_1\rangle, \dots, \langle x, e_n\rangle\right) \mid n \ge 1, \varphi \in C_b^2(\mathbb{R}^n)\,, \\ e_1, \dots, e_n \in D(A^*)\} \end{aligned}$$

of twice Frechet-differentiable finitely based cylindrical test functions. If $F \in \mathcal{F}C_b^2(D(A^*))$ admits the representation $F(x) = \varphi(\langle x, e_1\rangle, \dots, \langle x, e_n\rangle)$, then (5.4) reduces to the following expression

$$\begin{aligned} LF(x) = \frac{1}{2}&\sum_{k,l=1}^n \langle C(x)C(x)^*e_k, e_l\rangle \varphi_{x_k x_l}(\langle x, e_1\rangle, \dots, \langle x, e_n\rangle) \\ &+ \sum_{k=1}^n (\langle x, A^*e_k\rangle + \langle B(x), e_k\rangle)\varphi_{x_k}(\langle x, e_1\rangle, \dots, \langle x, e_n\rangle)\,. \end{aligned}$$

The importance of the Kolmogorov operator consists of the fact that it is explicitly given on smooth functions in contrast to the transition semigroup and that many quantities associated with (5.1) can be estimated using L rather than $(p_t)_{t\ge 0}$. The drawback of the Kolmogorov operator L however is that it might not uniquely determine the transition semigroup because the solution of the forward Kolmogorov equation might be nonunique. Rigorous results concerning existence and uniqueness of solutions of equation (5.3) in the infinite dimensional case can be found in the monograph [**6**].

Suppose that for all $x \in H$ (5.1) has a unique strong solution $(X(t,x))_{t\geq 0}$ with initial condition x. Then Ito's formula (see Theorem 4.17 in [**5**]) implies for suitable F in the domain of definition of the associated Kolmogorov operator L that

$$
\begin{aligned}
F(X(t,x)) = F(x) + \int_0^t LF(X(s,x))\,ds \\
+ \int_0^t \langle F'(X(s,x)), C(X(s,x))dW(s)\rangle\,.
\end{aligned}
\tag{5.5}
$$

If $E\left(\int_0^t |LF(X(s,x))| + \|C^*(X(s,x))F'(X(s,x))\|^2\,ds\right) < \infty$, $t \geq 0$, then (5.5) implies, taking expectations, that

$$
P_tF(x) = F(x) + \int_0^t P_sLF(x)\,ds
$$

so that similar to (5.3)

$$
\frac{d}{dt}P_tF(x) = P_tLF(x)
\tag{5.6}
$$

but with P_t and L interchanged. It requires further mathematical theory to identify sufficient conditions under which both equations (5.3) and (5.6) are in fact equivalent.

5.3. Invariant measures

Many important properties of random dynamical systems modelled by (5.1) can be read off from the long-time behaviour of its solution $(X(t))_{t\geq 0}$. One can distinguish between its pathwise behaviour and the statistical behaviour given in terms of the transition probabilities $(p_t)_{t\geq 0}$. In this section we will concentrate on the latter and introduce the basic concepts needed for a study of the long-time behaviour of $(p_t)_{t\geq 0}$ and the associated semigroup $(P_t)_{t\geq 0}$.

An important question in the understanding of the qualitative behaviour of $(X(t))_{t\geq 0}$ for large time is the existence of stationary states of its transition probabilities, since they correspond to a situation where the system described by $(X(t))_{t\geq 0}$ is in equilibrium. In contrast to the deterministic case, this is meant as a statistical equilibrium rather that a pathwise equilibrium. Further questions then address the qualitative properties of stationary states, e.g. their stability properties and support properties.

It is natural to define a stationary state for the transition probabilities $(p_t)_{t\geq 0}$ of (5.1) as a probability measure μ on H having the property that if the solution $X(t)$ has distribution μ at some time point, say t_0, then the distribution of $X(t)$ will be the same for all later times $t \geq t_0$, hence invariant under time evolution. Using the Markov property this implies in particular for all $s \geq 0$ that

$$
\begin{aligned}
\int P_sF(x)\,\mu(dx) &= E\left(P_sF(X(t_0))\right) = E\left(E\left(F(X(t_0+s)) \mid \mathcal{F}_{t_0}\right)\right) \\
&= E\left(F(X(t_0+s))\right) = \int F(x)\,\mu(dx)\,.
\end{aligned}
$$

An important subclass of stationary states are given by the class of (time-) reversible states for $(X(t))_{t\geq 0}$. By this we mean a probability measure μ with the property, that if the solution of (5.1) has distribution μ at time t, the joint distribution of

$(X(t), X(t+s))$ is the same as the joint distribution of the time-reversed process $(X(t+s), X(t))$. In particular,

$$E\left(F(X(t))G(X(t+s))\right) = E\left(F(X(t+s))G(X(t))\right)$$

or equivalently, using the Markov property,

$$\begin{aligned}\int F(x)P_sG(x)\,\mu(dx) &= E\left(F(X(t))P_sG(X(t))\right)\\ &= E\left(F(X(t))E\left(G(X(t+s))\mid \mathcal{F}_t\right)\right)\\ &= E\left(F(X(t))G(X(t+s))\right) = E\left(F(X(t+s))G(X(t))\right)\\ &= \int P_sF(x)G(x)\,\mu(dx)\,.\end{aligned}$$

Since the semigroup $(P_t)_{t\geq 0}$ associated with (5.1) is in most cases not explicitely given, it is convenient to look for alternative characterizations of invariance and reversibility in terms of the associated Kolmogorov operator L. This motivates the following

