# Lecture notes on SPDEs

Philipp Harms

February 10, 2017

## Contents

# 1 Introduction

## 1.1 Stochastic evolution equations

These lectures concentrate on stochastic partial differential equations (SPDEs) of evolutionary type. These are equations of the general form

$$dX_t = A(t, X_t)dt + B(t, X_t)dW_t,$$

where $X$, $A$, and $B$ are processes with values in some Banach space, and $W$ is Brownian motion on a Hilbert space.

## 1.2 Motivating examples

### 1.2.1 Random motion of a string

A random evolution of a string in $\mathbb{R}^d$ can be modeled as a process $X$ with values in $E := C([0,1];\mathbb{R}^d)$ solving

$$dX_t(u) = \Big(\Delta X_t(u) + f\big(X_t(u)\big)\Big)\mathrm{d}t + b\big(X_t(u)\big)\mathrm{d}W_t(u),$$

where the Laplace operator $\Delta$ models elastic forces within the string, the function $f : \mathbb{R}^d \to \mathbb{R}^d$ models an external forcing field acting on the string, the function $b : \mathbb{R}^d \to \mathbb{R}^d$ models the intensity of a random external force, and $W$ is a cylindrical Wiener process on $L^2([0,1])$ with the identity covariance operator. This equation is the limit of the following system of $N = 1/h$ interacting particles at positions $u_i \in [0,1]$:

$$\begin{aligned} dX_t(u_i) = \left(\frac{1}{2}\frac{X_t(u_{i-1}) - 2X_t(u_i) + X_t(u_{i+1})}{h^2} + f\big(X(u_i)\big)\right)\mathrm{d}t \\ + \sqrt{h}b\big(X(u_i)\big)\mathrm{d}W_t^i, \end{aligned}$$

where $W^1, \dots, W^N$ are independent scalar Brownian motions.

Many other examples of a similar flavor can be constructed by adding noise to physically motivated deterministic evolution equations such as Navier–Stokes, Korteweg de Vries, Schödinger, etc.

### 1.2.2 Zakai equation in non-linear filtering

There is an unobserved process $U$ and an observed process $Y$ given by

$$\begin{aligned} dU_t &= b(U_t)\mathrm{d}t + \sigma(U_t)\mathrm{d}W_t^1, \\ dY_t &= h(U_t)\mathrm{d}t + \mathrm{d}W_t^2, \end{aligned}$$

where all coefficients are in $C_b^\infty(\mathbb{R})$. Then $U$ is a Markov process with generator

$$Af(u) = b(u)\frac{\partial f(u)}{\partial u} + \frac{1}{2}\left(\sigma(u)\sigma(u)^\top\right)\frac{\partial^2 f(u)}{\partial u^2}.$$

Under an ellipticity assumption on $\sigma\sigma^\top$ and a regularity assumption on the law of $U_0$, the conditional law of $U_t$ given $\{Y_s, s \in [0,t]\}$ can be expressed in terms of an unnormalized density $X_t$ with respect to the Legesgue measure, which solves

$$dX_t(u) = A^* X_t(u)\mathrm{d}t + h(u)X_t(u)\mathrm{d}Y_t,$$

where $X_t$ takes values in a Sobolev space $E = H^k(\mathbb{R})$, $A^*$ is the (formal) adjoint of $A$, and $Y$ is Brownian motion under an equivalent probability measure. This stochastic evolution equation is studied in [Roz90].

The SPDE treatment has two advantages: first, it allows one to deduce and study regularity properties of the density, and second, it allows one to make use of numerical schemes for SPDEs.

### 1.2.3 Equations of population genetics

Let $X(t,u)$ denote the density of a population at time $t$ and location $u \in \mathbb{R}^d$. A model for unnormalized densities is the Dawson process

$$\mathrm{d}X(t,u) = a\Delta X(t,u)\mathrm{d}t + b\sqrt{X^+(t,u)}\mathrm{d}W_t(u),$$

and a model for normalized densities is the Fleming–Viot process

$$\mathrm{d}X(t,u) = \big(\Delta X(t,u) + aX(t,u) - b\big)\mathrm{d}t + b\sqrt{\frac{1}{2}X^+(t,u)\big(1 - X^+(t,u)\big)}\mathrm{d}W_t(u),$$

In these equations $X$ is a process with values in $L^2(\mathbb{R}^d)$, $a$ and $b$ are positive constants, and $W$ is cylindrical Brownian motion on some Hilbert space. The motivation behind these equations is a similar limiting procedure as in the example involving a random string.

**1 Reflection.** What processes does one get if the state space $\mathbb{R}^d$ of $u$ is replaced by a finite set $\{u_1, \dots, u_N\}$?

- CIR and Jacobi processes with interaction in the drift, if $\Delta$ is discretized as in the random string example.

Measure-valued processes can be used in many other applications such as stochastic portfolio theory, interest rate models with credit risk, energy delivery prices, etc.

## 1.3 Literature

We will use the following books and lecture notes as our main sources:

- G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions.* Cambridge university press, 2014
- J. Van Neerven. *Stochastic evolution equations.* ISEM lecture notes. 2007–2008
- A. Jentzen. *Stochastic Partial Differential Equations: Analysis and Numerical Approximations.* Lecture Notes, ETH Zürich. 2016
- C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations.* Vol. 1905. Springer, 2007

The motivating examples are taken from [DZ14].

# 2 Integration in Banach spaces

**2 Reflection.** Our aim is to define $\int f\mathrm{d}\mu$ for measures $\mu$ on $(A, \mathcal{A})$ and functions $f : A \to E$, where $E$ is a Banach space. What structural properties of the function space do we need?

- Functions can be approximated by simple functions.
- The space is closed under addition and scalar multiplication.

These issues will be addressed in the sequel.

## 2.1 Strong measurability

**3 Setting.** Let $(A, \mathcal{A}, \mu)$ be a finite measure space, let $E$ and $F$ be Banach spaces, and let $f, g : A \to E$.

**4 Definition.**

- $f$ is called *measurable* if $f^{-1}(B) \in \mathcal{A}$, for all $B \in \mathcal{B}(E)$.
- $f$ is called *weakly measurable* if $\langle f, x^* \rangle$ is measurable, for each $x \in E^*$.
- $f$ is called *simple* if it is measurable and the range of $f$ is finite.
- $f$ is called *strongly measurable* if it is measurable and the range of $f$ is separable.

**5 Reflection.** Is there a $\sigma$-algebra on $A$ such that $f$ is strongly measurable iff $f$ is measurable with respect to this $\sigma$-algebra?

- No. Indeed, it is easy to see that the $\sigma$-algebra generated by the strongly measurable functions is all of $\mathcal{A}$.

**6 Theorem** (Pettis measurability theorem)**.** *The following are equivalent:*

*(i)* $f$ *is strongly measurable;*

*(ii)* $f$ *is weakly measurable and has separable range;*

*(iii)* $f$ *is the point-wise limit of a sequence of simple functions;*

*(iv)* $f$ *is the point-wise limit of a sequence of strongly measurable functions.*

**7 Reflection.** What can go wrong if $f$ is only measurable instead of strongly measurable?

- If $f$ is a measurable function with non-separable range, then $f$ cannot be approximated by simple functions. This follows from Theorem 6.
- If $f$ and $g$ are measurable, it does not necessarily follow that $f + g$ is measurable:

$$\mathcal{A} \xrightarrow{(f,g)} \mathcal{B}(E) \otimes \mathcal{B}(E) \xleftarrow{I} \mathcal{B}(E \times E) \xrightarrow{+} \mathcal{B}(E)$$

  Writing $I = (\pi_1, \pi_2)$ for the identity and noting that the projections $\pi_1, \pi_2 : E \times E \to E$ are continuous, one sees that the mapping $I$ in the above diagram is measurable. Thus, $\mathcal{B}(E) \otimes \mathcal{B}(E) \subseteq \mathcal{B}(E \times E)$. If the inclusion is strict, $f + g$ might not be measurable (see [Els11, III.5.9–11] and [Gra16] for a discussion).

The proof of Theorem 6 requires some auxiliary lemmas.

**8 Lemma.** *The following are equivalent:*

*(i)* $f$ *is measurable;*

*(ii)* $f$ *is the pointwise limit of measurable functions.*

*Proof.* (i) $\Rightarrow$ (ii) is trivial. (ii) $\Rightarrow$ (i). Let $\phi \in C(E)$. Then $\phi \circ f$ is the pointwise limit of $\phi \circ f_n$. Thus, $\phi \circ f$ is measurable, as can be seen from

$$\begin{aligned}\{\phi \circ f \geq c\} &= \left\{\lim_{n\to\infty} \phi \circ f_n \geq c\right\} = \left\{\limsup_{n\to\infty} \phi \circ f_n \geq c\right\} \\ &= \left\{\lim_{n\to\infty} \sup_{m\geq n} \phi \circ f_m \geq c\right\} = \bigcap_{n\geq 0} \left\{\sup_{m\geq n} \phi \circ f_m \geq c\right\} \in \mathcal{A}.\end{aligned}$$

Thus, $f$ is Baire measurable, i.e., measurable with respect to the initial $\sigma$-algebra with respect to continuous functions on $E$. On metric spaces, the Baire and Borel algebras coincide. To see this, note that any open set $B$ can be written as $\phi^{-1}(\{0\})$, where $\phi \in C(E)$ is given by $\phi(x) = \mathrm{dist}(x, B^c)$. $\square$

**9 Lemma.** *The following properties hold:*

*(i) If $E$ is separable, then each subset $A \subseteq E$ is separable.*

*(ii) If $A \subseteq E$ is separable, then the closure of $A$ is separable.*

*(iii) If $A \subseteq E$ is separable, then the linear span of $A$ is separable.*

*Proof.* (i): Let $x_n$, $n \in \mathbb{N}$, be a dense sequence in $E$. Then there is a sequence $y_n$ in $A$ such that

$$\|y_n - x_n\| \leq 2^{-n} + \inf_{y\in A} \|y - x_n\|.$$

Then the sequence $y_n$ is dense in $A$ because one has for each $x \in A$

$$\begin{aligned}\inf_{n\in\mathbb{N}} \|x - y_n\| &\leq \inf_{n\in\mathbb{N}} \big(\|x - x_n\| + \|x_n - y_n\|\big) \\ &\leq \inf_{n\in\mathbb{N}} \big(\|x - x_n\| + 2^{-n} + \inf_{y\in A} \|y - x_n\|\big) \\ &\leq \inf_{n\in\mathbb{N}} \big(2\|x - x_n\| + 2^{-n}\big) = 0.\end{aligned}$$

This shows (i). (ii) and (iii) are trivial. $\square$

**10 Lemma.** *If $E$ is separable, then there exists a norming sequence $x_n^* \in E^*$, i.e., for every $x \in E$ one has*

$$\|x\| = \sup_n |\langle x, x_n^*\rangle|.$$

*Proof.* Let $x_n$ be a dense sequence in $E$, and let $\epsilon_n$ be a sequence of positive numbers converging to zero. For each $n$,

$$\|x_n\| = \sup_{\|x*\|\leq 1} |\langle x_n, x^*\rangle|.$$

Therefore, there is a unit vector in $x_n^* \in E^*$ such that

$$(1 - \epsilon_n)\|x_n\| \leq |\langle x_n, x_n^*\rangle|.$$

We claim that the sequence $x_n^*$ is norming. To see this, let $x \in E$ and let $x_{n_k}$ be a subsequence converging to $x$. Then

$$(1 - \epsilon_{n_k})\|x\| \leq (1 - \epsilon_{n_k})(\|x_{n_k}\| + \|x - x_{n_k}\|)$$

$$\leq |\langle x_{n_k}, x_{n_k}^* \rangle| + (1 - \epsilon_{n_k})\|x - x_{n_k}\|$$
$$\leq |\langle x, x_{n_k}^* \rangle| + (2 - \epsilon_{n_k})\|x - x_{n_k}\|$$

Taking the limes superior yields

$$\|x\| \leq \limsup_{k\to\infty} |\langle x, x_{n_k}^* \rangle| \leq \sup_n |\langle x, x_n^* \rangle|. \qquad \square$$

*Proof of Theorem 6.* (i) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (iii). Let $f$ be weakly measurable with separable range. Then the set $E_0 = \overline{\operatorname{span} f(A)}$ is separable by Lemma 9. Let $x_i$ be a dense sequence in $E_0$, and let $x_j^* \in E_0^*$ be the norming sequence provided by Lemma 10. For each $n$ let $f_n$ be the function mapping $a \in A$ to the point in $\{x_1, \ldots, x_n\}$ which is closest to $f(a)$; if several points $x_i$ have the same distance to $f(a)$, choose the point $x_i$ with the lowest index $i$. Note that

$$D_n(a) := \|f(a) - x_n\| = \sup_k |\langle f(a) - x_n, x_k^* \rangle|$$

is measurable in $a$, for each $n$. Therefore,

$$f_n^{-1}(x_k) = \left( \bigcap_{l=1}^{k-1} \{D_k < D_l\} \right) \cap \left( \bigcap_{l=1}^{n} \{D_k \leq D_l\} \right) \in \mathcal{A}.$$

Thus, $f_n$ is measurable. Moreover, $f_n$ has finite range, and the sequence $f_n$ converges pointwise to $f$. This shows (ii).

(iii) $\Rightarrow$ (iv) is trivial.

(iv) $\Rightarrow$ (i). Let $g_n$ be a sequence of strongly measurable functions converging pointwise to a function $f$. Then $f$ is measurable by Lemma 8. Moreover, the range of $f$ is contained in the set closure of the set $\bigcup_{n\in\mathbb{N}} g_n(A)$, which is separable by Lemma 9. This proves (iv). $\square$

## 2.2 Lebesgue-Bochner $L^p$ spaces

**11 Definition.** We call two functions $f, \tilde{f}$ versions of each other if $f = g$ holds $\mu$-almost everywhere; this defines an equivalence relation to be used in the sequel. For $0 \leq p < \infty$ we define $L^p(A; E)$ as the set of all equivalence classes of functions $f : A \to E$ having a strongly measurable version $\tilde{f}$ and satisfying

$$\|f\|_{L^p(A;E)} := \left( \int_A \|\tilde{f}\|^p \right)^{1/p} < \infty.$$

The topology on $L^0(A; E)$ is that of convergence in measure. Moreover, we define $L^\infty(A; E)$ as the set of all equivalence classes of functions $f : A \to E$ having a strongly measurable version and satisfying

$$\|f\|_{L^\infty(A;E)} := \inf \{r \geq 0 : \mu(\|f\| \geq r) = 0\} < \infty.$$

**12 Theorem.** *For each $1 \leq p \leq \infty$ the space $L^p(A; E)$ is a Banach space. For $1 \leq p < \infty$ the simple functions are dense in $L^p(A; E)$.*

*Proof.* Let $f, g$ be strongly measurable, let $F$ be a Banach space, and let $\phi \in C(E; F)$. Then Theorem 6.(iii) implies that $(f, g) : A \to E \times E$ and $\phi \circ f$ are strongly measurable. In particular, $f + g$ and $\lambda f$ are strongly measurable, for each $\lambda \in \mathbb{R}$.

The remaining part works as in the finite-dimensional case: one verifies that $\|\cdot\|_{L^p}$ is a norm and that $L^p$ is complete under this norm. The approximations constructed in Theorem 6 converge uniformly outside of sets of arbitrarily small measure by Egorov's theorem. This proves the density of simple functions. See [Lan93, Chapters VI and VII] for full proofs. □

## 2.3 Bochner integral

**13 Definition.** For each simple function $f = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n$ set

$$\int_A f \mathrm{d}\mu := \sum_{n=1}^{N} \mu(A_n) x_n.$$

**14 Theorem.** *There is a unique bounded linear operator*

$$\int_A \mathrm{d}\mu : L^1(A; E) \to E$$

*extending the linear mapping in Definition 13 such that for each $f \in L^1(A; E)$,*

$$\left\| \int_A f \mathrm{d}\mu \right\| \le \int_A \|f\| \mathrm{d}\mu =: \|f\|_{L^1(A;E)}.$$

*This operator is called the Bochner integral.*

*Proof.* The inequality is easy to see for simple functions:

$$\left\| \int_A f \mathrm{d}\mu \right\| = \left\| \sum_{n=1}^{N} \mu(A_n) x_n \right\| \le \sum_{n=1}^{N} \mu(A_n) \|x_n\| = \int_A \|f\| \mathrm{d}\mu.$$

It implies that the integral in Definition 13 is a bounded operator with operator norm $\le 1$. The set of simple functions is dense in $L^1(A; E)$ by Theorem 12. Therefore, the integral has a unique extension to $L^1(A; E)$ with the same operator norm. □

**15 Reflection.** Is there a dominated convergence theorem for Bochner integrals, and how would you prove it?

- If $f_n \to f$ a.s. and $\|f_n\| \le g \in L^1(\mu)$, then $\int \|f_n - f\| \mathrm{d}\mu \to 0$ by the scalar dominated convergence theorem, and $\int (f_n - f) \mathrm{d}\mu \to 0$ by the continuity of the integral.

**16 Reflection.** What are some examples of separable and non-separable Banach spaces?

- If $(A, \mathcal{A}, \mu)$ is countably generated and $E$ is separable, then $L^p(\mu; E)$ is separable for each $p \in [1, \infty)$, but $L^\infty(\mu; E)$ is in general not separable.

- $L(E; F)$ is in general not separable, even if $E$ and $F$ are separable Hilbert spaces. See exercises.

- If $K$ is a compact topological space, then $C(K)$ is separable. If $K$ is a locally compact Polish space, then $C_0(K)$ is separable, but $C(K)$ is in general not separable.
- For any Polish space $E$, $D(\mathbb{R}_+; E)$ with $\|\cdot\|_\infty$ is a non-separable Banach space, and $D(\mathbb{R}_+; E)$ with the Skorokhod $J_1$ metric is a Polish space, but addition is discontinuous.
- There are non-separable Hilbert spaces, but they do not appear often in practice.

### 2.4 Other notions of integrals

**17 Definition.** Let $f : A \to E$ be weakly measurable.

- The *Pettis integral* of $f$ over $A$, if it exists, is an element $x \in E$ satisfying

$$\langle x, x' \rangle = \int \langle f, x' \rangle \mathrm{d}\mu, \qquad \forall x' \in E'.$$

  Let $PI_1(A; E)$ be the space of Pettis-integrable functions, and let $\hat{P}I_1(A; E)$ be the completion of this space with respect to the norm

$$\|f\|_{PI_1(A;E)} = \sup \left\{ \int_A |\langle f, x^* \rangle| \mathrm{d}\mu : x^* \in E^*, \|x^*\| \leq 1 \right\}.$$

- The *Gelfand integral* is another name for the Pettis integral in the case where $E$ is the dual of a Banach space and carries the weak-* topology.
- The *Dunford integral* of $f$ over $A$, if it exists, is an element $x'' \in E''$ satisfying

$$\langle x'', x' \rangle = \int \langle f, x' \rangle \mathrm{d}\mu, \qquad \forall x' \in E'.$$

**18 Remark.** One always has

Bochner integrability $\Rightarrow$ Pettis integrability $\Rightarrow$ Dunford integrability.

In general these implications are strict: there are functions which are Dunford but not Pettis integrable [Rya02, p. 52] and functions which are Pettis but not Bochner integrable [Rya02, p. 53].

### 2.5 Literature

Most of this section is taken from [Van08] and [Jen16]. Integration of Banach-space valued functions is treated in detail in [Lan93] and [AB06].

## 3 Gaussian random variables

**19 Setting.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $E$ be a Banach space, and let $X : \Omega \to E$ be strongly measurable. More generally, all random variables are assumed to have a strongly measurable version.

**20 Reflection.** Our aim is to define Gaussian random variables with values in a Banach space. What are our options?

- Densities are not helpful because there is no Lebesgue measure on the Banach space.
- We could require all finite-dimensional projections to be Gaussian.
- We could require all one-dimensional projections to be Gaussian. (This turns out to be equivalent.)
- We could use some limiting procedure, where we build up vector-valued Gaussians from scalar ones.

**21 Definition.** $X$ is called centered Gaussian if for every $x^* \in E^*$ there exists $q \geq 0$ such that

$$\mathbb{E}\left[\exp\left(-i\xi\langle X, x^*\rangle\right)\right] = \exp\left(-\tfrac{1}{2}\xi^2 q\right), \qquad \xi \in \mathbb{R}.$$

**22 Reflection.** Can we write $\langle X, x^*\rangle \sim N(0, \sigma^2)$ for some $\sigma$?

- If $q = 0$ then $\langle X, x^*\rangle \sim \delta_0$;
- If $q > 0$ then $\langle X, x^*\rangle \sim N(0, q)$, i.e., $q = \sigma^2$.

## 3.1 Moments

**23 Theorem** (Moments)**.** *Let $X$ be centered Gaussian.*

*(i) Covariance: the law of $X$ is determined uniquely by the bilinear form*

$$q(x_1^*, x_2^*) = \mathbb{E}[\langle X, x_1^*\rangle\langle X, x_2^*\rangle], \qquad x_1^*, x_2^* \in E^*.$$

*(ii) Fernique's theorem: there exists $\beta > 0$ such that*

$$\mathbb{E}\left[\exp\left(\beta\|X\|^2\right)\right] < \infty.$$

*(iii) Covariance operator: there exists $Q \in L(E^*, E)$ such that*

$$q(x_1^*, x_2^*) = \langle Q x_1^*, x_2^*\rangle.$$

*(iv) Kahane-Khintchine inequality: for all $p, q \in [1, \infty)$ there is a universal constant $K_{p,q}$ such that $\|X\|_{L^p(\Omega;E)} \leq K_{p,q}\|X\|_{L^q(\Omega;E)}$.*

We need some auxiliary lemmas.

**24 Lemma** (Injectivity of the Fourier transform)**.** *If*

$$\mathbb{E}\left[\exp\left(-i\langle X_1, x^*\rangle\right)\right] = \mathbb{E}\left[\exp\left(-i\langle X_2, x^*\rangle\right)\right], \qquad x^* \in E^*,$$

*then $X_1$ is identically distributed to $X_2$.*

*Proof.* As $X_1$ and $X_2$ have strongly measurable versions, there is no loss of generality in assuming $E$ to be separable. Denote by $\mathrm{Cyl}(E)$ the algebra of cylinder sets on $E$, i.e., the initial algebra with respect to $E^*$. The laws of $X_1$ and $X_2$ agree on $\mathrm{Cyl}(E)$ because the Fourier-transform on the space of measures on $\mathbb{R}^n$ is injective. (It is even bijective on the much larger space of tempered distributions.) By Dynkin's Lemma it remains to show $\sigma(\mathrm{Cyl}(E)) = \mathcal{B}(E)$. As $\mathcal{B}(E)$ is generated by open sets, it is sufficient to show that open sets are contained in $\sigma(\mathrm{Cyl}(E))$. As open sets are countable unions of closed balls, it is sufficient to show that closed balls are contained in $\sigma(\mathrm{Cyl}(E))$. This can be seen by taking a norming sequence $x^* \in E^*$ and writing

$$B_R(x_0) = \bigcap_{n\in\mathbb{N}} \big\{x : |\langle x - x_0, x_n^*\rangle| \leq R\big\} \in \sigma(\mathrm{Cyl}(E)). \qquad \square$$

**25 Lemma** (Rotations)**.** *If $X_1$ and $X_2$ are iid centered Gaussian, then $Y_1 := (X_1 + X_2)/\sqrt{2}$ and $Y_2 := (X_1 - X_2)/\sqrt{2}$ are iid and have the same distribution as $X_1$ and $X_2$.*

*Proof.* We check that $(X_1, X_2)$ and $(Y_1, Y_2)$ have the same Fourier transforms. Setting $q(x^*) = \mathbb{E}[\langle X_1, x^*\rangle^2]$ we have

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(-i\langle Y_1, x_1^*\rangle - i\langle Y_2, x_2^*\rangle\Big)\Big] \\
&\quad= \mathbb{E}\Big[\exp\Big(-i\langle X_1 + X_2, x_1^*\rangle/\sqrt{2} - i\langle X_1 - X_2, x_2^*\rangle/\sqrt{2}\Big)\Big] \\
&\quad= \mathbb{E}\Big[\exp\Big(-i\langle X_1, x_1^* + x_2^*\rangle/\sqrt{2}\Big)\Big]\,\mathbb{E}\Big[\exp\Big(-i\langle X_2, x_1^* - x_2^*\rangle/\sqrt{2}\Big)\Big] \\
&\quad= \exp\Big(-\tfrac14 q(x_1^* + x_2^*)\Big)\exp\Big(-\tfrac14 q(x_1^* - x_2^*)\Big) \\
&\quad= \exp\Big(-\tfrac12 q(x_1^*) - \tfrac12 q(x_2^*)\Big) = \mathbb{E}\Big[\exp\Big(-i\langle X_1, x_1^*\rangle - i\langle X_2, x_2^*\rangle\Big)\Big]. \qquad \square
\end{aligned}$$

*Proof of Theorem 23.*

(i): This follows from Lemma 24.

(ii): We follow [Van08, Theorem 4.3]. By the triangle inequality, one has for each $x, y \in E$ that

$$\min\{\|x\|, \|y\|\} \geq \tfrac12\big(\|x+y\| - \|x-y\|\big).$$

Therefore, one has for each $0 < s \leq t$ that

$$\big(\|x-y\| \leq \sqrt{2}s, \|x+y\| \geq \sqrt{2}t\big) \Rightarrow \big(\|x\| \geq (t-s)/\sqrt{2}, \|y\| \geq (t-s)/\sqrt{2}\big).$$

Let $Y$ be an iid copy of $X$, possibly on an extended probability space. By Lemma 25 we have for each $s, t > 0$ that

$$\begin{aligned}
\mathbb{P}\left[\|X\| \leq s\right]\mathbb{P}\left[\|Y\| > t\right] &= \mathbb{P}\Big[\|(X-Y)/\sqrt{2}\| \leq s, \|(X+Y)/\sqrt{2}\| > t\Big] \\
&\leq \mathbb{P}\Big[\|X\| > (t-s)/\sqrt{2}, \|Y\| > (t-s)/\sqrt{2}\Big] = \mathbb{P}\Big[\|X\| > (t-s)/\sqrt{2}\Big]^2.
\end{aligned}$$

Choose $r > 0$ such that $\mathbb{P}[\|X\| \leq r] \geq 2/3$. For each $n \in \mathbb{N}$ set

$$t_0 = r, \qquad t_{n+1} = r + \sqrt{2}t_n, \qquad \alpha_n = \frac{\mathbb{P}[\|X\| > t_n]}{\mathbb{P}[\|X\| \leq r]}.$$

The above inequality with $s = r$ and $t = t_n$ yields

$$\alpha_n \leq \alpha_{n-1}^2 \leq \cdots \leq \alpha_0^{2^n} \leq \left( \frac{1 - \frac{2}{3}}{\frac{2}{3}} \right)^{2^n} = 2^{-2^n}.$$

Moreover, one can check that explicitly,

$$t_n = \left(1 + \sqrt{2} + \cdots + (\sqrt{2})^n\right) r = \frac{(\sqrt{2})^{n+1} - 1}{\sqrt{2} - 1} r \leq (\sqrt{2})^{n+4} r.$$

Set $\gamma = (\log 2)/2^6$. For any $t \geq r$ and $n \in \mathbb{N}$ such that $r \in [t_n, t_{n+1})$,

$$\begin{aligned}
\mathbb{P}[\|X\| > t] &\leq \mathbb{P}[\|X\| > t_n] \leq 2^{-2^n} \leq 2^{-(t_{n+1}/r)^2/32} \\
&= \exp\big( - 2\gamma t_{n+1}^2/r^2\big) \leq \exp\big( - 2\gamma t^2/r^2\big).
\end{aligned}$$

Integration by parts shows that for each $\beta \leq \gamma/r^2$

$$\begin{aligned}
\mathbb{E}[\exp(\beta\|X\|^2)] &\leq \mathbb{E}[\exp(\gamma\|X\|^2/r^2)] \\
&= \mathbb{P}[\|X\| \leq r] \exp(\gamma) + \int_r^\infty \mathbb{P}[\|X\| > t] \mathrm{d}\big(\exp(\gamma t^2/r^2)\big) \\
&\leq \exp(\gamma) + \int_r^\infty \exp\big( - 2\gamma t^2/r^2\big) \mathrm{d}\big(\exp(\gamma t^2/r^2)\big) \\
&= \exp(\gamma) + \int_1^\infty \exp\big( - 2\gamma s^2\big) \mathrm{d}\big(\exp(\gamma s^2)\big) \\
&= \exp(\gamma) + 2\gamma \int_1^\infty s \exp\big( - \gamma s^2\big) \mathrm{d}s =: K < \infty.
\end{aligned} \tag{1}$$

(iii): The operator

$$Q : E^* \to E, \qquad x^* \mapsto \mathbb{E}[\langle X, x^* \rangle X],$$

has the desired properties because $X \in L^2(\Omega; E)$.

(iv): By Jensen's inequality suffices to show that $\|\cdot\|_{L^{2n}(\Omega;E)} \leq K_{2n,1} \|\cdot\|_{L^1(\Omega;E)}$ for each $n \in \mathbb{N}$. Set $r = 3\|X\|_{L^1(\Omega;E)}$. By Chebychev,

$$\mathbb{P}[\|X\| \leq r] = 1 - \mathbb{P}[\|X\| > r] \geq 1 - \mathbb{E}[\|X\|]/r \geq 2/3.$$

Therefore, the choice of $r$ is admissible in the proof of Fernique's theorem, and we get from (1) that $\mathbb{E}[\exp(\gamma/r^2\|X\|^2)] \leq K$. For each $x \in \mathbb{R}_+$ one has

$$e^{\gamma\frac{x^2}{r^2}} \geq \frac{\gamma^n}{n!} \frac{x^{2n}}{r^{2n}}, \qquad x^{2n} \leq r^{2n} n! \gamma^{-n} e^{\gamma/r^2 x^2}.$$

Therefore,

$$\begin{aligned}
\|X\|_{L^{2n}(\Omega;E)} &\leq r \left( n! \gamma^{-n} r^{2n} \mathbb{E}\left[ e^{\gamma/r^2\|X\|^2} \right] \right)^{1/(2n)} \\
&\leq \|X\|_{L^1(\Omega;E)} \underbrace{3 \left(n! \gamma^{-n} r^{2n} K\right)^{1/(2n)}}_{K_{2n,1}}.
\end{aligned}$$

□

## 3.2 Convergence

**26 Theorem** (Convergence)**.** *Let $X_n$ be centered Gaussian $E$-valued random variables in $E$, let $Y_n$ be symmetric independent $E$-valued random variables, let $S_n = \sum_{i=1}^n Y_i$, and let $S$ and $X$ be $E$-valued random variables.*

*(i) Itō-Nisio theorem: the following are equivalent:*

- *for all $x^* \in E^*$ we have $\langle S_n, x^*\rangle \to \langle S, x^*\rangle$ almost surely;*
- *for all $x^* \in E^*$ we have $\langle S_n, x^*\rangle \to \langle S, x^*\rangle$ in probability;*
- *$S_n \to S$ almost surely;*
- *$S_n \to S$ in probability.*

*If $S \in L^p(\Omega; E)$ for some $p \in [1,\infty)$ then $S_n \to S$ in $L^p(\Omega; E)$.*

*(ii) If $\langle X_n, x^*\rangle \to \langle X, x^*\rangle$ in probability for each $x^*$, then $X$ is Gaussian.*

*(iii) If $X_n \to X$ in probability, then $X_n \to X$ in $L^p(\Omega; E)$ for each $p \in [1,\infty)$.*

*Proof.* (i): See [Van08, Theorem 2.17].

(ii): If for each $x^* \in E^*$, $\langle X_n, x^*\rangle \to \langle X, x^*\rangle$ in probability, then a subsequence converges almost surely, then by dominated convergence

$$\begin{aligned}\mathbb{E}[\exp(-i\xi\langle X, x^*\rangle)] &= \lim_{n\to\infty} \mathbb{E}[\exp(-i\xi\langle X_n, x^*\rangle)] = \lim_{n\to\infty} \exp(-\tfrac{1}{2}\xi^2\langle Q_n x^*, x^*\rangle) \\ &= \exp(-\tfrac{1}{2}\xi^2 q(x^*)),\end{aligned}$$

where $q(x^*) := \lim_{n\to\infty}\langle Q_n x^*, x^*\rangle$ exists because left-hand side converges.

(iii): See [Van08, Theorem 4.15]. □

## 3.3 Series representations

**27 Definition.** Let $X$ be centered Gaussian. A reproducing kernel Hilbert space of $X$ is a Hilbert space $H$ together with continuous embedding $i : H \to E$ such that

$$\mathbb{E}[\langle X, x^*\rangle\langle X, y^*\rangle] = \langle i^*x^*, i^*y^*\rangle, \qquad x^*, y^* \in E^*.$$

**28 Lemma.** *Let $X$ be centered Gaussian with covariance operator $Q$. The following are reproducing kernel Hilbert spaces of $X$:*

- *$H_X$ is the closure of the subspace $\{\langle X, x^*\rangle : x^* \in E^*\}$ of $L^2(\Omega)$, and $i_X(\langle X, x^*\rangle) = \mathbb{E}[\langle X, x^*\rangle X]$.*
- *$H_Q$ is the completion of* $\operatorname{ran}(Q)$ *with respect to the scalar product $\langle Qx^*, Qy^*\rangle_{H_Q} := \langle Qx^*, y^*\rangle_{E,E^*}$, and $i_Q(Qx^*) = Qx^*$.*
- *Assume that $E$ is a Hilbert space, and set $\Sigma = Q^{1/2}$. $H_\Sigma$ is the range of $\Sigma$ with $\langle\cdot,\cdot\rangle_{H_\Sigma} = \langle\Sigma^{-1}\cdot, \Sigma^{-1}\cdot\rangle$, where $\Sigma^{-1}$ is the pseudo-inverse of $\Sigma$, and $i_\Sigma(\Sigma x) = \Sigma x$.*

*The spaces are separable and isomorphic, and $i_X(H_X) = i_Q(H_Q) = i_\Sigma(H_\Sigma)$.*

**29 Remark.**

- The notations $H_Q$ and $H_\Sigma$ will not create any ambiguity because we will always use the letters $Q$ and $\Sigma$ to denote the covariance operator and its square root.

- The spaces $H_Q$ and $H_\Sigma$ can be constructed for any non-negative symmetric linear operators $Q$ and $\Sigma$, regardless of whether they correspond to the covariance of a Gaussian random variable or not.

*Proof.* We only show the statements for $H_X$, leaving the remaining statements as an exercise. Let $\mu_X$ be the law of $X$ on $E$, and let $E_0$ be a separable closed subspace of $E$ containing the essential range of $X$. Then $L^2(E_0, \mu_X)$ is separable because $\mathcal{B}(E_0)$ is countably generated. As every $X^* \in E^*$ can be seen as an element of $L^2(E, \mu_X)$, we have

$$L^2(E_0, \mu_X) \cong L^2(E, \mu_X) \supseteq \overline{E^*}^{L^2(E,\mu_X)} \cong H_X,$$

where the right-most isomorphism maps $x^* \in E^*$ to $\langle X, x^* \rangle \in H_X$. Therefore, $H_X$ is separable. The mapping $i_X$ is continuous because $X \in L^2(\Omega; E)$. The relation $\langle i_X h, x^* \rangle = \mathbb{E}[h\langle X, x^* \rangle]$ shows that its adjoint is given by $i_X^*(x^*) = \langle X, x^* \rangle$. Moreover, $i_X i_X^*(x^*) = \mathbb{E}[\langle X, x^* \rangle X] = Qx^*$, and $i_X$ is injective because $i_X^*$ has dense range. □

**30 Theorem** (Karhunen-Loève expansion)**.** *Let $X$ be centered Gaussian, and let $(\gamma_n)_{n\geq 1}$ be an orthonormal basis of $H_X$. Then $(\gamma_n)_{n\geq 1}$ is an iid standard Gaussian sequence and*

$$X = \sum_{n\geq 1} \gamma_n i_X \gamma_n,$$

*where convergence holds almost surely and in $L^p$ for each $p \in [1, \infty)$.*

*Proof.* $H_X$ consists of Gaussian random variables because $L^2$-limit of Gaussians are Gaussian by Theorem 26.(ii). The $L^2$-orthogonality of $(\gamma_n)_{n\geq 1}$ implies (via the injectivity of the Fourier transform) that $\gamma_n$ are mutually independent. For every $x^* \in E^*$,

$$\langle X, x^* \rangle = \sum_{n\geq 1} \gamma_n \mathbb{E}[\gamma_n \langle X, x^* \rangle] = \sum_{n\geq 1} \gamma_n \langle i_x \gamma_n, x^* \rangle,$$

where the first equality is the expansion of $\langle X, x^* \rangle$ with respect to the basis $(\gamma_n)_{n\geq 1}$ of $H_X$. The Itō-Nisio Theorem 26.(i) turns this weak convergence into $L^2(\Omega; E)$-convergence. □

## 3.4 Gaussians on Hilbert spaces

**31 Theorem** (Gaussians on Hilbert spaces)**.** *Let $H$ be a separable Hilbert space.*

*(i) Sazanov's theorem: $Q \in L(H)$ is the covariance operator of a centered Gaussian $H$-valued random variable $X$ if and only if $Q$ is symmetric, non-negative definite, and $Q$ is nuclear.*[1]

[1]See Section 4 below.

*(ii) In this case the Karhunen-Loève expansion of $X$ is*

$$X = \sum_{n\geq 1} \underbrace{\langle X, e_n/\sqrt{\lambda_n}\rangle_H}_{\gamma_n} \underbrace{e_n\sqrt{\lambda_n}}_{i_X\gamma_n},$$

*where $(e_n)_{n\geq 1}$ is an orthonormal basis of $(\ker Q)^\perp$ satisfying $Qe_n = \lambda_n e_n$ for eigenvalues $\lambda_n \in (0,\infty)$.*

*Proof.* (i): Let $X$ be centered Gaussian with covariance operator $Q$. Then $Q$ is symmetric and non-negative, and

$$\operatorname{Tr}(Q) = \sum_{n\geq 1}\langle Qe_n, e_n\rangle_H = \sum_{n\geq 1}\mathbb{E}[\langle X, e_n\rangle_H^2] = \mathbb{E}[\|X\|_H^2] < \infty.$$

Conversely, assume that $Q \in L_1(H)$ is symmetric and non-negative. As $Q$ is compact, there is an orthonormal basis $(e_n)_{n\geq 1}$ of $H$ such that $Qe_n = \lambda_n e_n$. Set

$$X = \sum_{n\geq 1}\gamma_n\sqrt{\lambda_n}e_n.$$

This series converges in $L^2(\Omega; E)$ because

$$\sum_{n\geq 1}\|\gamma_n\sqrt{\lambda_n}e_n\|_{L^2(\Omega;H)}^2 = \sum_{n\geq 1}\lambda_n\mathbb{E}[\gamma_n^2]\ \|e_n\|_H^2 = \sum_{n\geq 1}\lambda_n < \infty,$$

and the limit is Gaussian by Theorem 26.(ii).

(ii): The random variables $\gamma_n$ are orthonormal because

$$\mathbb{E}[\gamma_m\gamma_n] = \frac{\langle Qe_m, e_n\rangle_H}{\sqrt{\lambda_m\lambda_n}} = \mathbb{1}_{m=n}.$$

Let $(f_n)_{n\geq 1}$ be a basis of $\ker Q$. By the definition of $H_X$, the span of the random variables $\langle X, e_n\rangle$ and $\langle X, f_n\rangle$ is dense in $H_X$. But $\langle X, f_n\rangle = 0$ in $L^2(\Omega)$ because $\mathbb{E}[\langle X, f_n\rangle^2] = \langle Qf_n, f_n\rangle_H = 0$, which implies that the span of $\langle X, e_n\rangle_H$ is dense in $H_X$. Therefore, $(\gamma_n)_{n\geq 1}$ is an orthonormal basis. One has $i_X\gamma_n = \sqrt{\lambda_n}e_n$ because

$$\begin{aligned}\langle i_X\gamma_n, e_m\rangle_H &= \mathbb{E}[\gamma_n\langle X, e_m\rangle_H] = \mathbb{E}[\langle X, e_n/\sqrt{\lambda_n}\rangle_H\langle X, e_m\rangle_H] \\ &= \langle Qe_n, e_m\rangle_H/\sqrt{\lambda_n} = \mathbb{1}_{m=n}\sqrt{\lambda_n} = \langle\sqrt{\lambda_n}e_n, e_m\rangle_H.\end{aligned}$$

□

### 3.5 Literature

This section is a rearrangement of results in [Van08] and [Hai09].

## 4 Tensor products and operator ideals

**32 Setting.** Let $D$, $E$, $F$, and $G$ be Banach spaces, let $M \subseteq E$ and $N \subseteq F$ be finite-dimensional subspaces, let $H$, $H_1$, and $H_2$ be Hilbert spaces, let $K$ be a compact topological space, let $1 \leq p \leq \infty$, and let $1/p + 1/p' = 1$.

## 4.1 Tensor products and tensor norms

**33 Definition** (Algebraic tensor products)**.**

- Let $L(E, F; G)$ denotes the space of bounded bilinear mappings from $E \times F$ to $G$ with norm

$$\|T\|_{L(E,F;G)} = \sup\{\|T(x, y)\| : x \in E, y \in F, \|x\| \leq 1, \|y \leq 1\}.$$

- The algebraic tensor product $E \otimes F$ is the linear span of the functionals $x \otimes y \in L(E, F; \mathbb{R})^*$ given by $(x \otimes y)(T) := T(x, y)$, $T \in L(E, F; G)$.

**34 Definition** (Tensor norms)**.**

- A tensor norm $\alpha$ assigns to each pair of Banach spaces $(E, F)$ a norm on the algebraic tensor product $E \otimes F$ (shorthand: $E \otimes_\alpha F$ and $E \hat{\otimes}_\alpha F$ for the completion) such that the following properties hold:
  - (i) $\alpha$ is reasonable, i.e., one has for all $x \in E$, $y \in F$, $x^* \in E^*$, $y^* \in F^*$ that

$$\|x \otimes y\|_{E \otimes_\alpha F} \leq \|x\|_E \|y\|_F, \qquad \|x^* \otimes y^*\|_{(E \otimes_\alpha F)^*} \leq \|x^*\|_{E^*} \|y^*\|_{F^*}.$$

  - (ii) $\alpha$ satisfies the metric mapping property, i.e., for any $T_1 \in L(E_1; F_1)$ and $T_2 \in L(E_2; F_2)$ one has

$$\|T_1 \otimes T_2\|_{L(E_1 \otimes_\alpha E_2; F_1 \otimes \alpha F_2)} \leq \|T_1\| \|T_2\|.$$

  - (iii) $\alpha$ is finitely generated, i.e.,

$$\|u\|_{E \otimes_\alpha F} = \inf \left\{ \|u\|_{M \otimes_\alpha N} : u \in M \otimes N, \dim M, \dim N < \infty \right\}.$$

- For any tensor norm $\alpha$, the dual norm $\alpha'$ is the tensor norm

$$\|u\|_{E \otimes_{\alpha'} F} = \inf_{M,N} \sup_{\substack{v \in M^* \otimes N^* \\ \|v\|_{M^* \otimes_\alpha N^*} \leq 1}} |\langle u, v \rangle|,$$

  where the infimum is taken over all finite-dimensional spaces $M$ and $N$ such that $u \in M \otimes N$, and $\langle \cdot, \cdot \rangle$ denotes the pairing (trace) of $M \otimes N$ and $M^* \otimes N^*$.

- The projective tensor norm is given by

$$\|u\|_{E \otimes_\pi F} := \inf \left\{ \sum_{i=1}^n \|x_i\| \|y_i\| : u = \sum_{i=1}^n x_i \otimes y_i, x_i \in E, y_i \in F, n \in \mathbb{N} \right\}.$$

- The injective tensor norm is given by

$$\|u\|_{E \otimes_\varepsilon F} := \sup \left\{ |\langle u, x^* \otimes y^* \rangle| : x \in X^*, y \in Y^*, \|x^*\| \leq 1, \|y^*\| \leq 1 \right\}.$$

**35 Remark.**

- **Class exercise.** A norm $\alpha$ on $E \otimes F$ is reasonable if and only if it is sandwiched between the injective and projective norms, i.e.,

$$\| \cdot \|_{E\otimes_\varepsilon F} \le \| \cdot \|_{E\otimes_\alpha F} \le \| \cdot \|_{E\otimes_\pi F}.$$

  See [Rya02, Proposition 6.1].