Definition 5.2. A probability measure μ on (the Borel σ-algebra of) H is called

(i) **invariant** if

$$\int P_tF\,d\mu = \int F\,d\mu \quad \forall t\geq 0 \quad \forall F\in\mathcal{B}_b(H)\,,$$

(ii) **infinitesimally invariant** if

$$L(\mathcal{F}C_b^2(D(A^*)))\subset L^1(\mu) \text{ and } \int LF\,d\mu = 0 \quad \forall F\in\mathcal{F}C_b^2(D(A^*))\,,$$

(iii) **reversible** if

$$\int P_tF\,G\,d\mu = \int F\,P_tG\,d\mu \quad \forall t\geq 0 \quad \forall F,G\in\mathcal{B}_b(H)\,,$$

(iv) **symmetrizing** if $L(\mathcal{F}C_b^2(D(A^*)))\subset L^1(\mu)$ and

$$\int LF\,G\,d\mu = \int F\,LG\,d\mu \quad \forall F,G\in\mathcal{F}C_b^2(D(A^*))\,.$$

Remark 5.3. Under weak assumptions on the coefficients, the following implications hold:

$$\begin{array}{ccc}\mu \text{ reversible} & \Rightarrow & \mu \text{ symmetrizing}\\ \Downarrow & & \Downarrow\\ \mu \text{ invariant} & \Rightarrow & \mu \text{ infinitesimally invariant}\end{array}$$

The implications "μ reversible $\Rightarrow$ μ invariant" (resp. "μ symmetrizing $\Rightarrow$ μ infinitesimally invariant") are obvious: simply choose $G\equiv 1$, then μ reversible (for the semigroup $(P_t)_{t\geq 0}$) implies

$$\int P_tF\,d\mu = \int P_tF\,1\,d\mu = \int F\,P_t1\,d\mu = \int F\,d\mu$$

since $P_t 1 = 1$ (resp. $\int LF\,d\mu = \int LF\,1\,d\mu = \int F\,L1\,d\mu = 0$ since $L1 = 0$). The remaining two implications "μ reversible $\Rightarrow$ μ symmetrizing" (resp. "μ invariant $\Rightarrow$ μ infinitesimally invariant") follow from differentiating the semigroup in $t = 0$:

$$\int LF\,G\,d\mu = \frac{d}{dt}\int P_t F\,G\,d\mu_{|t=0} = \frac{d}{dt}\int F\,P_t G\,d\mu_{|t=0} = \int F\,LG\,d\mu$$

(resp. $\int LF\,d\mu = \frac{d}{dt}\int P_t F\,d\mu_{|t=0} = \frac{d}{dt}\int F\,d\mu_{|t=0} = 0$). None of the converse implications hold true in general. Some finite-dimensional counterexamples may be found in [**2**].

5.3.1. Existence of invariant measures – Krylov-Bogoliubov theory. The existence of invariant measures for equation (5.1) is strongly related to its stability properties. This was first observed and exploited in the finite dimensional case by R.Z. Hasminskii in [**9**]. A crucial role is played by the family μ_T, $T \geq 0$, of mean occupation time measures

$$\mu_T(A) := \frac{1}{T}\int_0^T P(X(t) \in A)\,dt\,, \quad A \in \mathcal{B}(H)$$

of the solution to (5.1). Indeed, if the system described by $(X(t))_{t\geq 0}$ approaches a stationary state it will spend with high probability most of the time in bounded, in fact even relatively compact, subregions of its state space. The resulting property for its mean occupation time measures is called tightness. This is the basic observation exploited in the Krylov-Bogoliubov theory.

THEOREM 5.4. *(Krylov-Bogoliubov)*
Assume that for some $T_0 > 0$

- *the transition semigroup* $(P_t)_{t\geq 0}$ *associated with* (5.1) *is Feller,*
- *the family* $(\mu_T)_{T\geq T_0}$ *of mean occupation time measures is tight, i.e., for all* $\varepsilon > 0$ *there exists a compact subset* $K_\varepsilon \subset H$ *with* $\mu_T(K_\varepsilon) \geq 1-\varepsilon$, $T \geq T_0$.

Then there exists an invariant measure μ *for* (5.1). *Moreover, every limit* μ_∞ *of some weakly convergent subsequence* $(\mu_{T_n})_{n\geq 1}$ *with* $T_n \to \infty$, *is an invariant measure.*

The tightness of $(\mu_T)_{T\geq T_0}$ can be deduced from the existence of a (proper) Lyapunov function $V : H \to \mathbb{R}\cup\{+\infty\}$, that is, a function with compact level sets $\{V \leq \alpha\}$, $\alpha \in \mathbb{R}$, that satisfies

$$\sup_{T\geq T_0}\int V\,d\mu_T \quad < \infty \tag{5.7}$$

or equivalently,

$$\sup_{T\geq T_0}\frac{1}{T}\int_0^T E(V(X(t)))\,dt < \infty\,.$$

To call V a Lyapunov function is motivated by the stability theory for deterministic systems. In fact, for a deterministic evolution system $(x(t))_{t\geq 0}$ a Lyapunov function is a (sufficiently regular) function V for which $V(x(t))$ is decreasing in time, so that

$$\frac{1}{T}\int_0^T V(x(t))\,dt \leq V(x(0))$$

remains bounded. The stochastic analogue therefore is a function V for which $E\left(V(X(t))\right)$ is decreasing in time, i.e., a function V for which $V(X(t))$ is decreasing in the statistical average. It follows that

$$\int V\, d\mu_T = \frac{1}{T}\int_0^T E\left(V(X(t))\right)\, dt \le E\left(V(X(0))\right)$$

remains bounded. For the purposes of Theorem 5.4 it then suffices to require only (5.7).