- **Class exercise.** Any reasonable norm $\alpha$ satisfies

$$\|x \otimes y\|_{E\otimes_\alpha F} = \|x\|_E \|y\|_F, \qquad \|x^* \otimes y^*\|_{(E\otimes_\alpha F)^*} = \|x^*\|_{E^*} \|y^*\|_{F^*}.$$

  See [Rya02, Proposition 6.1].

- **Class exercise.** The projective tensor product linearizes bilinear mappings in the sense that [Rya02, Theorem 2.9]

$$L(E, F; G) \cong L(E\hat{\otimes}_\pi F; G).$$

- One has $\epsilon' = \pi$, $\pi' = \epsilon$, and $\alpha'' = \alpha$ for each tensor norm $\alpha$.

- The following are embeddings of norm at most one: [Rya02, Chapter 7.1]

$$E\hat{\otimes}_\alpha F \hookrightarrow (E^*\hat{\otimes}_{\alpha'} F^*)^*, \qquad E^*\hat{\otimes}_\alpha F^* \hookrightarrow (E\hat{\otimes}_{\alpha'} F)^*.$$

## 4.2 Operator ideals and bilinear forms

**36 Definition** (Operator ideals and bilinear forms)**.**

- An operator ideal is an assignment to each pair of Banach spaces $E, F$ of Banach spaces $\mathfrak{A}(E; F) \subseteq L(E; F)$ such that the following properties hold:

  (i) for every $x^* \in E$ and $y^* \in F$ one has $x^* \otimes y^* \in A(E; F)$ and $\|x^* \otimes y^*\|_{\mathfrak{A}(E;F)} \le \|x^*\|\|y^*\|$, and

  (ii) for every $T \in L(D; E)$, $S \in \mathfrak{A}(E; F)$, and $R \in L(F; G)$ one has $RST \in \mathfrak{A}(D; G)$ and $\|RST\|_{\mathfrak{A}(D;G)} \le \|T\|_{L(D;E)} \|S\|_{\mathfrak{A}(E;F)} \|R\|_{L(F;G)}$.

- For any tensor norm $\alpha$, the set of $\alpha$-nuclear forms is

$$\mathfrak{N}_\alpha(E, F; \mathbb{R}) := \text{range}\left(J_\alpha : E^*\hat{\otimes}_\alpha F^* \to L(E, F; \mathbb{R})\right) \subseteq L(E, F; \mathbb{R}),$$

  and the operator ideal of $\alpha$-nuclear operators is

$$\mathfrak{N}_\alpha(E; F) := \text{range}\left(J_\alpha : E^*\hat{\otimes}_\alpha F \to L(E; F)\right) \subseteq L(E; F).$$

  The nuclear norms are defined such that the mappings $J_\alpha$ are metric surjections onto their range. Forms/operators in $\mathfrak{N}_\alpha$ are called nuclear for $\alpha = \pi$ and approximable for $\alpha = \varepsilon$.

- For any tensor norm $\alpha$, the set of $\alpha$-integral forms is

$$\mathfrak{I}_\alpha(E, F; \mathbb{R}) := (E\hat{\otimes}_{\alpha'} F)^* \subseteq L(E, F; \mathbb{R}) = L(E; F^*),$$

  and the operator ideal of $\alpha$-integral operators is

$$\mathfrak{I}_\alpha(E; F) := \mathfrak{I}_\alpha(E; F^{**}) \cap L(E, F) \subseteq \mathfrak{I}_\alpha(E; F^{**}) := (E\hat{\otimes}_{\alpha'} F^*)^*.$$

  Forms/operators in $\mathfrak{I}_\alpha$ are called integral for $\alpha = \pi$ and bounded for $\alpha = \varepsilon$.

**37 Remark.**

- If $\alpha$ dominates $\beta$, there are continuous embeddings $\mathfrak{N}_\alpha \hookrightarrow \mathfrak{N}_\beta$, $\mathfrak{I}_\alpha \hookrightarrow \mathfrak{I}_\beta$.
- For any tensor norm $\alpha$, there is a continuous embedding $\mathfrak{N}_\alpha \hookrightarrow \mathfrak{I}_\alpha$; see [Rya02, Section 8.1].

## 4.3 Trace

**38 Definition** (Trace)**.**

- The trace is the continuous linear mapping of norm one given by

$$E^* \hat{\otimes}_\pi E \to \mathbb{R}, \qquad x^* \otimes x \mapsto \langle x^*, x \rangle.$$

  If the mapping $J_\pi$ of Definition 36 is injective, then the trace is also defined for nuclear operators $\mathfrak{N}_\pi(E; E)$.

**39 Remark.**

- The mappings $J_\alpha$ of Definition 36 are injective for each tensor norm $\alpha$ if either $E^*$ or $F$ have the approximation property [Rya02, Proposition 8.7].
- A Banach space $E$ has the approximation property if for every compact subset $K \subset E$ and every $\epsilon > 0$ there exists a finite-rank operator $S \in L(E)$ such that

$$\sup_{x \in K} \|x - Sx\| < \epsilon.$$

  All Hilbert spaces and all Banach spaces with a Schauder basis have the approximation property [Rya02, Section 4.1].
- Approximable operators are compact, i.e., there is an embedding $\mathfrak{N}_\epsilon \hookrightarrow \mathfrak{K}$, because approximable operators are the closure of finite-rank operators under a stronger topology than $L$; see [Rya02, Section 4.1]. The converse holds, i.e., all compact operators from $E$ to $F$ are approximable, if either $E^*$ or $F$ has the approximation property; see [Rya02, Corollary 4.13].

## 4.4 Some further tensor norms and operator ideals

**40 Definition** (Some sequence spaces)**.**

- Let $\ell_p(E)$ be the set of sequences $(x_n)$ in $E$ such that

$$\|(x_n)\|_p := \left( \sum_{n=1}^{\infty} \|x_n\|_E^p \right)^{1/p} < \infty.$$

  Then $\ell_p(E)$ with this norm is a Banach space. We write $\ell_p$ for $\ell_p(\mathbb{R})$.
- Let $\ell_p^w(E)$ be the set of sequences $(x_n)$ in $E$ such that $\langle x^*, x_n \rangle \in \ell_p$ for each $x^* \in E^*$, and set

$$\|(x_n)\|_p^w := \sup \{ \|(\langle x^*, x_n \rangle)\|_p : x^* \in E^*, \|x^*\| \leq 1 \}.$$

  Then $\ell_p^w(E)$ is a Banach space because it is isomorphic to $L(c_0, E)$ for $p = 1$, to $L(\ell_{p'}, E)$ for $1 < p < \infty$, and to $\ell_\infty(E)$ for $p = \infty$.

- Let $c_0(E)$ be the set of sequences $(x_n) \in \ell_\infty(E)$ s.t. $\lim_{n\to\infty} \|x_n\| = 0$.

**41 Definition** (More tensor norms and operator ideals)**.**

- The Chevet-Saphar tensor norms are given by

$$\|u\|_{E\otimes_{d_p}F} = \inf\left\{ \|(x_i)\|_{p'}^w \|(y_i)\|_p : u = \sum_{i=1}^n x_i \otimes y_i, x_i \in E, y_i \in F, n \in \mathbb{N} \right\},$$
$$\|u\|_{E\otimes_{g_p}F} = \inf\left\{ \|(x_i)\|_p \|(y_i)\|_{p'}^w : u = \sum_{i=1}^n x_i \otimes y_i, x_i \in E, y_i \in F, n \in \mathbb{N} \right\}.$$

  Here $g$ stands for gauche and $d$ for droite. Forms/operators in $\mathfrak{I}_{d_p} =: P_p$ are called $p$-summing, and in $\mathfrak{N}_{g_p} =: L_p$ $p$-nuclear

- The Hilbertian tensor norms is given by

$$\|u\|_{E\otimes_{w_p}F} = \inf\left\{ \|(x_i)\|_p^w \|(y_i)\|_{p'}^w : u = \sum_{i=1}^n x_i \otimes y_i, x_i \in E, y_i \in F, n \in \mathbb{N} \right\}.$$

  Forms/operators in $\mathfrak{I}_{w_p}$ are called $p$-Hilbertian, in $\mathfrak{I}_{w_2}$ Hilbertian, and in $\mathfrak{I}_{w_p'}$ $p$-dominated.

- The operator ideal $HS(H_1; H_2)$ of Hilbert-Schmidt operators is defined for Hilbert spaces $H_1$ and $H_2$ and consists of all $T \in L(H_1; H_2)$ such that

$$\|T\|^2_{HS(H_1;H_2)} := \sum_{b\in B} \|Tb\|^2_{H_2} < \infty,$$

  for some (and hence all) orthonormal bases $B$ of $H_1$.

**42 Remark.**

- One has $d_1 = g_1 = \pi$. Moreover, for $p \le q$ one has $d_p \ge d_q$ and $g_p \ge g_q$. See [Rya02, Proposition 6.6]. Furthermore, $d_2' = g_2$ and $g_2' = d_2$ by [Rya02, Equation 7.7].

- **Class exercise.** If $H_1$ and $H_2$ are Hilbert spaces with orthonormal bases $(e_i)$ and $(f_j)$, respectively, then $H_1\hat{\otimes}_{w_2}H_2$ is a Hilbert space with orthonormal basis $(e_i \otimes f_j)$.

- If $H_1$ and $H_2$ are Hilbert spaces, then the following are isometries: [DF92, Proposition 11.6 and Section 26.6]

$$HS(H_1; H_2) \cong P_2(H_1; H_2) \cong L_2(H_1; H_2) \cong H_1\hat{\otimes}_{w_2}H_2.$$

- There are isomorphisms

| | |
|---|---|
| $L^1(A; E) \cong L^1(A)\hat{\otimes}_\pi E,$ | [Rya02, Example 2.19] |
| $\hat{P}I_1(A; E) \cong L^1(A)\hat{\otimes}_\varepsilon E,$ | [Rya02, Proposition 3.13] |
| $L^2(A; H) \cong L^2(A)\hat{\otimes}_{w_2} H,$ | **Class exercise.** |
| $C(K)\hat{\otimes}_\varepsilon E \cong C(K; E)$ | [Rya02, Section 3.2]. |

## 4.5 Operator ideals on Hilbert spaces

**43 Lemma.** *Let $H$ be a Hilbert space with orthonormal basis $\mathbb{B}$, let $\lambda : \mathbb{B} \to \mathbb{R}$ be a function, and let $T : D(T) \subseteq H \to H$ be the diagonal linear operator given by*

$$Tb = \lambda_b b, \qquad D(T) = \left\{ h \in H : \sum_{b \in \mathbb{B}} |\lambda_b|^2 \langle b, h \rangle_H^2 < \infty \right\}.$$

*We consider $B$ as a measure space with the counting measure $\#$. Then the following statements hold:*

*(i)* $T \in L(H)$ *iff* $\lambda \in L^\infty(\mathbb{B})$, *and* $\|T\|_{L(H)} = \|\lambda\|_{L^\infty(\mathbb{B})}$.

*(ii)* $T \in L_1(H)$ *iff* $\lambda \in L^1(\mathbb{B})$, *and* $\|T\|_{L_1(H)} = \|\lambda\|_{L^1(\mathbb{B})}$.

*(iii)* $T \in L_2(H)$ *iff* $\lambda \in L^2(\mathbb{B})$, *and* $\|T\|_{L_2(H)} = \|\lambda\|_{L^2(\mathbb{B})}$.

*Proof.* See exercises. $\square$

## 4.6 Literature

A good introduction to topological tensor products is [Rya02]. Further details can be found in [DF92]. Lemma 43 is taken from [Jen16].

# 5 Stochastic integration

**44 Setting.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+})$ be a filtered probability space satisfying the usual conditions, let $E$ be a Banach space, let $H$ and $U$ be separable Hilbert spaces, let $\mathcal{P}$ be the predictable $\sigma$-algebra on $\mathbb{R}_+ \times \Omega$, let $Q \in L(U)$, let $\Sigma = Q^{1/2}$, let $U_\Sigma$ be the Hilbert space $\Sigma U$ with scalar product $\langle u, v \rangle_{\Sigma U} = \langle \Sigma^{-1} u, \Sigma^{-1} v \rangle_U$, where $\Sigma^{-1}$ is the pseudo-inverse of $\Sigma$, and let $i_\Sigma : U_\Sigma \to U$ be given by $i_\Sigma(\Sigma u) = \Sigma u$.

## 5.1 Martingales

**45 Lemma** (Conditional expectation)**.** *For any $\sigma$-algebra $\mathcal{G}$ contained in $\mathcal{F}$, there exists a unique bounded linear operator of norm one,*

$$L^1(\mathcal{F}; E) \to L^1(\mathcal{G}; E), \qquad X \mapsto \mathbb{E}[X|\mathcal{G}],$$

*such that one has for each $A \in \mathcal{G}$ that*

$$\mathbb{E}[\mathbb{1}_A X] = \mathbb{E}[\mathbb{1}_A \mathbb{E}[X|\mathcal{G}]].$$

*Proof.* We use the fact that the conditional expectation of scalar random variables is a bounded linear operator of norm at most one,

$$\mathbb{E}[\cdot|\mathcal{G}] : L^1(\mathcal{F}) \to L^1(\mathcal{G}).$$

Recall that $L^1(\mathcal{F}; E) = L^1(\mathcal{F}) \hat{\otimes}_\pi E$. We claim that the desired operator is

$$T := \mathbb{E}[\cdot|\mathcal{G}] \hat{\otimes}_\pi I_E : L^1(\mathcal{F}) \hat{\otimes}_\pi E \to L^1(\mathcal{G}) \hat{\otimes}_\pi E.$$

By the metric mapping property of tensor products,

$$\|T\| \leq \|\mathbb{E}[\cdot|\mathcal{G}]\|_{L(L^1(\mathcal{F});L^1(\mathcal{G}))}\|I_E\|_{L(E;E)} \leq 1.$$

For each $A \in \mathcal{G}$ and $X = \mathbb{1}_B \otimes x \in L^1(\mathcal{F}) \otimes E$ one has

$$\mathbb{E}[\mathbb{1}_A X] = \mathbb{E}[\mathbb{1}_A \mathbb{1}_B]x = \mathbb{E}[\mathbb{1}_A \mathbb{E}[\mathbb{1}_B|\mathcal{G}]]x = \mathbb{E}[\mathbb{1}_A \mathbb{E}[\mathbb{1}_B|\mathcal{G}]x] = \mathbb{E}[\mathbb{1}_A TX].$$

By continuity, this relation holds for all $X \in L^1(\mathcal{F})\hat{\otimes}_\pi E$. Therefore, $T$ satisfies the properties stated in the lemma. Moreover, the last calculation shows that $T$ is determined uniquely on simple tensors and by continuity on all tensors. □

**46 Definition.** Let $M : \mathbb{R}_+ \times \Omega \to E$ be an $(\mathcal{F}_t)$-adapted process, and let $p \in [1, \infty]$.

- $M$ is called a martingale if $M_t \in L^1(\mathcal{F}_t; E)$ and $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ holds for all $0 \leq s \leq t$. If in addition $M_t \in L^p(\mathcal{F}_t; E)$ for all $t \in \mathbb{R}_+$, then we call $M$ a $L^p$-martingale.
- We write $\mathcal{M}^p$ for the space of uniformly integrable $L^p$-martingales, i.e., all martingales of the form $M_t = \mathbb{E}[M_\infty|\mathcal{F}_t]$ for some $M_\infty \in L^p(\Omega; E)$, endowed with the norm $\|M\|_{\mathcal{M}^p} := \|M_\infty\|_{L^p}$. The subspace of continuous martingales is denoted by $\mathcal{M}^p_c$.

**47 Lemma** (Doob's maximal inequality)**.** *Let $M$ be a right-continuous $E$-valued $(\mathcal{F}_t)$-martingale. Then the following statements hold.*

*(i) For each $p \in [1, \infty)$, $\|M_t\|^p$ is a real-valued $(\mathcal{F}_t)$-submartingale.*

*(ii) For each $p \in [1, \infty)$ and $\epsilon > 0$,*

$$\mathbb{P}\left[\|M\|_{L^\infty(\mathbb{R}_+;E)} > \epsilon\right] \leq \frac{1}{\epsilon^p}\|M\|^p_{L^\infty(\mathbb{R}_+;L^p(\Omega;H))}.$$

*(iii) For each $p \in (1, \infty)$,*

$$\|M\|_{L^p(\Omega;L^\infty(\mathbb{R}_+;E))} \leq \frac{p}{p-1}\|M\|_{L^\infty(\mathbb{R}_+;L^p(\Omega;E))}.$$

*Proof.* As the essential range of $M$ is separable, we assume without loss of generality that $E$ is separable. Let $(x_n^*)$ be a norming sequence in $E^*$. Then

$$\begin{aligned}\mathbb{E}\big[\|M_t\|\big|\mathcal{F}_s\big] &\geq \sup_n \mathbb{E}\big[\langle M_t, x_n^*\rangle\big|\mathcal{F}_s\big] = \sup_n \big\langle \mathbb{E}[M_t|\mathcal{F}_s], x_n^*\big\rangle \\ &= \sup_n \langle M_s, x_n^*\rangle = \|M_s\|.\end{aligned}$$

This shows that $\|M_t\|$ is a submartingale. By Jensen's inequality, $\|M_t\|^p$ is a submartingale for each $p \in [1, \infty)$. The remaining statements are Doob's maximal inequalities for non-negative real-valued submartingales. □

## 5.2 Brownian motion

**48 Definition** (Brownian motion)**.** Let $Q \in L_1(U)$.

- A $Q$-Brownian motion is a continuous process $W : \mathbb{R}_+ \times \Omega \to U$ with independent increments satisfying $W_0 = 0$ and $W_t - W_s \sim N(0, (t-s)Q)$ for all $0 \le s \le t$.
- A $Q$-Brownian motion with respect to $(\mathcal{F}_t)$ is a $Q$-Brownian motion $W$ such that $W_t$ is $\mathcal{F}_t$-measurable and $W_t - W_s$ is independent of $\mathcal{F}_s$, for all $0 \le s \le t$.

**49 Definition** (Cylindrical Brownian motion)**.** Let $Q \in L(U)$.

- A cylindrical $Q$-Brownian motion is a bounded linear mapping $W : L^2(\mathbb{R}_+; U) \to L^2(\Omega)$, whose values are centered Gaussians such that

  $$\mathbb{E}[W(\mathbb{1}_{[0,s]} \otimes u) W(\mathbb{1}_{[0,t]} \otimes v)] = \min(s,t)\langle Qu, v\rangle_U,$$

  for each $s, t \in \mathbb{R}_+$ and $u, v \in U$. We put

  $$W_t u := W(\mathbb{1}_{[0,t]} \otimes u), \qquad t \in \mathbb{R}_+, u \in U.$$

- A cylindrical $Q$-Brownian motion with respect to $(\mathcal{F}_t)$ is a $Q$-Brownian motion $W$ such that $W_t u$ is $\mathcal{F}_t$-measurable and $W_t u - W_s u$ is independent of $\mathcal{F}_s$, for all $0 \le s \le t \le T$ and $u \in U$.

**50 Remark.**

- If $W$ is a $Q$-Brownian motion, then (the extension of) the linear mapping

  $$\mathbb{1}_{[0,t]} \otimes u \mapsto \langle u, W_t\rangle_U \tag{2}$$

  is a cylindrical $Q$-Brownian motion.
- Conversely, let $W$ be a cylindrical $Q$-Brownian motion with $Q \in L_1(U)$. Then for each $t \ge 0$ the mapping

  $$W_t : U \to L^2(\Omega), \qquad u \mapsto W(\mathbb{1}_{[0,t]} \otimes u)$$

  satisfies

  $$\|W_t\|^2_{HS(U;L^2(\Omega))} = \sum_{b \in B} \|W_t b\|^2_{L^2(\Omega)} = t^2 \sum_{b \in B} \langle Qb, b\rangle_U \le t^2 \|Q\|_{L_1(U)} < \infty,$$

  where $B$ is an orthonormal basis of $H$. Therefore, $W_t \in HS(U; L^2(\Omega)) = L^2(\Omega)\hat{\otimes}_{w_2} U = L^2(\Omega; U)$. Selecting a continuous version of $(W_t)_{t \in \mathbb{R}_+}$ using the Kolmogorov continuity theorem gives a $Q$-Brownian motion, which represents $W$ as in (2).

## 5.3 Construction of the stochastic integral

**51 Definition.**

- The set $\mathcal{E}$ of elementary processes is the linear span of the set of predictable processes $X : \mathbb{R}_+ \times \Omega \to L(U, H)$ of the form
$$X = \mathbb{1}_{(t_1,t_2]} \otimes \mathbb{1}_A \otimes u \otimes h, \tag{3}$$
where $0 \le t_1 \le t_2$, $A \in \mathcal{F}_{t_1}$, $u \in U$, and $h \in H$.

- If $W$ is $Q$-Brownian motion with respect to $(\mathcal{F}_t)$ for some $Q \in L_1(U)$ and $X \in \mathcal{E}$ is as in (3), we define the stochastic integral
$$\int_0^\infty X_s dW_s := \mathbb{1}_A \langle u, W_{t_2} - W_{t_1} \rangle_H h.$$

- If $W$ is cylindrical $Q$-Brownian motion with respect to $(\mathcal{F}_t)$ for some $Q \in L(H)$ and $X \in \mathcal{E}$ is as in (3), we define the stochastic integral
$$\int_0^\infty X_s dW_s := \mathbb{1}_A (W_{t_2} u - W_{t_1} u) h.$$

- The Hilbert space $\mathcal{N}_W^2$ of stochastically square integrable processes consists of all predictable processes $X : \mathbb{R}_+ \times \Omega \to L_2(U_\Sigma; H)$ such that
$$\|X\|_{\mathcal{N}_W^2} := \|X\|_{L^2((\mathbb{R}_+ \times \Omega, \mathcal{P}, dt \otimes \mathbb{P}); L_2(U_\Sigma; H))} < \infty.$$

- The complete metric space $\mathcal{N}_W$ of stochastically integrable processes consists of all predictable processes $X : \mathbb{R}_+ \times \Omega \to L_2(U_\Sigma; H)$ such that
$$\mathbb{P}[\|X\|_{L^2(\mathbb{R}_+; L_2(U_\Sigma; H))} < \infty] = 1,$$
endowed with the Ky Fan metric
$$(X, Y) \mapsto \mathbb{E}[\|X - Y\|_{L^2(\mathbb{R}_+; L_2(U_\Sigma; H))} \wedge 1],$$
which induces the topology of convergence in probability.

**52 Remark.** Note that the integral with respect to cylindrical Brownian motion extends the integral with respect to Brownian motion.

**53 Lemma** (Density)**.** *The following statements holds.*

*(i)* $\mathcal{E}$ *is dense in* $\mathcal{N}_W^2$.

*(ii)* $\mathcal{N}_W^2$ *is dense in* $\mathcal{N}_W$.

*Proof.* (i): $\mathcal{E}$ is interpreted as a subspace of $\mathcal{N}_W^2$ via the (not necessarily injective) mapping
$$\mathcal{E} \to \mathcal{N}_W^2, \qquad \mathbb{1}_{(t_1,t_2]} \otimes \mathbb{1}_A \otimes u \otimes h \mapsto \mathbb{1}_{(t_1,t_2]} \otimes \mathbb{1}_A \otimes i_\Sigma^* u \otimes h,$$
where $i_\Sigma^* u \in U_\Sigma$ acts on $U_\Sigma$ via $\langle \cdot, \cdot \rangle_{U_\Sigma}$. We have
$$\mathcal{N}_W^2 = L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, dt \otimes \mathbb{P}) \hat{\otimes}_{w_2} U_\Sigma \hat{\otimes}_{w_2} H.$$

The span of the functions $\mathbb{1}_{(t_1,t_2]} \otimes \mathbb{1}_A$ is dense in $L^2(\mathbb{R}_+ \times \Omega, \mathcal{P}, dt \otimes \mathbb{P})$, and $i_\Sigma^*(U)$ is dense in $U_\Sigma$ because $i_\Sigma$ is injective. Therefore, $\mathcal{E}$ is dense in $\mathcal{N}_W^2$.

(ii): $\mathcal{N}_W^2$ is dense in $\mathcal{N}_W$. To see this, let $X \in \mathcal{N}_W$. For each $n \in \mathbb{N}$ set $T^n = \inf\{t \geq 0 : \|\mathbb{1}_{(0,t]}X\|_{N_W^2} > n\}$. By the right-continuity of $(\mathcal{F}_t)$, $T^n$ is a stopping time for each $n$. Moreover, $X^n := \mathbb{1}_{(0,T^n]}X \in \mathcal{N}_W^2$ and $\|X^n\|_{\mathcal{N}_W^2} \leq n$. Let $B$ be the set of probability one where $\|X\|_{L^2(\mathbb{R}_+;L_2(U_\Sigma;H))} < \infty$. Then $\lim_{n\to\infty} T^n = \infty$ on $B$ and $\lim_{n\to\infty} X^n = X$ on $\mathbb{R}_+ \times B$. By dominated convergence, $\lim_{n\to\infty} \|X^n - X\|_{L^2(\mathbb{R}_+;L_2(U_\Sigma;H))} = 0$ on $B$. As almost sure convergence implies convergence in probability, $X^n$ converges to $X$ in probability. Therefore, $\mathcal{N}_W^2$ is dense in $\mathcal{N}_W$. $\square$

**54 Theorem** (Stochastic integral)**.** *Let $Q \in L(U)$ and let $W$ be cylindrical $Q$-Brownian motion.*

*(i) The stochastic integral of Definition 51 has a unique extension to an isometry*

$$\mathcal{N}_W^2 \to L^2(\Omega; H), \qquad X \mapsto \int_0^\infty X_s dW_s.$$

*(ii) The stochastic integral of Definition 51 has a unique extension to a uniformly continuous linear mapping*

$$\mathcal{N}_W \to L^0(\Omega; H), \qquad X \mapsto \int_0^\infty X_s dW_s.$$

*Proof.* (i): We have to show that the stochastic integral of Definition 51 is an isometry on the subspace $\mathcal{E}$ of $\mathcal{N}_W^2$. Any $X \in \mathcal{E}$ can be written as a finite sum of the form

$$X = \sum_{k,i} \mathbb{1}_{(t_k,t_{k+1}]} \otimes \mathbb{1}_{A_{k,i}} \otimes u_{k,i} \otimes h_{k,i},$$

where $(t_k)$ are non-negative and strictly increasing, $A_{k,i} \in \mathcal{F}_{t_k}$, $u_{k,i} \in U$, and $h_{k,i} \in H$. Then

$$\mathbb{E}\left[\left\|\int_0^\infty X_s dW_s\right\|_H^2\right] = \sum_{k,l,i,j} \mathbb{E}\left[\left\langle \mathbb{1}_{A_{k,i}} W(\mathbb{1}_{(t_k,t_{k+1}]} \otimes u_{k,i})h_{k,i}, \mathbb{1}_{A_{l,j}} W(\mathbb{1}_{(t_l,t_{l+1}]} \otimes u_{l,j})h_{l,j}\right\rangle_H\right].$$

Conditioning on $\mathcal{F}_{t_k} \vee \mathcal{F}_{t_l}$ shows that

$$\mathbb{E}\left[\left\|\int_0^\infty X_s dW_s\right\|_H^2\right] = \sum_{k,l,i,j} \mathbb{E}[\mathbb{1}_{A_{k,i}} \mathbb{1}_{A_{l,j}}] \langle \mathbb{1}_{(t_k,t_{k+1}]}, \mathbb{1}_{(t_l,t_{l+1}]}\rangle_{L^2(\mathbb{R}_+)} \langle Qu_{k,i}, u_{l,j}\rangle_U \langle h_{k,i}, h_{l,j}\rangle_H.$$

By the reproducing property of the Hilbert space $U_\Sigma$ (see Definition 27) one has $\langle Qu_{k,i}, u_{l,j}\rangle = \langle i_\Sigma^* u_{k,i}, i_\Sigma^* u_{l,j}\rangle_{U_\Sigma}$, and therefore

$$\mathbb{E}\left[\left\|\int_0^\infty X_s dW_s\right\|_H^2\right] = \mathbb{E}[\mathbb{1}_{A_{k,i}} \mathbb{1}_{A_{l,j}}] \langle \mathbb{1}_{(t_k,t_{k+1}]}, \mathbb{1}_{(t_l,t_{l+1}]}\rangle_{L^2(\mathbb{R}_+)} \langle i_\Sigma^* u_{k,i}, i^*\Sigma u_{l,j}\rangle_{U_\Sigma} \langle h_{k,i}, h_{l,j}\rangle_H$$

$$= \|X\|^2_{\mathcal{N}^2_W}.$$

Therefore, the stochastic integral of Definition 51 is an isometry on the subspace $\mathcal{E}$ of $\mathcal{N}^2_W$. As $\mathcal{E}$ is dense in $\mathcal{N}^2_W$ by Lemma 53, there is a unique extension to an isometry $\mathcal{N}^2_W \to L^2(\Omega; H)$. This shows (i).

(ii): This follows from Theorem 55.(ii) below. □

**55 Theorem** (Integral process)**.** *Let $Q \in L(U)$ and let $W$ be cylindrical $Q$-Brownian motion. For each $t \geq 0$ set $\int_0^t X_s dW_s = \int_0^\infty \mathbb{1}_{(0,t]}(s) X_s dW_s$.*

*(i) The stochastic integral has a unique extension to an isometry*

$$\mathcal{N}^2_W \to \mathcal{M}^2_c, \qquad X \mapsto \int_0^\cdot X_s dW_s.$$

*(ii) The stochastic integral has a unique extension to a uniformly continuous linear mapping*

$$\mathcal{N}_W \to L^0(\Omega; C(\mathbb{R}_+; H)), \qquad X \mapsto \int_0^\cdot X_s dW_s.$$

*Proof.* (i): Let $X \in \mathcal{E}$ be given by (3). Then the scalar process $(W_t u)_{t \in \mathbb{R}_+}$ is mean-square continuous and has a continuous version by Kolmogorov's continuity theorem. Thus, the process $(M_t)_{t \in \mathbb{R}_+}$ has a continuous version because

$$M_t = \mathbb{1}_A (W_{t_2 \wedge t} u - W_{t_1 \wedge t} u) h.$$

Moreover, setting $M_\infty = \int_0^\infty X_s dW_s$, one has $\mathbb{E}[M_\infty | \mathcal{F}_t] = M_t$, as can be seen from

$$\mathbb{E}[M_\infty | \mathcal{F}_t] = \mathbb{E}[\mathbb{1}_A (W_{t_2} u - W_{t_1} u) h | \mathcal{F}_t] = \begin{cases} 0, & \text{if } t \in [0, t_1], \\ \mathbb{1}_A (W_t u - W_{t_1} u) h, & \text{if } t \in [t_1, t_2], \\ \mathbb{1}_A (W_{t_2} u - W_{t_1} u) h, & \text{if } t \in [t_2, \infty). \end{cases}$$

As $M_\infty \in L^2(\Omega; H)$, we have shown that $M$ has a version in $\mathcal{M}^2_c$. The stochastic integral of (i) defines an isometry $\mathcal{E} \to \mathcal{M}^2_c$ thanks to Theorem 54.(i). The space $\mathcal{M}^2_c$ is complete because it is a closed subspace of $\mathcal{M}^2$ thanks to Doob's maximal inequality (Lemma 47). Therefore, the stochastic integral has a unique extension to an isometry $\mathcal{N}^2_W \to \mathcal{M}^2_c$. This establishes (i).

(ii): Let $X \in \mathcal{N}^2_W$, let $\epsilon, \delta > 0$, and let

$$T = \inf \left\{ t \in [0, \infty) : \int_0^t \|X_s\|^2_{L_2(U_\Sigma; H)} ds > \delta \right\}.$$

Then one has by Chebychev's inequality, Doob's maximal inequality, and Itō's isometry, that

$$\begin{aligned} &\mathbb{P}\left[ \sup_{t \in [0,\infty]} \left\| \int_0^t X_s dW_s \right\|_H > \epsilon \right] \\ &\quad = \mathbb{P}\left[ \sup_{t \in [0,\infty]} \left\| \int_0^t X_s dW_s \right\|_H > \epsilon, \int_0^\infty \|X_s\|^2_{L_2(U_\Sigma; H)} ds > \delta \right] \end{aligned}$$

$$
\begin{aligned}
&+ \mathbb{P}\left[\sup_{t\in[0,\infty]}\left\|\int_0^t X_s dW_s\right\|_H > \epsilon, \int_0^\infty \|X_s\|^2_{L_2(U_\Sigma;H)} ds \le \delta\right] \\
&\le \mathbb{P}\left[\int_0^\infty \|X_s\|^2_{L_2(U_\Sigma;H)} ds > \delta\right] + \mathbb{P}\left[\sup_{t\in[0,\infty]}\left\|\int_0^t \mathbb{1}_{(0,T]}(s) X_s dW_s\right\|_H > \epsilon\right] \\
&\le \mathbb{P}\left[\int_0^\infty \|X_s\|^2_{L_2(U_\Sigma;H)} ds > \delta\right] + \frac{1}{\epsilon^2}\mathbb{E}\left[\int_0^\infty \|\mathbb{1}_{(0,T]} X_s\|^2_{L_2(U_\Sigma;H)} ds\right] \\
&= \mathbb{P}\left[\int_0^\infty \|X_s\|^2_{L_2(U_\Sigma;H)} ds > \delta\right] + \frac{1}{\epsilon^2}\mathbb{E}\left[\delta \wedge \int_0^\infty \|X_s\|^2_{L_2(U_\Sigma;H)} ds\right].
\end{aligned}
$$

This inequality is called Lenglart's inequality. It shows that the integral is continuous at zero with respect to the topology of convergence in probability. By linearity, it is uniformly continuous, and we have shown (ii). □

**56 Reflection.** For nuclear $Q$ one may use the following (larger) class of elementary processes $X$ to define the stochastic integral: [DZ14; Jen16; PR07]

$$X = \mathbb{1}_{(s,t]} \otimes \mathbb{1}_A \otimes Y,$$

where $0 \le s < t$, $A \in \mathcal{F}_s$, and $Y \in L(U;H)$. How are the spaces $L_2(U_\Sigma;H)$ and $L(U;H)$ related?

- $L(U,H)$ is continuously embedded in $L_2(U_\Sigma;H)$ because

  $$
  \begin{aligned}
  \|Y \circ i_\Sigma\|_{L^2(U_\Sigma;H)} &\le \|Y\|_{L(U;H)} \|i_\Sigma\|_{L_2(U_\Sigma;U)} = \|Y\|_{L(U;H)} \|\Sigma\|_{L_2(U)} \\
  &= \|Y\|_{L(U;H)} \|Q\|^{1/2}_{L_1(U)}.
  \end{aligned}
  $$

  To see this, let $\mathbb{B}$ be an orthonormal basis of $U$ and let $\lambda : \mathbb{B} \to [0,\infty)$ such that $Qb = \lambda_b b$. Then $\mathbb{B}_\Sigma := \{\Sigma b : b \in \lambda^{-1}((0,\infty))\}$ is an orthonormal basis of $U_\Sigma$. Then

  $$
  \begin{aligned}
  \|Y \circ i_\Sigma\|^2_{L_2(U_\Sigma;H)} &= \sum_{b\in\mathbb{B}_\Sigma} \|Y i_\Sigma b\|^2_H \le \|Y\|^2_{L(U;H)} \sum_{b\in\mathbb{B}_\Sigma} \|i_\Sigma b\|^2_U \\
  &= \|Y\|^2_{L(U;H)} \|i_\Sigma\|^2_{L_2(U_\Sigma;U)}, \\
  \|i_\Sigma\|^2_{L_2(U_\Sigma;U)} &= \sum_{b\in\mathbb{B}_\Sigma} \|b\|^2_U = \sum_{b\in\lambda^{-1}((0,\infty))} \|\Sigma b\|^2_U = \sum_{b\in\mathbb{B}} \|\Sigma b\|^2_U = \|\Sigma\|^2_{L_2(U)}, \\
  \|\Sigma\|^2_{L_2(U)} &= \sum_{b\in\mathbb{B}} \|\Sigma b\|^2_U = \sum_{b\in\mathbb{B}} \langle Qb, b\rangle_U = \sum_{b\in\mathbb{B}} \lambda_b = \|Q\|_{L_1(U)}.
  \end{aligned}
  $$

  The embedding may be strict, as $L_2(U_\Sigma;H)$ may contain unbounded operators on $U$.

## 5.4 Properties of the stochastic integral

**57 Lemma** (Localization)**.** *Let $X \in \mathcal{N}_W$ and let $T$ be a stopping time satisfying $\mathbb{P}[T < \infty] = 1$. Then*

$$\int_0^{\cdot\wedge T} X_s dW_s = \int_0^{\cdot} \mathbb{1}_{(0,T]}(s) X_s dW_s.$$

*Proof.* If $T$ is simple and $X \in \mathcal{E}$, this follows by inspection. In the general case, there is a sequence of simple stopping times $T_n \downarrow T$ and a sequence of elementary processes $X^m \xrightarrow{p} X$. Then

$$\begin{aligned}\int_0^{\cdot\wedge T} X_s dW_s &= \operatorname*{plim}_{n\to\infty} \int_0^{\cdot\wedge T^n} X_s dW_s = \operatorname*{plim}_{n\to\infty} \operatorname*{plim}_{m\to\infty} \int_0^{\cdot\wedge T^n} X_s^m dW_s \\ &= \operatorname*{plim}_{n\to\infty} \operatorname*{plim}_{m\to\infty} \int_0^{\cdot} \mathbb{1}_{(0,T^n]}(s) X_s^m dW_s = \int_0^{\cdot} \mathbb{1}_{(0,T]}(s) X_s dW_s. \qquad \square\end{aligned}$$

**58 Remark.** Lemma 57 allows one to define the stochastic integral for the localized classes of integrands $(\mathcal{N}_W^2)_{\mathrm{loc}}$ and $(\mathcal{N}_W)_{\mathrm{loc}}$.

**59 Lemma.** *Let $K$ be a topological space, let $E$ and $F$ be normed vector spaces, and let $f : E \to F$ be uniformly continuous on bounded sets. Then the following mapping is continuous:*

$$L^0(\Omega; C(K;E)) \to L^0(\Omega; C(K;F)), \qquad X \mapsto f \circ X.$$

*Proof.* Let $\epsilon > 0$ and let $X^n \to X$ in $L^0(\Omega; C(K;E))$. Then $(X^n)$ is bounded in probability (tight). Therefore, there is $R > 0$ such that

$$\sup_{n\in\mathbb{N}} \mathbb{P}\left[\|X^n\|_{C(K;E)} > R\right] < \epsilon/3, \qquad \mathbb{P}\left[\|X\|_{C(K;E)} > R\right] < \epsilon/3.$$

As $f$ is uniformly continuous on the ball of radius $R$, there is $\delta > 0$ such that

$$\big(\|x-y\|_E < \delta, \|x\|_E \le R, \|y\|_E \le R\big) \Rightarrow \|f(x) - f(y)\|_F < \epsilon/3.$$

As $X^n \to X$ in probability, one may choose $n$ large enough such that

$$\mathbb{P}\left[\|X^n - X\|_{C(K;E)} > \delta\right] < \epsilon/3.$$

Taken together, this yields

$$\begin{aligned}&\mathbb{P}\left[\|f\circ X^n - f\circ X\|_{C(K;F)} > \epsilon\right] \\ &\quad \le \mathbb{P}\left[\|f\circ X^n - f\circ X\|_{C(K;F)} > \epsilon, \|X^n\|_{C(K;E)} \le R, \|X\|_{C(K;E)} \le R\right] \\ &\qquad + \mathbb{P}\left[\|X^n\|_{C(K;E)} > R\right] + \mathbb{P}\left[\|X\|_{C(K;E)} > R\right] \\ &\quad \le \mathbb{P}\left[\|X^n - X\|_{C(K;E)} > \delta\right] + 2\epsilon/3 < \epsilon. \qquad \square\end{aligned}$$

**60 Lemma** (Itō's formula)**.** *Let $X_0 : \Omega \to H$ be strongly $\mathcal{F}_0$-measurable, let $F : \mathbb{R}_+ \times \Omega \to H$ be predictable and a.s. Bochner-integrable on $\mathbb{R}_+$, let $B : \mathbb{R}_+ \times \Omega \to L_2(U_\Sigma; H)$ be stochastically integrable, let*

$$X = X_0 + \int_0^{\cdot} F_s ds + \int_0^{\cdot} B_s dW_s,$$

*let $f : H \to H$ be a $C^2$ function such that $f, f_x, f_{xx}$ are uniformly continuous on bounded subsets of $H$, and let $\mathbb{U}$ be an orthonormal basis of $U_\Sigma$. Then*

$$\begin{aligned}f(X) = f(X_0) &+ \int_0^{\cdot} \Big(f_x(X_s)F_s + \frac{1}{2}\sum_{u\in\mathbb{U}} f_{xx}(X_s)(B_s u, B_s u)\Big) ds \\ &+ \int_0^{\cdot} f_x(X_s) B_s dW_s.\end{aligned}$$

*Proof.* By Itō's formula in finite dimensions, the statement of the lemma holds for elementary processes $F$ and $B$. The general case follows because all terms in Itō's formula are continuous with respect to the topology of convergence in probability thanks to Lemma 59. $\square$

**61 Lemma** (Stochastic Fubini theorem)**.** *Let $(E, \mathcal{E}, \mu)$ be a finite measure space, and let $X \in L^1(E; \mathcal{N}_W^2)$. Then the following iterated integrals are well-defined and coincide almost surely,*

$$\int_E \int_0^\infty X_s^x dW_s \mu(dx) = \int_0^\infty \int_E X_s^x \mu(dx) dW_s.$$

*Proof.* Recalling that $L^1(E; \mathcal{N}_W^2) = L^1(E) \hat{\otimes}_\pi \mathcal{N}_W^2$, let $X = \mathbb{1}_B Y$ for some $B \in \mathcal{E}$ and $Y \in \mathcal{N}_W^2$. Then the random function

$$\int_0^\infty X dW = \mathbb{1}_B \int_0^\infty Y dW : \Omega \times E \to H$$

is $\mathcal{F} \otimes \mathcal{E}$-measurable, and the process

$$\int_E X d\mu = \int_E \mathbb{1}_B d\mu \, Y : \mathbb{R}_+ \times \Omega \to L_2(U_\Sigma; H)$$

is predictable. By Minkowski's inequality and Itō's isometry,

$$\begin{aligned}
\left\| \int_E \int_0^\infty X dW d\mu \right\|_{L^2(\Omega; H)} &\leq \int_E \left\| \int_0^\infty X dW \right\|_{L^2(\Omega; H)} d\mu \\
&= \int_E \|X\|_{\mathcal{N}_W^2)} \, d\mu = \|X\|_{L^1(E; \mathcal{N}_W^2)}.
\end{aligned}$$

Moreover, the following integrals coincide,

$$\int_E \int_0^\infty X dW d\mu = \mu(B) \int_0^\infty Y dW = \int_0^\infty \int_E X^d \mu dW,$$

By continuity and density, this extends to all $X \in L^1(E; \mathcal{N}_W^2)$. $\square$

**62 Lemma** (Burkholder-Davis-Gundy inequality)**.** *For each $p \geq 2$, there exists a constant $c_p$ such that*

$$\left\| \int_0^\cdot X_s dW_s \right\|_{L^p(\Omega; L^\infty(\mathbb{R}_+; H))} \leq c_p \, \|X\|_{L^p(\Omega; L^2(\mathbb{R}_+; L_2(U_\Sigma; H)))} .$$

*Proof.* As the case $p = 2$ is covered by Doob's maximal inequality (Lemma 47.(iii)) and Itō's isometry, assume $p > 2$, and let $X$ be a predictable process in $L^p(\Omega; L^2(\mathbb{R}_+; L_2(U_\Sigma; H)))$. By Doob's maximal inequality,

$$\begin{aligned}
\left\| \int_0^\cdot X_s dW_s \right\|_{L^p(\Omega; L^\infty(\mathbb{R}_+; H))} &\leq \frac{p}{p-1} \left\| \int_0^\cdot X_s dW_s \right\|_{L^\infty(\mathbb{R}_+; L^p(\Omega; H))} \\
&\leq \frac{p}{p-1} \left\| \int_0^\infty X_s dW_s \right\|_{L^p(\Omega; H)} .
\end{aligned}$$