5.3.1.1. *Linear equations.* Consider the Ornstein-Uhlenbeck equation

$$dX(t) = AX(t)\, dt + C\, dW(t), \qquad X(0) = \xi\,. \tag{5.8}$$

Under the assumption that

$$Q_t = \int_0^t e^{sA} Q e^{sA^*}\, ds\,, \quad Q := CC^*$$

is a trace class operator, there exists a unique mild solution

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A} C\, dW(s)\,, \quad t \ge 0$$

of (5.8) (not necessarily having a pathwise continuous version) and the corresponding transition semigroup

$$P_t F(x) = E\left(F(X(t)) \mid X(0) = x\right)\,, \quad x \in H$$

is well-defined. Recall that

$$W_{A,C}(t) = \int_0^t e^{(t-s)A} C dW(s) \quad \sim \quad N\left(0, Q_t\right)$$

so that the following integral representation of the transition probabilities

$$\begin{aligned} P_t F(x) &= \int_H F(e^{tA}x + w) N\left(0, Q_t\right)(dw) \\ &= \int_H F(w) N\left(e^{tA}x, Q_t\right)(dw)\,, \end{aligned} \tag{5.9}$$

called Mehler's formula, holds. The integral representation (5.9) is extremely useful to deduce various further information on the process. In particular, a complete solution to the problem of existence of invariant measures can be deduced. The result is due to Snyders and Zakai in the finite-dimensional case (see [**15**]) and to Zabczyk in the general Hilbert space case (see [**4**]).

THEOREM 5.5. *The following conditions are equivalent:*

(i) $\exists$ *invariant measure for* (5.8)
(ii) $\int_0^\infty \|e^{tA}C\|^2_{L_2(U,H)}\, dt = \sup_{t\ge 0}\ tr\left(Q_t\right) < \infty$
(iii) $\exists\ Q_\infty \in L(H)$, $Q_\infty \ge 0$, *symmetric,* $tr\left(Q_\infty\right) < \infty$, *satisfying the equation*

$$2\langle Q_\infty A^* x, x\rangle + \langle Qx, x\rangle = 0 \quad \forall x \in D(A^*). \tag{5.10}$$

In this case, any invariant measure μ *for* (5.8) *admits the representation*

$$\mu = \nu * N\left(0, \int_0^\infty e^{tA} Q e^{tA^*}\, dt\right)\,,$$

where ν is invariant for $x' = Ax$ and

$$Q_\infty = \int_0^\infty e^{tA} Q e^{tA^*}\, dt$$

is the minimal nonnegative solution of (5.10).

5.3.1.2. *Semilinear equations.* The existence part in the previous theorem can be extended to the semilinear case under the following assumptions:

(A.1) A is self-adjoint, of negative type and having a compact resolvent

In the following denote by $V_\gamma = D((-A)^\gamma)$, $\gamma \in \mathbb{R}$, the real interpolation space equipped with the scalar product

$$\langle u, v\rangle_\gamma := \langle (-A)^\gamma u, (-A)^\gamma v\rangle_H + \langle u, v\rangle_H\,, \quad u, v \in V_\gamma .$$

(A.2) For some $\gamma_1 \geq 0$, B operates as a vector-field on V_{γ_1} and for any initial condition $x \in V_{\gamma_1}$ (5.1) has a unique mild solution with continuous trajectories in V_{γ_1}.
(A.3) The associated transition semigroup $(P_t)_{t\geq 0}$ is Feller on V_{γ_1}.

Theorem 5.6. *Suppose that for some $\gamma_0 > \gamma_1$*

(A.4) $\exists\, \varepsilon_0 > 0$ *such that*

$$\sup_{t\geq 0} E\left(e^{\varepsilon \|W_{A,C}(t)\|^2_{V_{\gamma_0}}}\right) < \infty \qquad \forall \varepsilon < \varepsilon_0$$

(A.5)
– $\exists\ \Psi : V_{\gamma_1} \to \mathbb{R}_+$ *Frechet-differentiable,*
– $\exists\ \Theta : V_{\gamma_0} \to \mathbb{R}_+$ *coercive,*
– $\exists\ \beta, \delta \in \mathbb{R}$, $\varepsilon_1 < \varepsilon_0$ *such that*

$$\langle Ay + B(y+w), \Psi'(y)\rangle_{V_{\gamma_1}} \leq -\Theta(y) + \beta e^{\varepsilon_1 \|w\|^2_{V_{\gamma_0}}} + \delta$$

for all $y \in V_{1+\gamma_1}$, $w \in V_{\gamma_0}$.

Then $(\mu_T)_{T\geq 1}$ is tight on V_{γ_1} for any initial condition $\xi \in V_{\gamma_1}$. Moreover, any limit point μ of some weakly convergent subsequence of $(\mu_T)_{T\geq 1}$ is an invariant measure of (5.1) *and*

$$\int F(x)\, \mu(dx) < \infty$$

for any continuous $F : V_\gamma \to \mathbb{R}_+$, $\gamma < \gamma_0$, satisfying the growth condition

$$F(y+w) \leq c_1 \Theta(y) + c_2 e^{\varepsilon_1 \|w\|^2_{V_{\gamma_0}}} + c_3 \qquad \forall y \in V_{\gamma_1}\,, w \in V_{\gamma_0}\,.$$

The proof of the Theorem (for the case of additive noise) can be found in [**7**].