Set $Y$ equal to the continuous modification of $\int_0^\cdot X_s dW_s$. For each $y \in H$, set

$$\begin{aligned} f(y) &= \|y\|_H^p, \\ f_y(y) &= p\|y\|_H^{p-2}\langle y, \cdot\rangle_H, \\ f_{yy}(y) &= p(p-2)\|y\|^{p-4}\langle y, \cdot\rangle_H \langle y, \cdot\rangle_H + p\|y\|^{p-2}\langle, \cdot\rangle_H, \end{aligned}$$

and note that for each $y \in H$

$$\|f_{yy}(y)\|_{L(H,H;\mathbb{R})} \leq p(p-1)\|y\|_H^{p-2}.$$

By Itō's formula, letting $\mathbb{U}$ be an orthonormal basis of $U_\Sigma$,

$$\|Y_\infty\|_{L^p(\Omega;H)}^p = \mathbb{E}\left[\int_0^\infty f_y(Y_s)X_s dW_s + \frac{1}{2}\int_0^\infty \sum_{u\in\mathbb{U}} f_{yy}(Y_s)(X_s u, X_s u)ds\right].$$

Assume temporarily that $Y$ is bounded. Then the stochastic integral above is a uniformly integrable martingale, and the bound on $f_{yy}$ gives

$$\begin{aligned} \|Y_\infty\|_{L^p(\Omega;H)}^p &\leq \frac{p(p-1)}{2}\mathbb{E}\left[\int_0^\infty \sum_{u\in\mathbb{U}} \|Y_s\|_H^{p-2}\|X_s u\|_H^2 ds\right] \\ &= \frac{p(p-1)}{2}\mathbb{E}\left[\int_0^\infty \|Y_s\|_H^{p-2}\|X_s\|_{L_2(U_\Sigma;H}^2 ds\right] \\ &\leq \frac{p(p-1)}{2}\mathbb{E}\left[\|Y_\infty\|_H^{p-2}\int_0^\infty \|X_s\|_{L_2(U_\Sigma;H}^2 ds\right]. \end{aligned}$$

By Hölder's inequality with exponents $p/(p-2)$ and $p/2$,

$$\begin{aligned} \|Y_\infty\|_{L^p(\Omega;H)}^p &\leq \frac{p(p-1)}{2}\mathbb{E}\left[\|Y_\infty\|_H^p\right]^{(p-2)/p}\mathbb{E}\left[\left(\int_0^\infty \|X_s\|_{L_2(U_\Sigma;H}^2 ds\right)^{p/2}\right]^{2/p} \\ &= \frac{p(p-1)}{2}\|Y_\infty\|_{L^p(\Omega;H)}^{p-2}\|X\|_{L^p(\Omega;L^2(\mathbb{R}_+;L_2(U_\Sigma;H)))}^2. \end{aligned}$$

Diving through $\|Y_\infty\|_{L^p(\Omega;H)}^{p-2}$ and taking the square root, one gets

$$\|Y_\infty\|_{L^p(\Omega;H)} \leq \sqrt{\frac{p(p-1)}{2}}\|X\|_{L^p(\Omega;L^2(\mathbb{R}_+;L_2(U_\Sigma;H)))}.$$

In the general case where $Y$ is not bounded, let for each $n \in \mathbb{N}$

$$T_n = \inf\{t \geq 0 : \|Y_t\|_H \geq n\}.$$

Then $Y^{T_n}$ is bounded, and $\|Y_\infty^{T_n}\|_H = \|Y_\infty\|_H \wedge n$ is monotonically increasing and converges to $\|Y_\infty\|_H$. By the monotone convergence theorem, the $L^p(\Omega;H)$-norm of $\|Y_\infty^{T_n}\|_H$ converges to the $L^p(\Omega;H)$-norm of $\|Y_\infty\|_H$. Thus, the estimate remains valid in the general case. □

## 5.5 Literature

The construction of the stochastic integral in Section 5.3 uses a smaller class of elementary processes than many other works [DZ14; PR07; Jen16]. This has the advantage that the stochastic integrals with respect to cylindrical and normal Brownian motion are immediately seen to be identical because they coincide for our small class of elementary processes.

Our presentation of Itō's formula (Lemma 60) and the stochastic Fubini theorem (Lemma 61) is similar to the one in [DZ14]. The Burkholder-Davis-Gundy inequality and its proof (see Lemma 62) is a corrected version of [DZ14, Theorem 4.36].

# 6 Stochastic evolution equations with Lipschitz coefficients

**63 Setting.** Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{R}_+})$ be a stochastic basis, let $H$ and $U$ be separable Hilbert spaces, let $Q \in L(U)$ by non-negative symmetric, let $\Sigma = Q^{1/2}$, let $\Sigma^{-1}$ be the pseudo-inverse of $\Sigma$, let $U_\Sigma = \Sigma(U)$ with $\langle u, v\rangle_{U_\Sigma} = \langle \Sigma^{-1}u, \Sigma^{-1}v\rangle_U$, and let $W$ be cylindrical $Q$-Brownian motion with respect to $(\mathcal{F}_t)$.

Let $T > 0$, let $\mathcal{P}$ be the predictable $\sigma$-algebra on $[0,T] \times \Omega$, let $\xi$ be an $H$-valued $\mathcal{F}_0$-measurable random variable, let $A : D(A) \subseteq H \to H$ be the generator of a strongly continuous semigroup on $H$, and let $F$ and $B$ be $\mathcal{P} \otimes \mathcal{B}(H)$-measurable mappings

$$\begin{aligned} F &: [0,T] \times \Omega \times H \to H, & (t,\omega,x) &\mapsto F_t(\omega,x), \\ B &: [0,T] \times \Omega \times H \to L_2(U_\Sigma; H), & (t,\omega,x) &\mapsto B_t(\omega,x). \end{aligned}$$

## 6.1 Solution concepts

**64 Definition.** Strong, weak, and mild solutions (in the analytical sense) of the stochastic evolution equation

$$dX_t = (AX_t + F_t(X_t))dt + B_t(X_t)dW_t, \qquad X_0 = \xi. \tag{4}$$

are defined as follows:

(i) A strong solution of (4) is a predictable process

$$X : [0,T] \times \Omega \to D(A)$$

which satisfies for each $t \in [0,T]$, almost surely,

$$X_t = \xi + \int_0^t \big(AX_s + F_s(X_s)\big)ds + \int_0^t B_s(X_s)dW_s.$$

(ii) A weak solution of (4) is a predictable process

$$X : [0,T] \times \Omega \to H$$

which satisfies for each $t \in [0,T]$ and $h \in D(A^*)$, almost surely,

$$\langle X_t, h\rangle_H = \langle \xi, h\rangle_H + \int_0^t \big(\langle X_s, A^*h\rangle_H + \langle F_s(X_s), h\rangle_H\big)ds + \int_0^t \langle B_s(X_s), h\rangle_H dW_s.$$

(iii) A mild solution of (4) is a predictable process

$$X : [0,T] \times \Omega \to H$$

which satisfies for every $t \in [0,T]$, almost surely,

$$X_t = e^{At}\xi + \int_0^t e^{A(t-s)} F_s(X_s) ds + \int_0^t e^{A(t-s)} B_s(X_s) dW_s.$$

**65 Remark.** It is part of the above definition that the Bochner integrals are well-defined a.s. and the stochastic integrals are well-defined.

**66 Remark.** The attributes weak and strong have the following additional unrelated meanings:

- A solution is strong in the stochastic sense if the Brownian motion and the stochastic basis are given, and weak in the stochastic sense if the Brownian motion and the stochastic basis may be chosen freely.
- The strong error of approximations $X^n$ of $X$ is $\mathbb{E}[\sup_{t\in[0,T]} \|X^n_t - X_t\|_H]$, and the weak error is $\mathbb{E}[f(X_T) - f(X^n_T)]$, where $f$ is a "nice" function.

**67 Lemma.**

*(i) Any strong solution is a weak and mild solution.*

*(ii) Any weak solution which satisfies*

$$\mathbb{P}\left[\int_0^T \left\|e^{A(T-t)}F_t(X_t)\right\|_H dt + \int_0^T \left\|e^{A(T-t)}B_t(X_t)\right\|^2_{L_2(U_\Sigma;H)} dt < \infty\right] = 1$$

*is a mild solution.*

*(iii) Any mild solution which satisfies for each $h \in D(A^*)$*

$$\mathbb{P}\left[\int_0^T |\langle F_t(X_t), h\rangle_H| dt + \int_0^T \|\langle B_t(X_t), h\rangle_H\|^2_{L_2(U_\Sigma;\mathbb{R})} dt < \infty\right] = 1$$

*is a weak solution.*

*Proof.* (i): Strong $\Rightarrow$ weak is trivial, and strong $\Rightarrow$ mild follows from (ii).

(ii): Let $X$ be a weak solution, let $D(A^*)$ be endowed with the graph norm, let $h \in D(A^*)$, let $f \in C^1([0,T])$, and let $Y = fh \in C^1([0,T]; D(A^*))$. Itō's formula applied to the product of $\langle X, h\rangle_H$ and $f$ yields

$$\begin{aligned}
\langle X_t, Y_t\rangle_H &= \langle X_t, h\rangle_H f_t \\
&= \langle \xi, h\rangle_H f_0 + \int_0^t \big(\langle X_s, A^* h\rangle_H + \langle F_s(X_s), h\rangle_H\big) f_s ds \\
&\quad + \int_0^t \langle X_s, h\rangle_H f'_s ds + \int_0^t \langle B_s(X_s), Y_s\rangle_H dW_s \\
&= \langle \xi, Y_0\rangle_H + \int_0^t \big(\langle X_s, A^* Y_s\rangle_H + \langle F_s(X_s), Y_s\rangle_H + \langle X_s, Y'_s\rangle_H\big) ds
\end{aligned}$$

$$+\int_0^t \langle B_s(X_s), Y_s\rangle_H dW_s.$$

By continuity, this relation extends to all functions $Y \in C^1([0,T]; D(A^*))$. In particular, the choice

$$Y_s = (e^{A(t-s)})^* h = e^{A^*(t-s)}h, \qquad Y_s' = -A^* Y_s$$

yields

$$\langle X_t, h\rangle_H = \langle e^{At}\xi, h\rangle_H + \int_0^t \langle e^{A(t-s)} F_s(X_s), h\rangle_H ds + \int_0^t \langle e^{A(t-s)} B_s(X_s), h\rangle_H dW_s.$$

Thus, by the integrability assumptions on $F(X)$ and $B(X)$ and by the denseness of $D(A^*)$, $X$ is a mild solution, and we have shown (ii).

(iii): Let $X$ be a mild solution, and let $h \in D(A^*)$. By Fubini's theorem for Bochner and stochastic integrals, after suitable localization,

$$\begin{aligned}
&\int_0^t \langle X_s, A^*h\rangle_H ds \\
&\quad= \int_0^t \left\langle e^{As}\xi + \int_0^s e^{A(s-u)} F_u(X_u)du + \int_0^s e^{A(s-u)} B_u(X_u)dW_u, A^*h \right\rangle_H ds \\
&\quad= \int_0^t \langle e^{As}\xi, A^*h\rangle_H ds + \int_0^t \int_u^t \langle e^{A(s-u)} F_u(X_u), A^*h\rangle_H ds du \\
&\qquad + \int_0^t \int_u^t \langle e^{A(s-u)} B_u(X_u), A^*h\rangle_H ds dW_u \\
&\quad= \int_0^t \langle \xi, e^{A^*s} A^*h\rangle_H ds + \int_0^t \int_u^t \langle F_u(X_u), e^{A^*(s-u)} A^*h\rangle_H ds du \\
&\qquad + \int_0^t \int_u^t \langle B_u(X_u), e^{A^*(s-u)} A^*h\rangle_H ds dW_u \\
&\quad= \langle \xi, (e^{A^*t} - I)h\rangle_H + \int_0^t \langle F_u(X_u), (e^{A^*(t-u)} - I)h\rangle_H du \\
&\qquad + \int_0^t \langle B_u(X_u), (e^{A^*(t-u)}h - I)\rangle_H dW_u \\
&\quad= \langle (e^{At} - I)\xi, h\rangle_H + \int_0^t \langle (e^{A(t-u)} - I) F_u(X_u), h\rangle_H du \\
&\qquad + \int_0^t \langle (e^{A(t-u)} - I) B_u(X_u), h\rangle_H dW_u \\
&\quad= \langle X_t, h\rangle - \langle \xi, h\rangle_H - \int_0^t \langle F_u(X_u), h\rangle_H du - \int_0^t \langle B_u(X_u), h\rangle_H dW_u.
\end{aligned}$$

Thus, $X$ is a weak solution, and we have shown (iii). □

## 6.2 Existence and uniqueness

**68 Theorem.** *If $F$ and $B$ are Lipschitz continuous in $x$, with a Lipschitz constant not depending on $(t, \omega)$, then there exists a mild solution of (4), which* *is unique up to modifications among the predictable processes satisfying*

$$\mathbb{P}\left[\int_0^T \|X_s\|_H^2 ds < \infty\right] = 1. \tag{5}$$

**69 Remark.**

- The integrals in (4) are well-defined if (5) holds, thanks to the Lipschitz continuity of $F$ and $B$.
- Uniqueness may fail if (5) is not imposed.

We need an auxiliary lemma.

**70 Lemma.** *Let $p \geq 2$, and let $E$ be the Banach space of $dt \otimes \mathbb{P}$-equivalence classes of predictable processes $X : [0,T] \times \Omega \to H$ satisfying*

$$\|X\|_E := \sup_{t\in[0,T]} \|X_t\|_{L^p(\Omega;H)} < \infty.$$

*If $F$ and $B$ are Lipschitz continuous in $x$, with a Lipschitz constant not depending on $(t,\omega)$, and $\xi \in L^p(\Omega; H)$, then the following statements hold:*

*(i) There is a constant $C_1$, which is given by (6), such that for any $X \in E$, the process*

$$KX_t = e^{At}\xi + \int_0^t e^{A(t-s)} F_s(X_s) ds + \int_0^t e^{A(t-s)} B_s(X_s) dW_s$$

*satisfies $KX \in E$ and $\|KX\|_E \leq C_1(1 + \|X\|_E)$.*

*(ii) There is a constant $C_2$, which is given by (7), such that for any $X, Y \in E$,*

$$\|KX - KY\|_E \leq C_2 \|X - Y\|_E .$$

*(iii) There is a constant $C_3$, which is given by (8), such that for any $X, Y \in E$,*

$$\|X - Y\|_E \leq C_3 \|X - KX - Y + KY\|_E .$$

*(iv) There exists a unique solution of (4) in $E$.*

*Proof.* By the Lipschitz property, there is a constant $M \in (0,\infty)$ such that for each $t \in [0,T]$, $x, y \in H$, and $\omega \in \Omega$,

$$\begin{aligned}
\left\|e^{At}\right\|_{L(H)} &\leq M,\\
\|F_t(\omega,x) - F_t(\omega,y)\|_H + \|B_t(\omega,x) - B_t(\omega,y)\|_{L_2(U_\Sigma;H)} &\leq M\|x-y\|_H,\\
\|F_t(\omega,x)\|_H + \|B_t(\omega,x)\|_{L_2(U_\Sigma;H)} &\leq M(1 + \|x\|_H).
\end{aligned}$$

(i): Let $X \in E$ and $t \in [0,T]$. The Bochner integral in Equation (4) satisfies each $t \in [0,T]$:

$$\left\|\int_0^t e^{A(t-s)} F_s(X_s) ds\right\|_{L^p(\Omega;H)}$$

$$
\begin{aligned}
&\leq \left\| e^{A(t-\cdot)} F(X) \right\|_{L^p(\Omega;L^1([0,t];H))} && \text{(Minkowski)}\\
&\leq M^2 \left\| 1 + \|X\|_H \right\|_{L^p(\Omega;L^1([0,t]))} && \text{(Lipschitz)}\\
&\leq M^2 t^{1-1/p} \left\| 1 + \|X\|_H \right\|_{L^p(\Omega;L^p([0,t]))} && \text{(Jensen)}\\
&\leq M^2 t^{1-1/p} \left\| 1 + \|X\|_H \right\|_{L^p([0,t];L^p(\Omega))} && \text{(Fubini).}
\end{aligned}
$$

and the stochastic integral in Equation (4), we get the following estimates for each $t \in [0,T]$ (using in particular the Burkholder-Davis-Gundy inequality of Lemma 62):

$$
\begin{aligned}
&\left\| \int_0^t e^{A(t-s)} B_s(X_s) dW_s \right\|_{L^p(\Omega;H)}\\
&\qquad \leq \left\| e^{A(t-\cdot)} B(X) \right\|_{L^p(\Omega;L^2([0,t];L_2(U_\Sigma;H)))} && \text{(BDG)}\\
&\qquad \leq M^2 \left\| 1 + \|X\|_H \right\|_{L^p(\Omega;L^2([0,t]))} && \text{(Lipschitz)}\\
&\qquad \leq M^2 t^{1/2-1/p} \left\| 1 + \|X\|_H \right\|_{L^p(\Omega;L^p([0,t]))} && \text{(Jensen)}\\
&\qquad = M^2 t^{1/2-1/p} \left\| 1 + \|X\|_H \right\|_{L^p([0,t];L^p(\Omega))} && \text{(Fubini),}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\|KX\|_E &\leq M \|\xi\|_{L^p(\Omega;H)} + M^2 (T^{1-1/p} + T^{1/2-1/p}) T^{1/p} (1 + \|X\|_E)\\
&\leq C_1 (1 + \|X\|_E),
\end{aligned}
$$

where

$$
C_1 = M \|\xi\|_{L^p(\Omega;H)} + M^2 (T + T^{1/2}). \tag{6}
$$

This shows (i).

(ii): Let $X, Y \in E$ and $t \in [0,T]$. The same steps as in (i) lead to the following estimates:

$$
\begin{aligned}
&\left\| \int_0^t e^{A(t-s)} (F_s(X_s) - F_s(Y_s)) ds \right\|_{L^p(\Omega;H)}\\
&\qquad \leq \left\| e^{A(t-\cdot)} (F(X) - F(Y)) \right\|_{L^p(\Omega;L^1([0,t];H))} && \text{(Minkowski)}\\
&\qquad \leq M^2 \|X - Y\|_{L^p(\Omega;L^1([0,t];H))} && \text{(Lipschitz)}\\
&\qquad \leq M^2 t^{1-1/p} \|X - Y\|_{L^p(\Omega;L^p([0,t];H))} && \text{(Jensen)}\\
&\qquad \leq M^2 t^{1-1/p} \|X - Y\|_{L^p([0,t];L^p(\Omega;H))} && \text{(Fubini),}
\end{aligned}
$$

and

$$
\begin{aligned}
&\left\| \int_0^t e^{A(t-s)} (B_s(X_s) - B_s(Y_s)) dW_s \right\|_{L^p(\Omega;H)}\\
&\qquad \leq \left\| e^{A(t-\cdot)} (B(X) - B(Y)) \right\|_{L^p(\Omega;L^2([0,t];L_2(U_\Sigma;H)))} && \text{(BDG)}\\
&\qquad \leq M^2 \|X - Y\|_{L^p(\Omega;L^2([0,t];H))} && \text{(Lipschitz)}\\
&\qquad \leq M^2 t^{1/2-1/p} \|X - Y\|_{L^p(\Omega;L^p([0,t];H))} && \text{(Jensen)}
\end{aligned}
$$

$$= M^2 t^{1/2-1/p} \|X-Y\|_{L^p([0,t];L^p(\Omega;H))} \qquad \text{(Fubini)}.$$

Therefore,

$$\|KX_t - KY_t\|_{L^p(\Omega;H)} \le M^2(T^{1-1/p} + T^{1/2-1/p}) \|X-Y\|_{L^p([0,t];L^p(\Omega;H))}.$$

Taking the supremum over $t \in [0,T]$ and using Hölder's inequality shows

$$\|KX - KY\|_E \le C_2 \|X-Y\|_E,$$

where

$$C_2 = M^2(T + T^{1/2}). \tag{7}$$

This shows (ii).

(iii): The last estimate above implies

$$\begin{aligned}\|X_t - Y_t\|_{L^p(\Omega;H)} &\le \|X_t - KX_t - Y_t + KY_t\|_{L^p(\Omega;H)} + \|KX_t - KY_t\|_{L^p(\Omega;H)} \\ &\le \|X - KX - Y + KY\|_E \\ &\quad + M^2(T^{1-1/p} + T^{1/2-1/p}) \|X-Y\|_{L^p([0,t];L^p(\Omega;H))}.\end{aligned}$$

Taking the $p$-th power and using $(a+b)^p \le 2^{p-1}(a^p + b^p)$ for any $a, b \ge 0$ yields

$$\begin{aligned}\|X_t - Y_t\|^p_{L^p(\Omega;H)} &\le 2^{p-1} \|X - KX - Y + KY\|^p_E \\ &\quad + 2^{p-1} M^{2p}(T^{1-1/p} + T^{1/2-1/p})^p \int_0^t \|X_s - Y_s\|^p_{L^p(\Omega;H)} ds.\end{aligned}$$

By Gronwalls inequality may be applied because both sides are bounded, and we get

$$\begin{aligned}\|X_t - Y_t\|^p_{L^p(\Omega;H)} &\le 2^{p-1} \|X - KX - Y + KY\|^p_E \\ &\quad \times \exp\left(2^{p-1} M^{2p}(T^{1-1/p} + T^{1/2-1/p})^p T\right).\end{aligned}$$

Taking the $p$-th root establishes (iii) with

$$C_3 = 2^{1-1/p} \exp\left(2^{1-1/p} M^2 (T + T^{1/2})\right). \tag{8}$$

(iv): For sufficiently small $T$, $K : E \to E$ is a contraction by (i) and (ii). In this case, Banach's fixed point theorem gives a unique solution of (4) in $E$. For arbitrary $T$, existence follows by concatenation of small time intervals, and uniqueness follows from (iii). □

*Proof of Theorem 68.* Let $p \ge 2$ be fixed.

Existence: For each $n \in \mathbb{N}$, let $\Gamma_n = \mathbb{1}_{\|\xi\|_H \le n}$, and let $X^n \in E$ satisfy for each $t \in [0,T]$

$$X^n_t = \Gamma_n e^{At}\xi + \int_0^t e^{A(t-s)} \Gamma_n F_s(X^m_s) ds + \int_0^t e^{A(t-s)} \Gamma_n B_s(X^m_s) dW_s.$$

Then one has for for all $m \ge n$ that $\Gamma_n X^n = \Gamma_n X^m \in E$ because all processes $\Gamma_n X^m$ satisfy one and the same equation

$$\Gamma_n X^m_t = \Gamma_n e^{At}\xi + \int_0^t e^{A(t-s)} \Gamma_n F_s(\Gamma_n X^m_s) ds + \int_0^t e^{A(t-s)} \Gamma_n B_s(\Gamma_n X^m_s) dW_s.$$

Thus, for each $t \in [0,T]$, the limit $X_t := \lim_{n\to\infty} \Gamma_n X_t^n$ exists a.s. The $dt \otimes \mathbb{P}$-equivalence class of $X$ contains a predictable process, which satisfies (5) and is a mild solution of (4). This establishes existence.

Uniqueness: Let $X$ and $Y$ be mild solutions of (4) satisfying (5). For each $n \in \mathbb{N}$, let

$$S_n = \inf\left\{t \in [0,T] : \|X_s\|^2 + \|Y_s\|^2 \geq n\right\} \wedge T,$$

let $X_t^n = \Gamma_n X_{t\wedge S_n}$, let $Y_t^n = \Gamma_n Y_{t\wedge S_n}$, and define for each $Z \in E$ and $t \in [0,T]$

$$\begin{aligned} K_n Z_t = e^{A(t\wedge S_n)}\Gamma_n\xi &+ \int_0^t e^{A(t-s)}\mathbb{1}_{(0,S_n]}\Gamma_n F_s(Z_s)ds \\ &+ \int_0^t e^{A(t-s)}\mathbb{1}_{(0,S_n]}\Gamma_n B_s(Z_s)dW_s. \end{aligned}$$

Then $X^n$ and $Y^n$ belong to $E$ and satisfy $X^n = K_n X^n$, $Y^n = K_n Y^n$. By Lemma 70.(iii), $X^n = Y^n \in E$. Therefore, $X = Y$ up to modifications. □

## 6.3 Existence of continuous modifications

**71 Theorem.** *The mild solution provided by Theorem 68 has a continuous modification.*

We need some auxiliary lemmas.

**72 Lemma.** *Let $p > 2$, let $\xi \in L^p(\Omega; H)$, let $X$ be the mild solution provided by Lemma 70, and let $\alpha < 1/2$. Then the process*

$$Z_t = \int_0^t (t-s)^{-\alpha} e^{A(t-s)} B_s(X_s)dW_s.$$

*belongs to $L^p(\Omega; L^p([0,T]; H))$.*

*Proof.* Let $M \in (0,\infty)$ be as in the proof of Lemma 70, and let $I_{[0,T]}$ denote the identity on $[0,T]$. Then

$$\begin{aligned}
\|Z\|_{L^p(\Omega;L^p([0,T];H))} &= \int_0^T \mathbb{E}\left[\|Z_t\|_H^p\right] dt && \text{(Fubini)}\\
&\leq c_p^p \int_0^T \mathbb{E}\left[\left(\int_0^t (t-s)^{-2\alpha}\|e^{A(t-s)}B_s(X_s)\|^2_{L_2(U_\Sigma;H)}ds\right)^{p/2}\right] dt && \text{(BDG)}\\
&\leq c_p^p M^{2p}\int_0^T \mathbb{E}\left[\left(\int_0^t (t-s)^{-2\alpha}(1+\|X_s\|_H)^2 ds\right)^{p/2}\right] dt && \text{(Lipschitz)}\\
&= c_p^p M^{2p}\mathbb{E}\left[\left\|I^{-2\alpha} \star (1+\|X\|_H)^2\right\|^{p/2}_{L^{p/2}([0,T])}\right] && \text{(Fubini)}\\
&\leq c_p^p M^{2p}\mathbb{E}\left[\left\|I^{-2\alpha}\right\|^{p/2}_{L^1([0,T])}\left\|(1+\|X\|_H)^2\right\|^{p/2}_{L^{p/2}([0,T])}\right] && \text{(Young)}\\
&= c_p^p M^{2p}\left\|I^{-2\alpha}\right\|^{p/2}_{L^1([0,T])}\int_0^T \mathbb{E}\left[(1+\|X_t\|_H)^p\right] dt && \text{(Fubini)}
\end{aligned}$$

The right-hand side is finite thanks to the condition $\alpha < 1/2$ and Lemma 70. □

**73 Lemma.** *For any $p > 1$ and $\alpha > 1/p$,*

$$(G_\alpha z)_t = \int_0^t (t-s)^{\alpha-1} e^{A(t-s)} z_s ds$$

*defines a bounded linear operator $G_\alpha : L^p([0,T];H) \to C([0,T];H)$.*

*Proof.* Let $q$ be such that $1/p + 1/q = 1$, and let $I_{[0,T]}$ be the identity on $[0,T]$. For any $t \in [0,T]$,

$$\begin{aligned} \|(G_\alpha z)_t\|_H &\le M \int_0^t (t-s)^{\alpha-1} \|z_s\|_H ds && \text{(Minkowski)} \\ &\le M \left\| I_{[0,T]}^{\alpha-1} \right\|_{L^q} \|z\|_{L^p([0,T];H)} && \text{(Hölder)} \end{aligned}$$

and similarly, for any $s, t \in [0,T]$,

$$\begin{aligned} \|(G_\alpha z)_s - (G_\alpha z)_t\|_H &= \left\| \int_0^T u^{\alpha-1} e^{Au} \big( \mathbb{1}_{[0,s]}(u) z_{s-u} - \mathbb{1}_{[0,t]}(u) z_{t-u} \big) du \right\|_H \\ &\le M \left\| I_{[0,T]}^{\alpha-1} \right\|_{L^q} \|\mathbb{1}_{[0,s]}(u) z_{s-u} - \mathbb{1}_{[0,t]}(u) z_{t-u}\|_{L^p([0,T];H)}. \end{aligned}$$

This shows that $G_\alpha z$ belongs to $C([0,T];H)$ if $z$ is continuous. By the density of continuous functions in $L^p([0,T];H)$, this holds for general $z$. $\square$

**74 Lemma.** *Let $p > 2$, assume that $\xi \in L^p(\Omega;H)$, let $X$ be the mild solution provided by Lemma 70, let $\alpha \in (1/p, 1/2)$, and let $Z$ be as in Lemma 72. Then the process*

$$\frac{\sin(\alpha\pi)}{\pi} G_\alpha Z$$

*is a continuous modification of the process*

$$\int_0^{\cdot} e^{A(\cdot - s)} B_s(X_s) dW_s.$$

*Proof.* We use Euler's reflection formula for the Beta function, which states that for each $0 \le s \le t$ and $\alpha \in (0,1)$,

$$\int_s^t (t-u)^{\alpha-1} (u-s)^{-\alpha} du = \int_0^1 (1-u)^{\alpha-1} u^{-\alpha} du = \mathrm{Beta}(\alpha, 1-\alpha) = \frac{\pi}{\sin(\alpha\pi)}.$$

For each $t \in [0,T]$, assuming that we may apply the stochastic Fubini theorem,

$$\begin{aligned} G_\alpha Z_t &= \int_0^t (t-u)^{\alpha-1} e^{A(t-u)} \int_0^u (u-s)^{-\alpha} e^{A(u-s)} B_s(X_s) dW_s du \\ &= \int_0^t \int_0^u (t-u)^{\alpha-1} (u-s)^{-\alpha} e^{A(t-s)} B_s(X_s) dW_s du \\ &= \int_0^t \int_s^t (t-u)^{\alpha-1} (u-s)^{-\alpha} du \; e^{A(t-s)} B_s(X_s) dW_s = \frac{\pi}{\sin(\alpha\pi)} Y_t. \end{aligned}$$

To verify the condition of Fubini's theorem, we estimate

$$\int_0^t \left\| s \mapsto (t-u)^{\alpha-1} e^{A(t-u)} (u-s)^{-\alpha} e^{A(u-s)} B_s(X_s) \right\|_{L^2([0,u];L_2(U_\Sigma;H))} du$$

$$\leq M \int_0^t (t-u)^{\alpha-1} \left\| s \mapsto (u-s)^{-\alpha} e^{A(u-s)} B_s(X_s) \right\|_{L^2([0,u];L_2(U_\Sigma;H))} du$$
$$= M \int_0^t (t-u)^{\alpha-1} \|Z_u\|_{L^2(\Omega;H))} du \qquad \text{(Itō's isometry)}$$
$$\leq M \left\| I_{[0,T]}^{\alpha-1} \right\|_{L^q} \|Z\|_{L^p([0,t];L^2(\Omega;H)} \qquad \text{(Lemma 73)}.$$

The right-hand side is finite by Jensen's inequality and Lemma 72. □

*Proof of Theorem 71.* For each $n \in \mathbb{N}$, let $\Gamma_n = \mathbb{1}_{\|\xi\|_H \leq n}$, and let $X^n = \Gamma_n X$. We saw in the proof of Theorem 68 that $X^n$ is an $E$-valued solution (where $E$ is defined with respect to some fixed $p > 2$) of the truncated SPDE

$$X_t^n = \Gamma_n e^{At}\xi + \int_0^t e^{A(t-s)} \Gamma_n F_s(X_s^m) ds + \int_0^t e^{A(t-s)} \Gamma_n B_s(X_s^m) dW_s.$$

The Bochner integral is continuous by Lemma 73, and the stochastic integral has a continuous modification by Lemma 74. As this holds for each $n \in \mathbb{N}$, $X$ has a continuous modification. □

## 6.4 Literature

Theorem 68 and the supporting Lemma 70 correspond to parts of [DZ14, Theorem 7.2 and 7.5]. Section 6.3 is taken from [DZ14].

# 7 Heath–Jarrow–Morton equation

## 7.1 Bond prices and interest rates

**75 Definition.** We denote time by $t \in \mathbb{R}_+$, time to maturity by $x \in \mathbb{R}_+$, and maturity by $T = t + x$, the price of zero-coupon bonds by $P(t,T) = P_t(x)$, the yields (or spot rate) by $Y(t,T) = Y_t(x)$, the instantaneous forward rate by $f(t,T) = f_t(x)$, the short rate by $r_t = f_t(0)$, and the bank account by $B_t = \exp(\int_0^t r_s ds)$. The following fundamental relation holds:

$$P(t,T) = \exp\big(-(T-t)Y(t,T)\big) = \exp\left(-\int_t^T f(t,s)ds\right).$$

**76 Remark.** The Euro-Area yield curves are public and can be viewed at `http://ecb.europa.edu/stats/money/yc/`.

## 7.2 Existence and uniqueness

**77 Setting.** We want to analyze the equation

$$df_t = \big(Af_t(x) + \alpha_t(f)\big)dt + \sigma_t(f)dW_t$$

subject to the following assumptions:

(i) $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{R}_+})$ is a stochastic basis, $U$ is a separable Hilbert space, and $W$ is an $I_U$-cylindrical Brownian motion on $U$ with respect to $(\mathcal{F}_t)$.

(ii) $H$ is a separable Hilbert space consisting of continuous functions $\mathbb{R}_+ \to \mathbb{R}$ such that point evaluations

$$\delta_x : H \to \mathbb{R}, f \mapsto f(x)$$

are continuous for all $x \in \mathbb{R}_+$.

(iii) For each $f \in H$ and $t \in \mathbb{R}_+$, the function $S_t f := f(t+\cdot)$ belongs to $H$, and $S : \mathbb{R}_+ \to L(H)$ is a strongly continuous semigroup on $H$. The generator of $S$ is denoted by $A$.

(iv) There is a linear subspace $H_0$ of $H$ such that the following mapping is sounded bilinear:

$$m : H_0 \times H_0 \to H, \qquad m(f,g)(x) = f(x) \int_0^x g(y)dy.$$

(v) The functions

$$\lambda : \mathbb{R}_+ \times \Omega \times H \to U, \qquad \sigma : \mathbb{R}_+ \times \Omega \times H \to L_2(U;H)$$

are $\mathcal{P} \otimes \mathcal{B}(H)$-measurable, and

$$\alpha : \mathbb{R}_+ \times \Omega \times H \to H$$

is given by

$$\alpha_t(f) = \operatorname{Tr}\Big(m\big(\sigma_t(f), \sigma_t(f)\big)\Big) + \sigma_t(f)\lambda_t(f).$$

**78 Remark.** The trace in the definition of $\alpha$ is well-defined because

$$\begin{aligned}\operatorname{Tr}\Big(m\big(\sigma_t(f), \sigma_t(f)\big)\Big) &= \sum_{n=1}^{\infty} \Big\|(m\big(\sigma_t(f)e_n, \sigma_t(f)e_n\big)\Big\|_H \\ &\leq \|m\|_{L(H_0,H_0;H}\|\sigma_t(f)\|^2_{L_2(U;H_0}.\end{aligned}$$

**79 Theorem.** *Assume that Setting 77 holds, let $\sigma$ be bounded, let $\sigma$ and $\lambda$ be Lipschitz in $f$, and let $f_0 \in L^0(\mathcal{F}_0; H)$. Then there is an up to modifications unique predictable process $f : \mathbb{R}_+ \times \Omega \to H$ which satisfies for each $t \in \mathbb{R}_+$ that*

$$f_t = S_t f_0 + \int_0^t S_{t-s}\alpha_s(f_s)ds + \int_0^t S_{t-s}\sigma_s dW_s$$

*and*

$$\mathbb{P}\left[\forall t \in \mathbb{R}_+ : \int_0^t \|f_s\|_H^2 ds < \infty\right] = 1.$$

*Proof.* For each $f, g \in H$ set

$$\alpha_t(f,g) = \operatorname{Tr}\Big(m\big(\sigma_t(f), \sigma_t(g)\big)\Big) + \sigma_t(f)\lambda_t(g)$$

and observe that

$$\|\alpha_t(f) - \alpha_t(g)\|_H = \frac{1}{2}\|\alpha_t(f-g, f+g) + \alpha_t(f+g, f-g)\|_H$$

$$
\begin{aligned}
&\le \|m\|_{L(H_0,H_0;H}\|\sigma_t(f-g)\|_{L_2(U;H_0)}\|\sigma_t(f+g)\|_{L_2(U;H_0)} \\
&\qquad + \frac{1}{2}\|\sigma_t(f-g)\|_{L_2(U;H_0)}\|\lambda_t(f+g)\|_U \\
&\qquad + \frac{1}{2}\|\sigma_t(f+g)\|_{L_2(U;H_0)}\|\lambda_t(f-g)\|_U \\
&\le C\|f-g\|_H
\end{aligned}
$$

for some constant $C \in \mathbb{R}_+$ which does not depend on $f, g$. Thus, $\alpha$ and $\sigma$ are Lipschitz continuous, and the result follows from Theorem 68. □

## 7.3 Absence of arbitrage

**80 Definition.** (i) A probability measure $\mathbb{Q}$ on $\bigcup_{t\in\mathbb{R}_+}\mathcal{F}_t$ is called a local martingale measure if for every $T \in \mathbb{R}_+$ the process $(B_t^{-1}P(t,T))_{t\in[0,T]}$ is a $\mathbb{P}$-local martingale.

(ii) An interest rate model is called free of arbitrage if there exists a local martingale measure $\mathbb{Q}$ on $\bigcup_{t\in\mathbb{R}_+}\mathcal{F}_t$ which is equivalent to $\mathbb{P}$ on $\mathcal{F}_t$ for every $t \in \mathbb{R}_+$.

**81 Theorem.** *Let $\mathbb{P}$ and $f$ be as in Theorem 79. Then $\mathbb{P}$ is a local martingale measure if and only if $\sigma_t(f)\lambda_t(f)$ vanishes $dt \otimes \mathbb{P}$-almost surely.*

*Proof.* As a first step we show that the theorem holds under the additional assumption that $f_0 \in L^2(\mathcal{F}_0; H)$. For each $T \in \mathbb{R}_+$ and $f \in H$ let

$$I_T f = \int_0^T f(x)dx.$$

Then $I_T : H \to \mathbb{R}$ is continuous because

$$|I_T f| \le |T| \sup_{x\in[0,T]} |f(x)| \le |T| \sup_{x\in[0,T]} \|\delta_x\|_{H^*}\|f\|_H$$

and $\sup_{x\in[0,T]} \|\delta_x\|_{H^*} < \infty$ by the Banach-Steinhaus theorem. Therefore,

$$
\begin{aligned}
-\log P(t,T) &= \int_0^{T-t} f_t(x)dx = I_{T-t}f_t \\
&= I_{T-t}S_t f_0 + \int_0^t I_{T-t}S_{t-s}\alpha_s(f_s)ds + \int_0^t I_{T-t}S_{t-s}\sigma_s(f_s)dW_s \\
&= (I_T - I_t)f_0 + \int_0^t (I_{T-s} - I_{t-s})\alpha_s(f_s)ds + \int_0^t (I_{T-s} - I_{t-s})\sigma_s(f_s)dW_s \\
&= I_T f_0 + \int_0^t I_{T-s}\alpha_s(f_s)ds + \int_0^t I_{T-s}\sigma_s(f_s)dW_s \\
&\qquad - I_t f_0 - \int_0^t \int_s^t \delta_{u-s}\alpha_s(f_s)duds - \int_0^t \int_s^t \delta_{u-s}\sigma_s(f_s)dudW_s.
\end{aligned}
$$

To verify that we may apply Fubini's theorem to the last two integrals, note that the solution $f$ lies in the space $E$ of Lemma 70 with $p = 2$. Thus, letting

$|\alpha|_{\mathrm{Lip}}$ and $|\sigma|_{\mathrm{Lip}}$ denote the Lipschitz constants of $\alpha$ and $\sigma$ with respect to $f$, we have

$$\mathbb{E}\left[\int_0^t \int_0^u |\delta_{u-s}\alpha_s(f_s)|dsdu\right] \le t^2 \sup_{x\in[0,t]} \|\delta_x\|_{H^*}|\alpha|_{\mathrm{Lip}}\|f\|_E < \infty$$

and

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \left(\int_0^u \|\delta_{u-s}\sigma_s(f_s)\|^2_{L_2(U;\mathbb{R})}ds\right)^{1/2} du\right] \\
&\qquad \le \int_0^t \left(\int_0^u \mathbb{E}\left[\|\delta_{u-s}\sigma_s(f_s)\|^2_{L_2(U;\mathbb{R})}\right] ds\right)^{1/2} du \\
&\qquad \le t^{3/2} \sup_{x\in[0,t]} \|\delta_x\|_{H^*}|\sigma|_{\mathrm{Lip}}\|f\|_E < \infty.
\end{aligned}$$

Therefore, we can apply Fubini's theorem and get

$$\begin{aligned}
-\log P(t,T) &= I_T f_0 + \int_0^t I_{T-s}\alpha_s(f_s)ds + \int_0^t I_{T-s}\sigma_s(f_s)dW_s \\
&\quad - I_t f_0 - \int_0^t\int_0^u \delta_{u-s}\alpha_s(f_s)dsdu - \int_0^t\int_0^u \delta_{u-s}\sigma_s(f_s)dW_s du \\
&= I_T f_0 + \int_0^t I_{T-s}\alpha_s(f_s)ds + \int_0^t I_{T-s}\sigma_s(f_s)dW_s \\
&\quad - \int_0^t \delta_0\left(S_u f_0 + \int_0^u S_{u-s}\alpha_s(f_s)ds + \int_0^u S_{u-s}\sigma_s(f_s)dW_s\right) du \\
&= I_T f_0 + \int_0^t I_{T-s}\alpha_s(f_s)ds + \int_0^t I_{T-s}\sigma_s(f_s)dW_s - \int_0^t r_u du.
\end{aligned}$$

By Itō's formula, the processes $P(t,T)$ and $\tilde{P}(t,T) := B_t^{-1}P(t,T)$ satisfy

$$\begin{aligned}
P(t,T) &= \exp\left(-I_T f_0 - \int_0^t I_{T-s}\alpha_s(f_s)ds - \int_0^t I_{T-s}\sigma_s(f_s)dW_s + \int_0^t r_u du\right) \\
&= P(0,T) - \int_0^t P(s,T)I_{T-s}\sigma_s(f_s)dW_s \\
&\quad + \int_0^t P(s,T)\left(r_s - I_{T-s}\alpha_s(f_s) + \frac{1}{2}\operatorname{Tr}\left(I_{T-s}\sigma_s(f_s), I_{T-s}\sigma_s(f_s)\right)\right) ds,
\end{aligned}$$

and

$$\begin{aligned}
\tilde{P}(t,T) &= \tilde{P}(0,T) - \int_0^t \tilde{P}(s,T)I_{T-s}\sigma_s(f_s)dW_s \\
&\quad + \int_0^t \tilde{P}(s,T)\left(-I_{T-s}\alpha_s(f_s) + \frac{1}{2}\operatorname{Tr}\left(I_{T-s}\sigma_s(f_s), I_{T-s}\sigma_s(f_s)\right)\right) ds.
\end{aligned}$$

For each $g \in H$ one has

$$m(g,g)(x) = g(x)\int_0^x g(y)dy = \frac{d}{dx}\frac{1}{2}\left(\int_0^x g(y)dy\right)^2,$$

which implies that

$$I_{T-s}\,\mathrm{Tr}\,\big(m(\sigma_s(f_s),\sigma_s(f_s))\big) = \frac{1}{2}\,\mathrm{Tr}\,\big(I_{T-s}\sigma_s(f_s), I_{T-s}\sigma_s(f_s)\big).$$

Therefore, by the definition of $\alpha$,

$$\begin{aligned}\tilde{P}(t,T) &= \tilde{P}(0,T) - \int_0^t \tilde{P}(s,T) I_{T-s}\sigma_s(f_s)\lambda_s ds - \int_0^t \tilde{P}(s,T) I_{T-s}\sigma_s(f_s) dW_s \\ &= \tilde{P}(0,T)\,\mathcal{E}\bigg(\int_0^{\cdot} I_{T-s}\sigma_s(f_s)(dW_s - \lambda_s ds)\bigg)_t,\end{aligned} \tag{9}$$

where $\mathcal{E}$ denotes the stochastic exponential. This process is a local martingale for each $T$ if and only if for each $T \in \mathbb{R}_+$ the following is a $dt \otimes \mathbb{P}$-null set:

$$\{(t,\omega) \in [0,T] \times \Omega : I_{T-t}\sigma_t(f_t)\lambda_t \neq 0\}.$$

Differentiation with respect to $T$ shows that this is equivalent to $\sigma_t(f_t)\lambda_t$ vanishing $dt \otimes \mathbb{P}$-almost surely. This establishes the theorem under the assumption that $f_0 \in L^2(\mathcal{F}_0; H)$. The general case follows using the localization technique in the existence proof of Lemma 70. □

**82 Corollary.** *Let $\mathbb{P}$ and $f$ be as in Theorem 79, and assume that for each $t \in \mathbb{R}_+$*

$$\mathbb{E}\left[\exp\left(-\frac{1}{2}\int_0^t \|\lambda_s(f_s)\|_U^2 ds + \int_0^t \langle \lambda_s(f_s), \cdot\rangle_U dW_s\right)\right] = 1.$$

*Then the bond market is free of arbitrage.*

*Proof.* Let $\mathbb{Q}$ be the probability measure on $\bigcup_{t\in\mathbb{R}_+}\mathcal{F}_t$ which satisfies for each $t \in \mathbb{R}_+$ that

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = \exp\left(-\frac{1}{2}\int_0^t \|\lambda_s(f_s)\|_U^2 ds + \int_0^t \langle \lambda_s(f_s), \cdot\rangle_U dW_s\right).$$

By Girsanov's theorem, the process $W_t - \int_0^t \lambda_s ds$ is a Brownian motion under $\mathbb{Q}$. It follows from (9) that the discounted bond prices are local martingales under $\mathbb{Q}$. Thus, $\mathbb{Q}$ is an equivalent local martingale measure, and the bond market is free of arbitrage. □

## 7.4 Examples

The following lemma shows that the presence of the shift semigroup in the HJM equation is due to the parametrization of forward rates as functions of time to maturity $x$ instead of maturity $T$. Moreover, it shows how mild solutions of the HJM equation can be constructed forward rate processes which are defined pointwise for every maturity $T$.