Example 5.7. In the example of the stochastic reaction diffusion equation

$$dX(t) = [\Delta X(t) - f(X(t))]\, dt + (-\Delta)^{-\beta}\, dW(t)$$

on $L^2([0,1])$ with $f(t) = a_{2n+1}t^{2n+1} + \ldots + a_1 t$ with $a_{2n+1} < 0$, let Δ be the Dirichlet-Laplacian. Then for $\beta > 0$, $\frac{1}{4} < \gamma < \frac{1}{4} + \beta$, $\beta < \frac{\gamma}{2(2n+1)}$, the evolution equation has an invariant measure μ satisfying the moment estimates

- $\int \|x\|^2_{V_\gamma}\, \mu(dx) < \infty$
- $\int e^{\varepsilon \|x\|^{\frac{2}{(2n+1)(n+1)}}}\, \mu(dx) < \infty$
- $\int \|x^r\|^2\, \mu(dx) < \infty\ \forall r \geq 1$

(see [**7**]).

REMARK 5.8. Comments on the assumptions (A.4) and (A.5):

(i) The exponential moment estimate

$$E\left(\exp\left(\varepsilon\left\|\int_0^T \Phi(s)\,dW(s)\right\|_H^2\right)\right) \le \frac{1}{1-\varepsilon A^2 2eK} < \infty \tag{5.11}$$

for $\varepsilon < \frac{1}{A^2 2eK}$ (see [**10**] and Exercises to Chapter 3) implies that the exponential moment estimate required for the stochastic convolution $W_{A,C}(t)$ in (A.4) is satisfied for bounded dispersion coefficients

$$\int_0^\infty \sup_{x\in H} \|e^{tA}C(x)\|^2_{L_2(U,V_{\gamma_1})}\,dt < \infty\,.$$

(ii) In the case of a classical Lyapunov function of polynomial type

$$\langle Ay + B(y+w), y\rangle_H \le -\alpha\|y\|^2_{V_{\frac{1}{2}}} + \beta\|w\|^{2s}_{V_{\gamma_0}} + \delta$$

for some $s \ge 1$, $\alpha > 0$, $\beta,\delta \in \mathbb{R}$, the theorem implies the following exponential moment estimate

$$K(\varepsilon) = \int e^{\varepsilon\|x\|_H^{\frac{2}{s}}}\,\mu(dx) < \infty \qquad \text{for } \varepsilon \le \varepsilon_1\,.$$

In particular, μ has (sub-) Gaussian tails $\mu\left(\|x\|_H \ge R\right) \le K(\varepsilon_1)e^{-\varepsilon_1 R^{\frac{2}{s}}}$, $R > 0$.

5.3.1.3. *Further examples.* **(a) symmetrizing measures**

It is well-known that for a smooth potential $\varphi : \mathbb{R}^d \to \mathbb{R}$ satisfying $Z := \int e^{2\varphi}\,dx < \infty$ and $\int |\nabla\varphi| e^{2\varphi}\,dx < \infty$ the measure $\mu := Z^{-1}e^{2\varphi(x)}\,dx$ is the unique invariant (and in fact reversible, hence symmetrizing) measure for the transition semigroup associated with the stochastic (ordinary) differential equation

$$dX(t) = \nabla\varphi(X(t))\,dt + dW(t) \quad \in \mathbb{R}^d \tag{5.12}$$

(see [**2**]). The example ultimately goes back to Kolmogorov, who in his classical paper [**13**], fully characterized reversible measures of diffusions on compact manifolds in terms of the coefficients of the associated generator. To see that μ is symmetrizing, note that it satisfies the following integration by parts formula

$$\int \partial_{x_k}F(x)\,\mu(dx) = -2\int F(x)\partial_{x_k}\varphi(x)\,\mu(dx) \quad \forall F \in C_b^1(\mathbb{R}^d) \tag{5.13}$$

so that for the Kolmogorov operator

$$LF(x) = \frac{1}{2}\Delta F(x) + \langle\nabla\varphi(x), \nabla F(x)\rangle$$

we obtain that

$$\int LF\,G\,d\mu = -\int\langle\nabla F, \nabla G\rangle\,d\mu = \int F\,LG\,d\mu\,.$$

This observation can be extended to the infinite dimensional setting as follows: A sufficient (and in many cases also necessary) condition for the existence of a symmetrizing measure for the stochastic evolution equation (5.1) is the existence of a measure μ satisfying appropriate integrability conditions and the integration by parts formula

$$\int \langle C(x)C(x)^* F'(x), l\rangle \, \mu(dx) = -2 \int \left(\langle x, A^* l\rangle + \langle B(x), l\rangle\right) F(x) \, \mu(dx) \tag{5.14}$$

for any $l \in D(A^*)$.

Illustrations Assume that $C(x)C(x)^* \equiv Q > 0$ is constant and Q^{-1} bounded. Then (5.14) is equivalent with

$$\int \partial_l F(x) \, \mu(dx) = -2 \int \left(\langle x, A^* Q^{-1} l\rangle + \langle B(x), Q^{-1} l\rangle\right) F(x) \, \mu(dx) \tag{5.15}$$

for any l satisfying $Q^{-1} l \in D(A^*)$. Under the additional assumption that

$$AQ = QA^* \tag{5.16}$$

and $Q^{-1}A$ is invertible with bounded inverse, the linear part $\langle x, A^* Q^{-1} l\rangle$ in the logarithmic derivative can be identified as the logarithmic derivative of the Gaussian measure $N(0, \frac{1}{2} Q A^{-1})$ which implies that it is the, in fact unique, symmetrizing measure of the linear stochastic evolution equation

$$dX(t) = AX(t) \, dt + C \, dW(t) \, .$$

Suppose now that $Q = Id_H$, so that (5.16) implies that $A = A^*$, i.e., A self-adjoint, and $B = \varphi'$ for some potential $\varphi : H \to \mathbb{R}$ satisfying some growth condition (e.g. bounded from above). Then, similar to the finite-dimensional case,

$$2\langle x, Al\rangle + \langle \varphi'(x), l\rangle \, , \qquad l \in D(A)$$

is the logarithmic derivative of the measure

$$\mu(dx) \propto e^{2\varphi(x)} N\left(0, \frac{1}{2} A^{-1}\right)(dx) \tag{5.17}$$

(see [**16**], [**14**]). The above factorization of the symmetrizing measure is not unique and might be in fact not appropriate in given examples.