**83 Lemma.** *Let Setting 77.(i) be in place. For every $T \in \mathbb{R}_+$ let*

$$\alpha(\cdot,T) : [0,T] \times \Omega \to \mathbb{R}, \qquad \sigma(\cdot,T) : [0,T] \times \Omega \to L_2(U;\mathbb{R})$$

*be predictable processes such that*

$$\int_0^T |\alpha(s,T)|ds < \infty, \qquad \int_0^T \|\sigma(s,T)\|^2_{L_2(U;\mathbb{R})} ds < \infty.$$

*Then the following statements hold:*

(i) *The equation*

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dW_s,$$

*holds for every* $t \in [0,T]$ *and every* $T \in \mathbb{R}_+$ *if and only if the equation*

$$f_t(x) = S_t f_0(x) + \int_0^t S_{t-s}\alpha_s(x)ds + \int_0^t S_{t-s}\sigma_s(x)dW_s$$

*holds for every* $t, x \in \mathbb{R}_+$*, where* $S$ *is the shift semigroup on functions* $\mathbb{R}_+ \to \mathbb{R}$ *and where*

$$f_t(x) = f(t, t+x), \qquad \alpha_t(x) = \alpha(t, t+x), \qquad \sigma_t(x) = \sigma(t, t+x).$$

(ii) *If additionally Setting 77.(ii)–(iii) holds, if there are predictable processes*

$$f : \mathbb{R}_+ \times \Omega \to H, \qquad \alpha : \mathbb{R}_+ \times \Omega \to H, \qquad \sigma : \mathbb{R}_+ \times \Omega \to L_2(U; H)$$

*such that for every* $x \in \mathbb{R}_+$*,*

$$f_t(x) = \delta_x f_t, \qquad \alpha_t(x) = \delta_x \alpha_t, \qquad \sigma_t(x) = \delta_x \sigma_t,$$

*and if for every* $t \in \mathbb{R}_+$*,*

$$\int_0^t \|\alpha_s\|_H ds < \infty, \qquad \int_0^t \|\sigma_s\|^2_{L_2(U;H)} ds < \infty,$$

*then* $f$ *is a mild solution of the HJM equation on* $H$*.*

*Proof.* See exercises or [Fil01, Sections 4.1 and 4.2]. □

The following is a common choice of a Hilbert space containing the forward rate curves.

**84 Lemma.** *Let* $\alpha > 3$*, and let*

$$H = \left\{ f \in H^1_{\text{loc}} : \|f\|_H < \infty \right\},$$

*where*

$$\|f\|^2_H = |f(0)|^2 + \int_0^\infty |f'(x)|^2 w(x)dx, \qquad w(x) = (1+x)^\alpha.$$

*Then* $H$ *satisfies Setting 77.(ii)–(iii).*

*Proof.* See exercises or [Fil01, Theorem 5.1.1 and Example 5.1.2]. □

**85 Lemma.** *Let $H$ be as in Lemma 84, let $\theta \in H$, let $\beta < 0$, let $a > 0$, let $U = \mathbb{R}$, let $W$ be $I_U$-Brownian motion, let $(r_t)_{t \in \mathbb{R}_+}$ be the unique solution of*

$$dr_t = (\theta(t) + \beta r_t)dt + \sqrt{a} dW_t, \tag{10}$$

*and define for each $t, x \in \mathbb{R}_+$*

$$\begin{aligned} P_t(x) &= \mathbb{E}\left[\exp\left(-\int_t^{t+x} r(s)ds\right)\middle|\mathcal{F}(t)\right], \\ f_t(x) &= -\frac{d}{dx}\log P_t(x). \end{aligned}$$

*Then $f$ is the unique mild solution of the HJM equation with $\lambda = 0$ and*

$$\sigma_t(f)(u)(x) = \sqrt{a} e^{\beta x} u, \qquad t \in \mathbb{R}_+, f \in H, u \in U, x \in \mathbb{R}_+.$$

*Proof.* See exercises or [Fil09, Section 5.4.1] and Lemma 83. □

**86 Lemma.** *Let $H$ be as in Lemma 84, let $\theta \in H$ satisfy $\theta(x) \geq 0$ for each $x \in \mathbb{R}_+$, let $\beta < 0$, let $\alpha > 0$, let $U = \mathbb{R}$, let $W$ be $I_U$-Brownian motion, let $(r_t)_{t \in \mathbb{R}_+}$ be the unique solution of*

$$dr_t = (\theta(t) + \beta r_t)dt + \sqrt{\alpha r_t} dW_t, \tag{11}$$

*and define $P_t(x)$ and $f_t(x)$ as in Lemma 85. Then $f$ is a mild solution of the HJM equation with $\lambda = 0$ and*

$$\sigma_t(f)(u)(x) = -\sqrt{\alpha f(0)}\Psi'(x)u, \qquad t \in \mathbb{R}_+, f \in H, u \in U, x \in \mathbb{R}_+,$$

*where for each $x \in \mathbb{R}_+$,*

$$\Psi(x) = \frac{-2\left(e^{\gamma x} - 1\right)}{\gamma\left(e^{\gamma x} + 1\right) - \beta\left(e^{\gamma x} - 1\right)}, \qquad \gamma = \sqrt{\beta^2 + 2\alpha}.$$

*Proof.* See exercises or [Fil09, Section 5.4.2] and Lemma 83. □

### 7.5 Literature

This section is similar to [CT07, Chapter 6] and [Fil01, Chapters 4 and 5]. Further information can be found in the textbook [Fil09] and in the extensive monograph [BM07].

## 8 Stochastic evolution equations with unbounded coefficients

### 8.1 Interpolation spaces

**87 Definition.** Let $H$ be a Hilbert space, let $A : D(A) \subseteq H \to H$ be a symmetric diagonal linear operator with $\inf \sigma_P(A) > 0$.

(i) For each $r \geq 0$, $A^r : D(A^r) \subseteq H \to H$ is the linear operator which is defined on

$$D(A^r) = \left\{ v \in H : \sum_{\lambda \in \sigma_P(A)} \left\| \lambda^r P_{\ker(\lambda - A)}(v) \right\|_H^2 < \infty \right\},$$

and which satisfies for each $v \in D(A^r)$ that

$$A^r v = \sum_{\lambda \in \sigma_P(A)} P_{\ker(\lambda - A)}(v).$$

(ii) A family of interpolation spaces associated to $A$ is a family $(H_r)_{r \in \mathbb{R}}$ of Hilbert spaces such that for all $q \geq 0$, all $r \geq s$, and all $v \in H_r$, $H_q = D(A^q)$, $H_r \subseteq H_s \subseteq \overline{H_r}^{H_s}$, and $\|v\|_{H_r} = \|A^{r-s} v\|_{H_s}$.

**88 Remark.**

- Interpolation spaces as in Definition 87 exist and are unique [Jen15, Theorem 3.5.24].
- For each $r \in \mathbb{R}$ the scalar product $\langle \cdot, \cdot \rangle_H$ admits an extension to a non-degenerate continuous bilinear form $H_r \times H_{-r} \to \mathbb{R}$.
- Interpolation spaces for more general operators are defined in [EN99, Section 2.5].

## 8.2 Smoothing effect of the semigroup

**89 Lemma.** *Let $H$ be a Hilbert space, let $A : D(A) \subseteq H \to H$ be a symmetric diagonal linear operator with $\sup \sigma_P(A) < 0$. Then*

*(i) $A$ is the generator of a strongly continuous semigroup, and*

*(ii) For each $r \geq 0$ it holds that*

$$\sup_{t \in [0,\infty)} \left\| (-tA)^r e^{At} \right\|_{L(H)} \leq \left( \frac{r}{e} \right)^r < \infty.$$

*Proof.* (i): As $A$ is diagonal, there is an orthonormal basis $\mathbb{B}$ of $H$ and a function $\lambda : \mathbb{B} \to \mathbb{R}$ such that for each $v \in D(A)$,

$$Av = \sum_{b \in \mathbb{B}} e^{\lambda_b t} \langle b, v \rangle_H b.$$

The assumption that $\sup \sigma_P(A) < 0$ ensures that the following function is well-defined,

$$S : [0, \infty) \to L(H), \qquad S_t v = \sum_{b \in \mathbb{B}} e^{\lambda_b t} \langle b, v \rangle_H b.$$

$S$ is a semigroup because for each $s, t \geq 0$ and $v \in H$,

$$S_s S_t v = S_s \left( \sum_{b \in \mathbb{B}} e^{\lambda_b t} \langle b, v \rangle_H b \right) = \sum_{b \in \mathbb{B}} e^{\lambda_b t} \langle b, v \rangle_H S_s b$$

$$= \sum_{b\in\mathbb{B}} e^{\lambda_b t}\langle b, v\rangle_H e^{\lambda_b s} b = S_{s+t} v,$$

and $S$ is strongly continuous because one has for each $v \in H$ by the dominated convergence theorem that

$$\begin{aligned}\lim_{t\searrow 0} \|S_t v - v\|_H^2 &= \lim_{t\searrow 0} \left\| \sum_{b\in\mathbb{B}} \left(e^{\lambda_b t} - 1\right) \langle b, v\rangle_H b \right\|_H^2 \\ &= \lim_{t\searrow 0} \sum_{b\in\mathbb{B}} \left|e^{\lambda_b t} - 1\right|^2 |\langle b, v\rangle_H|^2 = 0.\end{aligned}$$

The generator of $S$ is an extension of $A$ because for each $v \in D(A)$, by the dominated convergence theorem,

$$\begin{aligned}\lim_{t\searrow 0} \left\| \frac{S_t v - v}{t} - Av \right\|_H^2 &= \lim_{t\searrow 0} \left\| \sum_{b\in\mathbb{B}} \left( \frac{e^{\lambda_b t} - 1}{t} - \lambda_b \right) \langle b, v\rangle_H b \right\|_H^2 \\ &= \lim_{t\searrow 0} \sum_{b\in\mathbb{B}} \left| \frac{e^{\lambda_b t} - 1}{t} - \lambda_b \right|^2 |\langle b, v\rangle_H|^2 = 0.\end{aligned}$$

Conversely, $A$ extends the generator of $S$: if $v$ belongs to the domain of the generator of $S$, there exists $w \in H$ such that

$$0 = \lim_{t\searrow 0} \left\| \frac{S_t v - v}{t} - w \right\|^2 = \lim_{t\searrow 0} \sum_{b\in\mathbb{B}} \left| \frac{e^{\lambda_b t} - 1}{t} \langle b, v\rangle_H - \langle b, w\rangle_H \right|^2,$$

which implies that $\langle b, w\rangle_H = \lambda_b \langle b, v\rangle_H$ and

$$\sum_{b\in\mathbb{B}} |\lambda_b|^2 |\langle b, v\rangle_H|^2 = \sum_{b\in\mathbb{B}} |\langle b, w\rangle_H|^2 < \infty,$$

and therefore $v \in D(A)$ and $w = Av$. Therefore, the generator of $S$ equals $A$.

(ii): For each $r, t \geq 0$ one has

$$\left\|(-tA)^r e^{At}\right\|_{L(H)} = \sup_{b\in\mathbb{B}} \left|(-t\lambda_b)^r e^{\lambda_b t}\right| \leq \sup_{x\in(0,\infty)} \left|x^r e^{-x}\right| \leq \left(\frac{r}{e}\right)^r < \infty. \quad \square$$

## 8.3 Existence and uniqueness

**90 Setting.** We want to analyze the equation

$$dX_t = \big(AX_t + F(X_t)\big)dt + B(X_t)dW_t$$

subject to the following assumptions:

(i) $T \in [0, \infty)$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ is a stochastic basis, $U$ is a separable Hilbert space, and $(W_t)_{t\in[0,T]}$ is an $I_U$-cylindrical Brownian motion on $U$ with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

(ii) $H$ is a separable Hilbert space, $A : D(A) \subseteq H \to H$ is a symmetric diagonal linear operator with $\sup \sigma_P(A) < 0$, and $(H_r)_{r\in\mathbb{R}}$ is a family of interpolation spaces associated to $-A$.

(iii) $\gamma \in \mathbb{R}$, $p \in [2, \infty)$, and $\xi \in L^p(H_\gamma)$.

(iv) $\eta \in [0, 1)$, and $F$ and $B$ are $\mathcal{P} \otimes \mathcal{B}(H_\gamma)$-measurable mappings

$$\begin{aligned}
&F : [0,T] \times \Omega \times H_\gamma \to H_{\gamma-\eta}, && (t, \omega, x) \mapsto F_t(\omega, x),\\
&B : [0,T] \times \Omega \times H_\gamma \to L_2(U; H_{\gamma-\eta/2}), && (t, \omega, x) \mapsto B_t(\omega, x).
\end{aligned}$$

**91 Theorem.** *Let Setting 90 hold, and let $F$ and $B$ be Lipschitz continuous in $x$ with a Lipschitz constant not depending on $(t, \omega)$. Then there exists a unique up to modifications predictable processes $X : [0,T] \times \Omega \to H_\gamma$ which satisfies for each $t \in [0,T]$ that*

$$X_t = e^{At}\xi + \int_0^t e^{A(t-s)} F_s(X_s) ds + \int_0^t e^{A(t-s)} B_s(X_s) dW_s$$

*and*

$$\mathbb{P}\left[\int_0^T \|X_s\|_{H_\gamma}^2 ds < \infty\right] = 1.$$

*Proof.* The proof is similar to Lemma 70. We set up a fixed point problem on the Banach space $E$ of $dt \otimes \mathbb{P}$-equivalence classes of predictable processes $X : [0,T] \times \Omega \to H$ satisfying

$$\|X\|_E := \sup_{t \in [0,T]} \|X_t\|_{L^p(\Omega; H_\gamma)} < \infty.$$

There are three key estimates. First, by the smoothing effect of the semigroup (Lemma 89.(ii)), one has for each $t \geq 0$ and $r \in [0, e]$ that

$$\begin{aligned}
\left\|e^{At}\right\|_{L(H_{\gamma-r}; H_\gamma)} &= \left\|(-A)^\gamma e^{At} (-A)^{r-\gamma}\right\|_{L(H;H)}\\
&= t^{-\eta} \left\|(-tA)^r e^{At}\right\|_{L(H;H)} \leq t^{-r} \left(\frac{r}{e}\right)^r \leq t^{-r}.
\end{aligned}$$

Second, for any predictable process $F : [0,T] \times \Omega \to H_{\gamma-\eta}$ one has

$$\begin{aligned}
&\left\|\int_0^t e^{A(t-s)} F_s ds\right\|_{L^p(\Omega; H_\gamma)}\\
&\quad\leq \int_0^t \left\|e^{A(t-s)} F_s\right\|_{L^p(\Omega; H_\gamma)} ds && \text{(Minkowski)}\\
&\quad\leq \int_0^t (t-s)^{-\eta} \|F_s\|_{L^p(\Omega; H_{\gamma-\eta})} ds && \text{(smoothing)}\\
&\quad\leq \sqrt{\int_0^t (t-s)^{-\eta} ds \int_0^t (t-s)^{-\eta} \|F_s\|_{L^p(\Omega; H_{\gamma-\eta})}^2 ds} && \text{(C–S)}\\
&\quad= \sqrt{\frac{t^{1-\eta}}{1-\eta} \int_0^t (t-s)^{-\eta} \|F_s\|_{L^p(\Omega; H_{\gamma-\eta})}^2 ds}\\
&\quad\leq \frac{t^{1-\eta}}{1-\eta} \sqrt{\int_0^t \|F_s\|_{L^p(\Omega; H_{\gamma-\eta})}^2 ds} && \text{(Young).}
\end{aligned}$$

Third, for any predictable process $B : [0, T] \times \Omega \to L_2(U; H_{\gamma-\eta})$ one has

$$
\begin{aligned}
&\left\| \int_0^t e^{A(t-s)} B_s dW_s \right\|_{L^p(\Omega; H_\gamma)} \\
&\qquad \leq \sqrt{\frac{p(p-1)}{2} \int_0^t \left\| e^{A(t-s)} B_s \right\|^2_{L^p(\Omega; L_2(U; H_\gamma))} ds} && \text{(BDG)} \\
&\qquad \leq \sqrt{\frac{p(p-1)}{2} \int_0^t (t-s)^{-\eta} \left\| B_s \right\|^2_{L^p(\Omega; L_2(U; H_{\gamma-\eta/2}))} ds} && \text{(smoothing)} \\
&\qquad \leq \sqrt{\frac{p(p-1)}{2} \frac{t^{1-\eta}}{1-\eta} \int_0^t \left\| B_s \right\|^2_{L^p(\Omega; L_2(U; H_{\gamma-\eta/2}))} ds} && \text{(Young).}
\end{aligned}
$$

These estimates imply the local contraction property of the fixed point mapping and therefore existence and uniqueness similarly to the proof of Lemma 70. □

**92 Remark.**

- The condition $\xi \in L^p((\Omega, \mathcal{F}_0, \mathbb{P}); H_\gamma)$ can be relaxed to $\xi \in L^0((\Omega, \mathcal{F}_0, \mathbb{P}); H_\gamma)$ using localization as in the proof of Theorem 68.
- A priori estimates can be obtained using a generalized Gronwall inequality [Jen15, Corollary 1.4.6] applied to the expressions one step before Young's inequality (see [Jen15, Proposition 7.1.4].

## 8.4 Literature

This section follows the setup and notation of [Jen15, Section 7]. At the expense of heavier notation most results can be extended to analytic semigroups with generators that are not necessarily diagonal linear operators [DZ14, Section 6.5].

# 9 Stochastic heat equation

## 9.1 Existence and uniqueness

**93 Setting.** We want to analyze the equation

$$dX_t(x) = \big(\theta\, \Delta X_t(x) + f(x, X_t(x))\big)dt + b(x, X_t(x))dW_t(x), \qquad x \in (0,1),$$

in the following setting:

(i) $T \in [0, \infty)$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ is a stochastic basis, $U = L^2((0,1))$, and $(W_t)_{t\in[0,T]}$ is an $I_U$-cylindrical Brownian motion with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

(ii) $H = U$, $\theta > 0$, and $\Delta : H_0^2((0,1)) \subseteq H \to H$ is the Laplace operator.

(iii) $f, b : (0,1) \times \mathbb{R} \to \mathbb{R}$ are measurable functions, are Lipschitz in the second variable, and satisfy $f(\cdot, 0), b(\cdot, 0) \in L^2((0,1))$.

**94 Theorem.** *Under the assumptions of Setting 93, the following statements hold:*

*(i)* $A := \theta\Delta$ *is a diagonal linear operator, which satisfies* $\sup \sigma_P(A) < 0$ *and generates a strongly continuous semigroup.*

*(ii)* *For any* $\beta < -1/4$*, the following mappings are well-defined and Lipschitz continuous,*

$$\begin{aligned} &F : H \to H, && F(v)(x) = f(x, v(x)), \\ &B : H \to L_2(U; H_\beta), && B(v)(u)(x) = b(x, v(x))u(x), \end{aligned}$$

*where* $H_\beta$ *denotes an interpolation space associated to* $-A$*.*

*(iii)* *For each* $\xi \in L^2((\Omega, \mathcal{F}_0, \mathbb{P}); H)$ *there exist a unique up to modifications predictable process* $X : [0,T] \times \Omega \to H$ *which satisfies for each* $t \in [0,T]$ *that*

$$X_t = e^{At}\xi + \int_0^t e^{A(t-s)} F(X_s)ds + \int_0^t e^{A(t-s)} B(X_s)dW_s$$

*and*

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

*Proof.* We follow [Jen15, Section 7.2.1].

(i): For each $n \in \mathbb{N}$ and $x \in (0,1)$ let

$$e_n(x) = \sqrt{2}\sin(n\pi x).$$

Then $(e_n)_{n=1,2,\dots}$ is an orthonormal basis of $H$, each $e_n$ is contained in $D(A)$, and $Ae_n = -\theta\pi^2 n^2 e_n$. Thus, $A$ is a diagonal linear operator, which satisfies $\sup \sigma_P(A) = -\theta\pi^2 < 0$. By Lemma 89 it generates a strongly continuous semigroup $S$. This proves (i).