(b) infinitesimally invariant measures

An important class of stochastic evolution equations with invariant measures is provided by equations where the nonlinear part of the drift has an infinitesimally conserved quantity in the sense of the following two examples:

Example 5.9. (i) Stochastic Burgers equation

$$dX(t) = \left[\gamma \Delta X(t) + \partial_\xi (X(t)^2)\right] \, dt + C \, dW(t)$$

on the space $L^2([0,1])$ with Dirichlet boundary conditions. In this case, the L^2-norm is an infinitesimally conserved quantity for $\partial_\xi \left(X(t)^2\right)$, since $\int_0^1 \partial_\xi (u^2) \, u \, d\xi = 0$ for $u \in H_0^{1,2}([0,1])$.

(ii) Stochastic Navier-Stokes equations

$$dX(t) = \left[\gamma \Delta_S X(t) - \Pi \left(X(t) \cdot \nabla\right) X(t)\right] \, dt + C \, dW(t) \, , \ \text{div } X(t) = 0$$

on the space $L_0^2([0,1]^d; R^d)$ of divergence-free, square-integrable vector-fields. In this case, the L^2-norm is an infinitesimally conserved quantity for the convection term, since $\int_{[0,1]^d} \langle (u \cdot \nabla)\, u, u \rangle \, d\xi = 0$ for smooth, divergence-free vector-fields u with periodic boundary conditions.

In both examples existence of stationary martingale solutions can be obtained in the case where the covariance operator $Q = CC^*$ has finite trace, using the compactness method, first presented in the paper [**8**] (see also [**11**] for the case of stochastic Navier-Stokes Coriolis equations). The corresponding invariant distribution μ satisfies the exponential moment estimate

$$\int \left(1 + \|x\|^2_{V_{\frac{1}{2}}}\right) e^{\varepsilon \|x\|^2_H} \, \mu(dx) < \infty \qquad \text{for } \varepsilon < \frac{\gamma}{\|Q\|_{L(U,H)}} \,.$$

Bibliography

[1] R.M. Blumenthal, R.K. Getoor, Markov processes and potential theory, Pure and Applied Mathematics, Vol. 29 Academic Press, New York, 1968.

[2] V.I. Bogachev, M. Röckner, W. Stannat, Uniqueness of Invariant Measures and Maximal Dissipativity of Diffusion Operators on L^1, In: Infinite dimensional stochastic analysis, Editors: P. Clement et al. Amsterdam, Royal Netherlands Academy of Arts and Sciences, 39-54, 2000.

[3] Z. Brzezniak, D. Gatarek, Martingale solutions and invariant measures for stochastic evolution equations in Banach spaces, Stochastic Process. Appl. 84, 187-225, 1999.

[4] G. Da Prato, J. Zabczyk, Ergodicity for infinite-dimensional systems, London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, 1996.

[5] G. Da Prato, J. Zabczyk, Stochastic equations in infinite dimensions, Cambridge University Press, 1992.

[6] G. Da Prato, J. Zabczyk, Second Order Partial Differential Equations in Hilbert Spaces, London Mathematical Society Lecture Note Series, vol. 293, Cambridge University Press, 2002.

[7] A. Es-Sarhir, W. Stannat, Invariant measures for semilinear SPDE's with local Lipschitz drift coefficients and applications, Journal of Evolution Equations 8, 129–154, 2008.

[8] F. Flandoli, D. Gatarek, Martingale and stationary solutions for stochastic Navier-Stokes Equations, Probab. Theory Rel. Fields 102, 367–391, 1995.

[9] R.Z. Hasminskii, Ergodic properties of recurrent diffusion processes and stabilitation of the solution of the Caudy problem for parabolic equations, Theory Probab. Appl. 5, 179–196, 1960

[10] E. Hausenblas, J. Seidler, Stochastic Convolutions Driven by Martingales: Maximal Inequalities and Exponential Integrability, Stochastic Analysis and Applications 26, 98-119, 2008.

[11] M. Hieber, W. Stannat, Stochastic Stability of the Ekman spiral, manuscript, submitted, 2010.

[12] A. N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, Math. Ann. 104, 415-458, 1931.

[13] A. N. Kolmogorov, Zur Umkehrbarkeit der statistischen Naturgesetze, Math. Ann. 113, 766-772, 1937.

[14] M. Röckner, L^p-analysis of finite and infinite dimensional diffusion operators, in: Stochastic PDE's and Kolmogorov equations in Infinite Dimensions, Ed. G Da Prato, Lecture Notes in Math. 1715, Springer, Berlin, 1999.

[15] J. Snyders, M. Zakai, Stationary probability measures for linear differential equations driven by white noise, Journal of Differential Equations 8, 27–32, 1970.

[16] J. Zabczyk, Symmetric solutions of semilinear stochastic equations, in: Stochastic partial differential equations and applications, II (Trento, 1988), Lecture Notes in Math. 1390, Springer, Berlin, 1989.

APPENDIX A

Elements of linear functional analysis

We briefly discuss basic terminology from linear functional analysis, which is the mathematical of linear operators on general vector spaces. All this can be taken from elementary textbooks on functional analysis.

A.1. Banach and Hilbert spaces

Definition A.1. Let E be a real vector space. A map $E \to \mathbb{R}$, $x \mapsto \|x\|$, is called a **norm on** E if

(i) $\|x\| \geq 0$ for all $x \in E$ and $\|x\| \geq 0$ iff $x = 0$;
(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in E$ and $\alpha \in \mathbb{R}$;
(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in E$ (triangle inequality).

$(E, , \|\cdot\|)$ or simply E is called a **normed vector space**.

The following is obvious: Let $(E, , \|\cdot\|)$ be a normed vector space. Then

$$d(x, y) := \|x - y\|$$

for all $x, y \in E$ defines a metric on E.

Definition A.2. A normed space which is complete w.r.t. the metric induced by the norm is called a **Banach space**.

In other words: A normed vector space E is a Banach space iff every Cauchy-sequence $(x_n) \subset E$ is convergent (in E).

Definition A.3. Let E be a real vector space. A map $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$, is called a **scalar product on** E if

(i) $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in E$;
(ii) $\langle x, x \rangle \geq 0$ for all $x \in E$ and $\langle x, x \rangle = 0$ iff $x = 0$;
(iii) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ for all $x, y, z \in E$ and $\alpha, \beta \in \mathbb{R}$.

$(E, , \langle \cdot, \cdot \rangle)$ or simply E is called a **inner product vector space**.

Theorem A.4. *(Cauchy-Schwarz inequality) Let $(E, \langle \cdot, \cdot \rangle)$ be an inner product vector space. Then*

$$|\langle x, y \rangle| \leq \|x\|\|y\| \tag{A.1}$$

for all $x, y \in E$ with equality iff x and y are linearly dependent.

The following is obvious from (A.1): Let $(E, , \langle \cdot, \cdot \rangle)$ be an inner product vector space. Then

$$\|x\| := \sqrt{\langle x, x \rangle}$$

for all $x \in E$ defines a norm on E.

Definition A.5. (Hilbert space) An inner product vector space which is complete with respect to the induced norm is called a **Hilbert space**.

A.2. Orthogonality, orthonormal systems, Parseval's identity

Recall that two vectors u and v of an inner product space are called orthogonal if $\langle u, v\rangle = 0$.

DEFINITION A.6. Let H be an inner product space with inner product $\langle \cdot, \cdot\rangle$ and induced norm $\|\cdot\|$. Let $M \subset H$ be a non-empty subset.

(i) M is called an orthogonal system if $\langle u, v\rangle = 0$ for all u,$v \in M$ with $u \neq v$.
(ii) M is called an orthonormal system if it is an orthogonal system and$\|u\| = 1$ for all $u \in M$.
(iii) M is called a complete orthonormal system or orthonormal basis of H if it is an orthogonal system and $\overline{\operatorname{span}(M)} = H$.

A.3. Linear operators

As already mentioned several times, a linear operator $L : E \to F$ between two Banach spaces need not be continuous if the domain E is infinite dimensional. A classical example is the derivative

$$Lf := \begin{cases} f' & \text{if } f \in C^1([0,1]) \\ 0 & \text{otherwise} \end{cases}$$

as a linear operator on the Banach space $C([0, 1])$ of continuous functions on $[0, 1]$ (endowed with the supremum-norm). Here, $C^1([0, 1])$ denotes the subspace of one times continuously differentiable functions. Then, e.g. $f_n(x) := \frac{1}{n}x^n \to 0$ in $C([0, 1])$, but $Lf_n(x) = x^{n-1}$ is not converging to 0.

In the following let E, F be two (real) normed vector spaces. We denote with $L(E, F)$ the space of **continuous** linear operators $L : E \to F$. We then have the following fundamental theorem of linear operators between normed vector spaces:

THEOREM A.7. *For $L : E \to \mathbb{F}$ linear, the following statements are equivalent:*

(i) *L is uniformly continuous;*
(ii) *$L \in L(E, F)$, i.e., L is continuous;*
(iii) *L is continuous at $x = 0$;*
(iv) *L is bounded;*
(v) *There exists $\alpha > 0$ such that $\|Lx\|_F \leq \alpha\|x\|_E$ for all $x \in E$.*

PROOF. The implications $(i) \Rightarrow (ii)$ and $(ii) \Rightarrow (iii)$ are obvious. Suppose that L is continuous at $x = 0$ and let $U \subset E$ be any bounded subset, and let $M := \sup_{x\in M} \|x\|_E$ be its upper bound. We have to show that $L(U) \subset F$ is bounded. Since L is continuous at $x = 0$, we can find $\delta > 0$ such that $\|x\|_E \leq \delta$ implies $\|Lx\|_F \leq 1$. Consequently, if $x \in U$,

$$\|Lx\|_F = \frac{M}{\delta}\|L(\underbrace{\frac{\delta}{M}x}_{\|\cdot\|\leq\delta})\| \leq \frac{M}{\delta},$$

which shows that $L(U) \subset B^F_{\frac{M}{\delta}}(0)$. Suppose now that L is bounded, then there exists in particular $\alpha > 0$ such that $L(B^E_1(0)) \subset B^F_\alpha(0)$, and thus if $x \in E \setminus \{0\}$, then

$$\|Lx\|_F = \|x\|_E\|L(\underbrace{\frac{x}{\|x\|}}_{\|\cdot\|\leq 1})\| \leq \alpha\|x\|.$$

Since this is trivial if $x = 0$, we obtain (v). Finally, suppose that (v) holds. Then by linearity of L

$$\|Lx - Ly\|_F = \|L(x-y)\|_F \leq \alpha\|x-y\|_E,$$

showing the uniform continuity of L. □

Definition A.8 (operator norm). For $L \in L(E,F)$ we define

$$\|L\|_{L(E,F)} := \inf\{\alpha > 0 \mid \|Lx\|_F \leq \alpha\|x\|_E \text{ for all } x \in E\}.$$

We call $\|L\|_{L(E,F)}$ the **operator norm of** L.

The following Proposition summarizes all general equivalent characterizations of $\|L\|_{L(E,F)}$.

Proposition A.9. *Let $L \in L(E,F)$. Then $\|Lx\|_F \leq \|L\|_{L(E,F)}\|x\|_E$ for all $x \in E$. Moreover,*

$$\|L\|_{L(E,F)} = \sup_{x\in E\setminus\{0\}} \frac{\|Lx\|_F}{\|x\|_E} = \sup_{\|x\|_E=1} \|Lx\|_F = \sup_{\|x\|_E\leq 1} \|Lx\|_F = \sup_{\|x\|_E<1} \|Lx\|_F.$$

The space $L(E,F)$, endowed with the operator norm $\|\cdot\|_{L(E,F)}$, is again a normed vector space, and it is Banach space as soon as F (!) is a Banach space.

A.4. Duality

Definition A.10 (Dual space). If E is a (real) normed vector space, then

$$E' = L(E,\mathbb{R})$$

is the **dual space** of E. Elements of E' are often also referred to as **bounded linear functionals on** E. The operator norm on $L(E,\mathbb{R})$ is called the dual norm on E' and is denoted by $\|\cdot\|_{E'}$. If $f \in E'$, then we define

$$\langle f, x\rangle := f(x).$$

Remark A.11. By definition of the operator norm, the dual norm is given by

$$\|f\|_{E'} = \sup_{x\in E\setminus\{0\}} \frac{|\langle f,x\rangle|}{\|x\|_E} = \sup_{\|x\|_E\leq 1} |\langle f,x\rangle|$$

and as a consequence

$$|\langle f,x\rangle| \leq \|f\|_{E'}\|x\|_E.$$

A.5. Duality in Hilbert spaces

Suppose now that H is a (real) Hilbert space (with inner product $\langle\cdot,\cdot\rangle$). Since for all $y \in H$, the linear mapping

$$J(y) : H \to \mathbb{R}, x \mapsto \langle x,y\rangle$$

is continuous, we have that $J(y) \in H'$ with

$$\|J(y)\|_{H'} = \sup_{\|x\|_H\leq 1} |\underbrace{\langle J(y),x\rangle}_{=\langle x,y\rangle}| \leq \|y\|_H.$$

It follows that $y \mapsto J(y)$ is continuous, but in fact also the converse inequality holds, since we can choose the uni vector $x = \frac{y}{\|y\|_H}$ to conclude that

$$\|J(y)\|_{H'} = \sup_{\|x\|_H \leq 1} |\underbrace{\langle J(y), x\rangle}_{=\langle x,y\rangle}| \geq |\langle \frac{y}{\|y\|_H}, y\rangle| = \frac{\|y\|^2}{\|y\|} = \|y\|_H.$$

Consequently, $\|J(y)\|_{H'} = \|y\|_H$.

THEOREM A.12. *(Riesz representation theorem) The map $J : H \to H'$ is a linear isometric isomorphism. It is called the* ***Riesz isomorphism****.*

A.6. Reflexivity

The dual space $E'' := (E')'$ of E' is called the bi-dual space. Note that any $x \in E$ induces an element in the bi-dual space of E by considering the continuous linear functional

$$f \mapsto \langle f, x\rangle, f \in E'.$$

It follows that

$$\|x\|_{E''} = \sup_{\|f\|_{E'} \leq 1} |\langle f, x\rangle| \leq \sup_{\|f\|_{E'} \leq 1} \|f\|_{E'} \|x\|_E \leq \|x\|_E$$

hence the embedding $E \subset E''$ is continuous. In fact, due to the Hahn-Banach Theorem, there exists $f_0 \in E'$ such that $\|f_0\|_{E'} \leq 1$ and $\langle f_0, x\rangle = \|x\|$, which implies the converse inequality

$$\|x\|_E = |\langle f_0, x\rangle| \leq \sup_{\|f\|_{E'} \leq 1} |\langle f, x\rangle| = \|x\|_{E''},$$

so that

$$\|x\|_E = \|x\|_{E''}$$

and thus the embedding $E \subset E''$ is an isometry w.r.t. the respective norms. We say that a normed vector space E is **reflexive** if $E = E''$.

The most important class of reflexive spaces are Hilbert spaces (!), so in particular $L^2(\Omega, \mathcal{A}, \mu)$. Other important examples for reflexive function spaces are $L^p(\Omega, \mathcal{A}, \mu)$ for $p \in (1, \infty)$, since for $1 \leq p < \infty$ the space $L^{p^*}(\Omega, \mathcal{A}, \mu)$ with $p^* = \frac{p}{p-1}$ (and $p^* = \infty$ if $p = 1$) is the dual space of $L^p(\Omega, \mathcal{A}, \mu)$ and $p^{**} = p$ for $p \in (1, \infty)$.

APPENDIX B

Weak Topology

We briefly discuss basic functional analytic facts concerning the weak topology on (reflexive) Banach and Hilbert spaces. In particular we will recall the basic fact that weak convergence + norm convergence implies strong convergence in Hilbert spaces. We also formulate Banach-Alaoglu's theorem, and its important consequence for reflexive Banach spaces.

B.1. Weak convergence

Definition B.1. Let E be a real Banach space and (x_n) be a sequence in E. We say that (x_n) converges weakly to $x \in E$ if

$$\lim_{n\to\infty} \langle f, x_n \rangle = \langle f, x \rangle$$

for all $f \in E'$. We write $x_n \rightharpoonup x$ in that case.

Remark B.2. (i) Clearly, $x_n \to x$ (strongly) in E implies $x_n \rightharpoonup x$ weakly, since

$$|\langle f, x_n \rangle - \langle f, x \rangle| = |\langle f, x_n - x \rangle| \leq \|f\|_{E'} \|x_n - x\|_E \to 0$$

(ii) The converse is not true, that is weak convergence does not imply strong convergence. As an example consider the sequence $e_n := (\delta_{kn})_{k\in\mathbb{N}}$. Then $\|e_n\|_{\ell^2} = 1$ and $\langle f, e_n \rangle = f_n$ for all $f = (f_k)_{k\in\mathbb{N}} \in \left(\ell^2\right)' = \ell_2$. Hence for all $f \in \left(\ell^2\right)'$

$$\langle f, e_n \rangle = f_n \to 0$$

as $n \to \infty$ weakly. However, (e_n) does not converge w.r.t. the norm.

(iii) The weak limit of a sequence is unique. To see this note that if $\langle f, x \rangle = 0$ for all $f \in E'$, then $x = 0$, since otherwise the Hahn-Banach Theorem implies the existence of $f \in E'$ with $\langle f, x \rangle \neq 0$.

Although weak convergence does not imply strong convergence we have an upper bound of $\|x\|_E$.

Proposition B.3. *Suppose that E is a Banach space and that $x_n \rightharpoonup x$ weakly in E. Then (x_n) is bounded and*

$$\|x\|_E \leq \liminf_{n\to\infty} \|x_n\|_E.$$

Proof. We consider x_n as an element of the bi-dual space E'', the dual space of E'. By assumption $\langle f, x_n \rangle \to \langle f, x \rangle$ for all $f \in E'$. But that means that x_n, considered as linear functionals on E' are pointwise bounded and therefore bounded in the operator norm by the uniform boundedness principle, i.e., $\sup_n \|x_n\|_{E''} < \infty$. Since $\|x_n\|_{E''} = \|x_n\|_E$, it follows that (x_n) is bounded in E. Now,

$$|\langle f, x_n \rangle| \leq \|f\|_{E'} \|x_n\|_E$$

implies

$$|\langle f, x\rangle| = \lim_{n\to\infty} |\langle f, x_n\rangle| \le \|f\|_{E'} \liminf_{n\to\infty} \|x_n\|_E$$

for all $f \in E'$. Now, let $f \in E'$ be such that $\langle f, x\rangle = \|x\|_E$ and $\|f\|_{E'} = 1$, such a functional exists due to the Hahn-Banach Theorem, we then obtain

$$\begin{aligned}\|x\|_E = |\langle f, x\rangle| &= \lim_{n\to\infty} |\langle f, x_n\rangle| \le \|f\|_{E'} \liminf_{n\to\infty} \|x_n\|_E \\ &= \liminf_{n\to\infty} \|x_n\|_E.\end{aligned}$$

□

In the case of Hilbert spaces, the proof of the above Proposition becomes much easier. Indeed, suppose that H is a Hilbert space and that $x_n \rightharpoonup x$ weakly in H. Then

$$\|x\|_H^2 = \langle x, x\rangle = \lim_{n\to\infty} \langle x, x_n\rangle \le \|x\|_H \liminf_{n\to\infty} \|x_n\|_H$$

and so

$$\|x\|_H \le \liminf_{n\to\infty} \|x_n\|_H.$$

In the case of Hilbert spaces we also have the following important converse:

Proposition B.4. *Suppose that H is a Hilbert space and that $x_n \rightharpoonup x$ weakly in H is such that in addition $\|x_n\|_H \to \|x\|_H$. Then $x_n \to x$ strongly in H.*

Proof. Since $\langle x, x_n\rangle \to \|x\|_H^2$ and $\|x_n\|_H^2 \to \|x\|_H^2$, it follows that

$$\lim_{n\to\infty} \|x - x_n\|_H^2 = \lim_{n\to\infty} \|x\|_H - 2\langle x, x_n\rangle + \|x_n\|_H^2 = \|x\|_H - 2\|x\|^2 + \|x\|_H^2 = 0.$$

□

B.2. Banach-Alaoglu

In its general form, the Banach-Alaoglu theorem states that the unit ball $B_1^{E'}$ in the dual E' of some normed vector space E is compact in the weak*-topology. The weak*-topology on E' is the topology induced by the weak*-convergence. A sequence $(f_n) \subset E'$ is said to be **weak*-convergent** to some element $f \in E'$ if

$$\langle f_n, x\rangle \to \langle f, x\rangle \text{ for all } x \in E,$$

i.e. we just interchange the convergence in the duality mapping $\langle f, x\rangle$.

Theorem B.5. *(Banach-Alaoglu, Weak*-compactness) The unit ball (and thus any bounded subset) in a separable Banach space E is weakly* sequentially compact, i.e. any bounded sequence contains a weak*-convergent subsequence. If E is in addition reflexive, i.e., E can be identified with its bi-dual E'', e.g. a Hilbert space, then any bounded subset is sequentially compact w.r.t. the weak topology.*