(ii): Let $|f|_{\mathrm{Lip}}$ and $|b|_{\mathrm{Lip}}$ denote the Lipschitz constants of $f$ and $b$ in the second variable. The mapping $F : H \to H$ is well-defined and Lipschitz continuous because for each $u, v \in H$,

$$\begin{aligned} \|F(u)\|_{L^2((0,1))} &\le \|f(\cdot,0)\|_{L^2((0,1))} + \|f(\cdot,u(\cdot)) - f(\cdot,0)\|_{L^2((0,1))} \\ &\le \|f(\cdot,0)\|_{L^2((0,1))} + |f|_{\mathrm{Lip}}\|u\|_{L^2((0,1))} < \infty, \\ \|F(u) - F(v)\|_{L^2((0,1))} &\le |f|_{\mathrm{Lip}}\|u - v\|_{L^2((0,1))} < \infty. \end{aligned}$$

The mapping $B : H \to L_2(H; H_\beta)$ is well-defined because for each $v \in H$,

$$\begin{aligned} \|B(v)\|^2_{L_2(U;H_\beta} &= \sum_{k=1}^\infty \|b(\cdot, v(\cdot))e_k(\cdot)\|^2_{H_\beta} \\ &= \sum_{k=1}^\infty \|(-A)^\beta b(\cdot, v(\cdot))e_k(\cdot)\|^2_H && ((-A)^\beta \text{ is isometry}) \\ &= \sum_{k,l=1}^\infty |\langle e_l, (-A)^\beta b(\cdot, v(\cdot))e_k(\cdot)\rangle_H|^2 && (\text{Parseval}) \\ &= \sum_{l=1}^\infty \|(-A)^\beta e_l\|^2_H \sum_{k=1}^\infty |\langle e_l, b(\cdot, v(\cdot))e_k(\cdot)\rangle_H|^2 && (e_l \text{ is eigenvector}) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{l=1}^{\infty} \|(-A)^{\beta} e_l\|_H^2 \|b(\cdot, v(\cdot)) e_l(\cdot)\|_H^2 && \left(\underset{\text{Parseval}}{\langle e_l, b e_k\rangle_H = \langle b e_l, e_k\rangle_H,}\right) \\
&\leq 2 \sum_{l=1}^{\infty} \|(-A)^{\beta} e_l\|_H^2 \|b(\cdot, v(\cdot))\|_H^2 && (\|e_l\|_{L^\infty((0,1))} = \sqrt{2}) \\
&= 2\|b(\cdot, v(\cdot))\|_H^2 \|(-A)^{\beta}\|_{L_2(U;H)}^2 < \infty. && (\textstyle\sum_{k=1}^{\infty} k^{4\beta} < \infty).
\end{aligned}
$$

Moreover, $B : H \to L_2(H; H_\beta)$ is Lipschitz continuous because for each $v, w \in H$, by a similar estimate as above,

$$
\begin{aligned}
\|B(v) - B(w)\|_{L_2(U;H_\beta}^2 &\leq 2\|b(\cdot, v(\cdot)) - b(\cdot, w(\cdot))\|_H^2 \|(-A)^{\beta}\|_{L_2(U;H)}^2 \\
&\leq 2|b|_{\mathrm{Lip}} \|v - w\|_H^2 < \infty.
\end{aligned}
$$

(iii): This follows from Theorem 91 with $\gamma = 0$, $p = 2$, $\eta \in (1/2, 1)$ using (i)–(ii). ☐

## 9.2 Literature

This section is taken from [Jen15, Section 7.2.1]. The stochastic heat equation is also called continuous-time parabolic Anderson model. Higher-dimensional analogues of the equation are studied in the context of stochastic quantization (see [DZ14, Section 13.7] for an overview).

# 10 Stochastic wave equation

## 10.1 Overview of wave equations

The following is an overview of some well-known equations for the height $u(x,t)$ of a wave at time $t$ and location $x$. Many qualitative properties of the equation can be seen by studying traveling wave solutions, i.e., solutions of the form $u(x,t) = f(x - ct)$, where $f$ is a function and $c$ is a constant.

| Name | Equation | Traveling wave solutions |
|---|---|---|
| Transport | $u_t + u_x = 0$ | $u(x,t) = f(x - t)$ |
| Wave | $u_{tt} - u_{xx} = 0$ | $u(x,t) = f(x \pm t)$ |
| Klein-Gordon | $u_{tt} - u_{xx} + u = 0$ | $u(x,t) = a\cos(k(x - ct - x_0))$ |
| | | $c^2 < 1$, $k = (1 - c^2)^{-1/2}$ |
| sine-Gordon | $u_{tt} - u_{xx} + \sin(u) = 0$ | $u(x,t) = \frac{2}{\pi}\arctan(e^{\sqrt{2\pi}k(x-ct-x_0)})$ |
| | | $c^2 < 1$, $k = \pm(1 - c^2)^{-1/2}$ |
| Airy | $u_t + u_{xxx} = 0$ | $u(x,t) = a\cos(k(x - ct - x_0)) + b$ |
| | | $c < 0$, $k = (-c)^{1/2}$ |
| KdV | $u_t + uu_x + u_{xxx} = 0$ | $u(x,t) = 3c\,\mathrm{sech}^2(k/2(x - ct - x_0))$ |
| | | $c > 0$, $k = c^{1/2}$ |
| Schrödinger | $\sqrt{-1}u_t = -u_{xx}$ | $u(x,t) = e^{\sqrt{-1}k(x-ct-x_0)}$ |
| | | $c \neq 0$, $k = c$ |

One speaks of dispersion if there are multiple possible values of $c$, of causality if compactly supported initial data leads to compactly supported solutions, of a conservation law if the equation is of the form $u_t = \partial_x(\dots)$, of a wave train if

$f$ is periodic, of a wave front if $f$ is monotonic, and of a solitary wave if $f$ has a unique local maximum.

## 10.2 Existence and uniqueness

**95 Setting.** We want to analyze the equation

$$\partial_t^2 X_t(x) = \theta \Delta X_t(x) + f(x, X_t(x)) + b(x, X_t(x)) \dot{W}_t(x), \qquad x \in (0,1),$$

where $\dot{W}$ is space-time white noise, by rewriting it as a first-order system

$$\begin{cases} dX_t(x) = Y_t(x) dt, \\ dY_t(x) = \big(\theta \Delta X_t(x) + f(x, X_t(x))\big) dt + b(x, X_t(x)) dW_t(x), \end{cases}$$

in the following setting:

(i) $T \in [0, \infty)$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0,T]})$ is a stochastic basis, $U = L^2((0,1))$, and $(W_t)_{t \in [0,T]}$ is a $I_U$-Brownian motion with respect to $(\mathcal{F}_t)_{t \in [0,T]}$.

(ii) $H = U$, $\theta > 0$, and $\Delta : H_0^2((0,1)) \subset H \to H$ is the Laplace operator.

(iii) $f, b : (0,1) \times \mathbb{R} \to \mathbb{R}$ are measurable functions, are Lipschitz in the second variable, and satisfy $f(\cdot, 0), b(\cdot, 0) \in L^2((0,1))$.

**96 Theorem.** *Under the assumptions of Setting 95, the following statements hold:*

*(i) The linear operator $\theta\Delta$ is diagonal, satisfies* $\sup \sigma_P(\theta\Delta) < 0$, *and has an associated family $(H_r)_{r \in \mathbb{R}}$ of interpolation spaces.*

*(ii) The linear operator*

$$A : H_{1/2} \times H_0 \subset H_0 \times H_{-1/2} \to H_0 \times H_{-1/2}, \qquad (v, w) \mapsto (w, \theta\, \Delta v),$$

*generates a strongly continuous group of isometries on $H_0 \times H_{-1/2}$.*

*(iii) The following mappings are well-defined and Lipschitz continuous,*

$$\begin{aligned} &F : H_0 \times H_{-1/2} \to H_0 \times H_{-1/2}, && F(v,w)(x) = \big(0, f(x, v(x))\big), \\ &B : H_0 \times H_{-1/2} \to L_2(U; H_0 \times H_{-1/2}), && B(v,w)(u)(x) = \big(0, b(x, v(x)) u(x)\big). \end{aligned}$$

*(iv) For each $\xi \in L^2((\Omega, \mathcal{F}_0, \mathbb{P}); H_0 \times H_{-1/2})$ there exists a unique up to modifications predictable process $(X, Y) : [0,T] \times \Omega \to H_0 \times H_{-1/2}$ which satisfies for each $t \in [0,T]$ that*

$$(X_t, Y_t) = e^{At}\xi + \int_0^t e^{A(t-s)} F(X_s, Y_s) ds + \int_0^t e^{A(t-s)} B(X_s, Y_s) dW_s$$

*and*

$$\mathbb{P}\left[ \int_0^T \left( \|X_s\|_{H_0}^2 + \|Y_s\|_{H_{-1/2}}^2 \right) ds < \infty \right] = 1.$$

*Proof.* (i): This was shown in Theorem 94.(i).

(ii): $A^*$ is an extension of $-A$ because for each $(v_1, w_1), (v_2, w_2) \in D(A)$,

$$\begin{aligned}
&\langle A(v_1, w_1), (v_2, w_2)\rangle_{H_0 \times H_{-1/2}} = \langle w_1, v_2\rangle_{H_0} + \langle \theta\Delta v_1, w_2\rangle_{H_{-1/2}} \\
&\quad = \langle (-\theta\Delta)^{1/2} w_1, (-\theta\Delta)^{1/2} v_2\rangle_{H_{-1/2}} + \langle (-\theta\Delta)^{-1/2}\theta\Delta v_1, (-\theta\Delta)^{-1/2} w_2\rangle_{H_0} \\
&\quad = \langle w_1, -\theta\Delta v_2\rangle_{H_{-1/2}} - \langle v_1, w_2\rangle_{H_0} = -\langle (v_1, w_1), A(v_2, w_2)\rangle_{H_0 \times H_{-1/2}}.
\end{aligned}$$

To see that $A^* = -A$ let $(v, w) \in D(A^*)$. Then the following linear mapping is bounded:

$$H_{1/2} \times H_0 \subset H_0 \times H_{-1/2} \to \mathbb{R}, \qquad (h, k) \mapsto \langle A(h, k), (v, w)\rangle_{H_0 \times H_{-1/2}}.$$

Rewriting the last expression as

$$\begin{aligned}
\langle A(h, k), (v, w)\rangle_{H_0 \times H_{-1/2}} &= \langle k, v\rangle_{H_0} + \langle \theta\Delta h, w\rangle_{H_{-1/2}} \\
&= \langle k, v\rangle_{H_0} + \langle -(-\theta\Delta)^{1/2} h, (-\theta\Delta)^{-1/2} w\rangle_{H_0}
\end{aligned}$$

and using that $(-\theta\Delta)^{1/2} : H_0 \to H_{-1/2}$ is an isometry shows that the following linear mappings are bounded,

$$\begin{aligned}
&H_0 \subset H_{-1/2} \to \mathbb{R}, &&k \mapsto \langle k, v\rangle_{H_0}, \\
&H_0 \subset H_{-1/2} \to \mathbb{R}, &&h \mapsto \langle h, (-\theta\Delta)^{-1/2} w\rangle_{H_0}.
\end{aligned}$$

By [NJW15, Lemma 3.10.(ii)] this implies that $v$ and $(-\theta\Delta)^{-1/2} w$ belong to $H_{1/2}$, which is equivalent to $(v, w) \in H_{1/2} \times H_0 = D(A)$. This proves that $A^* = -A$. It follows from a theorem of Stone [EN99, Theorem 3.24] that $A$ generates a strongly continuous group of isometries. This proves (ii).

(iii): This follows from Theorem 94.(ii).

(iv): This follows from Theorem 91 using (ii)–(iii). □

### 10.3 Literature

The overview of wave equations is inspired by Peter D. Miller's lecture notes [Mil06]. The stochastic wave equation is also called continuous-time hyperbolic Anderson model. An extensive analysis of the stochastic wave equation can be found in [NJW15].

## 11 Stochastic Schrödinger equation

### 11.1 Quantum mechanics

**97 Definition.** According to John von Neumann's and Paul Dirac's axiomatization of quantum mechanics and Schrödinger's view of quantum dynamics, a physical system is described by the following ingredients:

- **States** are one-dimensional subspaces of a separable complex Hilbert space $H$. Thus, for every non-zero $v \in H$, $\operatorname{span}_{\mathbb{C}}\{v\}$ is a state.
- **Observables** are self-adjoint linear operators on $H$. The expectation (in the sense of probability) of an observable $T$ under a state $\operatorname{span}_{\mathbb{C}}\{v\}$ with $\|v\| = 1$ is defined as $\langle Tv, v\rangle_H$.

- **Dynamics** of the states $X : \mathbb{R}_+ \to H$ are encoded in Schrödinger's equation
$$\sqrt{-1}\hbar dX_t = AX_t dt,$$
where $\hbar$ is a constant and $A$ is a self-adjoint linear operator on $H$, which is called the Hamiltonian. Note that Schrödinger's equation is norm-preserving, i.e., $-\sqrt{-1}A$ generates a strongly continuous group of isometries on $H$ by a theorem of Stone [EN99, Theorem 3.24].

**98 Example** (Harmonic oscillator)**.** $H = L^2(\mathbb{R};\mathbb{C})$ and

$$A(v)(x) = -\frac{\hbar^2}{2m}\Delta v(x) + \frac{1}{2}m\omega^2 x^2 v(x),$$

where $m$ and $\omega$ are constants. A typical observable is the position operator $Tv(x) = xv(x)$.

**99 Remark.** There are several extensions and modifications of Schrödinger's equation. For example, one may add a nonlinearities and noise. The noise is typically added in Stratonovich form to preserve the property that $\|X_t\|_H = \|X_0\|_H$. This leads to equations of the form

$$\sqrt{-1}\hbar dX_t = \big(AX_t + F(X_t)\big)dt + B(X_t) \circ dW_t,$$

where $F : H \to H$, $B : H \to L_2(U;H)$, $W$ is real-valued Brownian motion on a Hilbert space $U$, and $\circ dW_t$ is the Stratonovich integral. This equation can be recast in Itō form as

$$\sqrt{-1}\hbar dX_t = \left(AX_t + F(X_t) + \frac{1}{2}\operatorname{Tr}_{U_Q} B'(X_t)B(X_t)\right) dt + B(X_t) \circ dW_t,$$

provided that

$$\operatorname{Tr}_{U_Q} B'(X_t)B(X_t) = \sum_{u\in\mathbb{U}} \Big(B'(X_t)\big(B(X_t)u\big)\Big)u \in H,$$

where $\mathbb{U}$ is an orthonormal basis of the reproducing kernel Hilbert space $U_Q$ of the cylindrical Brownian motion $W$.

## 11.2 Existence and uniqueness

**100 Setting.** We want to analyze the equation

$$\sqrt{-1}\, dX_t(x) = \big(\Delta X_t(x) + f(x, X_t(x))\big)dt + b(x, X_t(x))dW_t(x), \qquad x \in \mathbb{R}^d,$$

in the following setting:

(i) $T \in [0,\infty)$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ is a stochastic basis, $d \in \mathbb{N}\backslash\{0\}$, $r > d/2$, $U = H^r(\mathbb{R}^d;\mathbb{C})$, $Q \in L_1(U)$, and $(W_t)_{t\in[0,T]}$ is a $Q$-Brownian motion with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

(ii) $H = L^2(\mathbb{R}^d;\mathbb{C})$ and $\Delta : H^2(\mathbb{R}^d;\mathbb{C}) \subset H \to H$ is the Laplace operator.

(iii) $f, b : (0,1) \times \mathbb{R} \to \mathbb{R}$ are measurable functions, are Lipschitz in the second variable, and satisfy $f(\cdot, 0), b(\cdot, 0) \in L^2((0,1))$.

**101 Theorem.** *Under the assumptions of Setting 93, the following statements hold:*

*(i)* $A := \sqrt{-1}\,\Delta$ *generates a strongly continuous group of isometries on* $H$.

*(ii) The following mappings are well-defined and Lipschitz continuous,*

$$\begin{aligned} &F : H \to H, && F(v)(x) = -\sqrt{-1} f(x, v(x)), \\ &B : H \to L_2(U_Q; H), && B(v)(u)(x) = -\sqrt{-1} b(x, v(x))u(x). \end{aligned}$$

*(iii) For each* $\xi \in L^2((\Omega, \mathcal{F}_0, \mathbb{P}); H)$ *there exist a unique up to modifications predictable process* $X : [0, T] \times \Omega \to H$ *which satisfies for each* $t \in [0, T]$ *that*

$$X_t = e^{At}\xi + \int_0^t e^{A(t-s)} F(X_s)ds + \int_0^t e^{A(t-s)} B(X_s)dW_s$$

*and*

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

*Proof.* (i): Let ˆ denote the Fourier transform. For each $v \in H$ and $t \in \mathbb{R}$ let $S_t v$ be the unique element of $H$ which satisfies for each $\xi \in \mathbb{R}$ that

$$\widehat{S_t v}(\xi) = \exp(-it\xi^2)\widehat{v}(\xi). \tag{12}$$

Thus, in the Fourier domain, $S_t$ is a multiplication operator by a function of absolute value one. It is easy to verify that $S$ is a strongly continuous group of isometries on $H$, whose generator is $A$ [EN99, Propositions 4.11 and 4.12].

(ii): The Lipschitz continuity of $F$ follows by the same argument as in the proof of Theorem 94.(ii). The Lipschitz continuity of $B$ can be seen as follows. First, the mapping

$$H \to H, \qquad v \mapsto b(\cdot, v(\cdot))$$

is Lipschitz by the same argument as in the proof of Theorem 94.(ii). Second, by the Sobolev embedding theorem $U$ is continuously embedded in $L^\infty(\mathbb{R}^d; \mathbb{C})$, and therefore multiplication $H \times U \to H$ is continuous. Third, the embedding $i_Q : U_Q \to U$ is Hilbert-Schmidt because for any orthonormal basis $\mathbb{B}$ of $U$ and any orthonormal basis $\mathbb{U}$ of $U_Q$,

$$\begin{aligned} \|i_Q\|_{L_2(U_Q;U)} &= \sum_{u\in\mathbb{U}} \|i_Q u\|_U^2 = \sum_{u\in\mathbb{U}}\sum_{v\in\mathbb{B}} |\langle i_Q u, v\rangle_U|^2 = \sum_{u\in\mathbb{U}}\sum_{v\in\mathbb{B}} |\langle u, Qv\rangle_{U_Q}|^2 \\ &= \sum_{v\in\mathbb{B}} \|Qv\|_{U_Q}^2 = \sum_{v\in\mathbb{B}} \langle Qv, v\rangle_U^2 = \operatorname{Tr}(Q) \le \|Q\|_{L_1(U)} < \infty. \end{aligned}$$

The mapping $B$ is a composition of these three mappings, and it follows that $B$ is Lipschitz continuous. This shows (ii).

(iii): This follows from Theorem 91 using (i)–(ii). ☐

### 11.3 Literature

The following variant of Schrödinger's equation is a basic model for nonlinear waves,

$$\sqrt{-1}\, dX_t(x) = \big(\Delta X_t(x) + |X_t(x)|^{2\sigma})\big)dt + X_t(x) \circ dW_t(x), \qquad x \in \mathbb{R}^d,$$

where $\sigma > 0$ is a constant and $\circ dW_t$ is a Stratonovich integral (see [DZ14, Section 13.23] for an overview). This equation does not fit into Setting 100 because of the non-Lipschitz drift. Solutions exist only locally in time and blow up with positive probability in finite time [BD02].

## 12 Stochastic linearized Korteweg–de Vries equation

### 12.1 History of the Korteweg–de Vries equation

In the 19th century much research was devoted to the study of water waves, particularly in England and France (see [Jag] for a historical account).

- In 1834 naval architect Scott Russel observed a solitary wave, which was traveling in the Union Canal between Edinburgh and Glasgow at a speed of about 13 km/h without changing its shape. This led him to perform extensive experiments and to search for a mathematical model of such "great traveling waves."
- Some of Russel's contemporaries, including foremost Airy, dismissed Russel's observation as impossible. Others, including Rayleigh, Boussinesq, and finally Korteweg and de Vries [KV95] took up the quest and derived new shallow-water limits of the Navier-Stokes equations. These new equations admit solitary traveling waves, similar to the ones observed by Russel.

### 12.2 Existence and uniqueness

**102 Setting.** We want to study a perturbed linearized KdV (or Airy) equation,

$$dX_t(x) = \big(AX_t(x) + f\big(x, X_t(x)\big)\big)\, dt + b\big(x, X_t(x)\big)dW_t, \qquad x \in \mathbb{R},$$

in the following setting:

(i) $T \in [0,\infty)$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ is a stochastic basis, $d \in \mathbb{N}\backslash\{0\}$, $r > 1/2$, $U = H^r(\mathbb{R})$, $Q \in L_1(U)$, and $(W_t)_{t\in[0,T]}$ is a $Q$-Brownian motion with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

(ii) $H = L^2(\mathbb{R})$ and $A : H^3(\mathbb{R}) \subset H \to H$ is given by $u \mapsto -u'''$.

(iii) $f, b : (0,1) \times \mathbb{R} \to \mathbb{R}$ are measurable functions, are Lipschitz in the second variable, and satisfy $f(\cdot, 0), b(\cdot, 0) \in L^2((0,1))$.

**103 Theorem.** *Under the assumptions of Setting 102, the following statements hold:*

*(i) $A$ generates a strongly continuous group of isometries on $H$.*

*(ii) The following mappings are well-defined and Lipschitz continuous,*

$$\begin{aligned} &F : H \to H, && F(v)(x) = f(x, v(x)), \\ &B : H \to L_2(U_Q; H), && B(v)(u)(x) = b(x, v(x))u(x). \end{aligned}$$

*(iii) For each $\xi \in L^2((\Omega, \mathcal{F}_0, \mathbb{P}); H)$ there exist a unique up to modifications predictable process $X : [0,T] \times \Omega \to H$ which satisfies for each $t \in [0,T]$ that*

$$X_t = e^{At}\xi + \int_0^t e^{A(t-s)} F(X_s)ds + \int_0^t e^{A(t-s)} B(X_s)dW_s$$

*and*

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

*Proof.* (i): Let $\hat{}$ denote the Fourier transform, and let $i = \sqrt{-1}$. For each $v \in H$ and $t \in [0, \infty)$ let $S_t v$ be the unique element of $H$ which satisfies for each $\xi \in \mathbb{R}$ that

$$\widehat{S_t v}(\xi) = \exp(it\xi^3)\widehat{v}(\xi). \tag{13}$$

Thus, in the Fourier domain, $S_t$ is a multiplication operator by a function of absolute value one. It is easy to verify that $S$ is a strongly continuous group of isometries on $H$, whose generator is $A$ [EN99, Propositions 4.11 and 4.12].

(ii): This follows as in Theorem 101.(ii).

(iii): This follows from Theorem 91 using (i)–(ii). □

### 12.3 Literature

The historical account is inspired by [Jag]. It is possible to replace the colored noise by white noise and to add the nonlinearity $uu_x$, but this requires a different fixed point argument [BD98; BDT99].

# References


[AB06] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis. A Hitchhiker's guide.* 3rd ed. Springer, 2006.

[BD02] A. de Bouard and A. Debussche. "On the effect of a noise on the solutions of the focusing supercritical nonlinear Schrödinger equation". In: *Probability theory and related fields* 123.1 (2002), pp. 76–96.

[BD98] A. de Bouard and A. Debussche. "On the stochastic Korteweg–de Vries equation". In: *Journal of Functional Analysis* 154.1 (1998), pp. 215–251.

[BDT99] A. de Bouard, A. Debussche, and Y. Tsutsumi. "White noise driven Korteweg–de Vries equation". In: *Journal of Functional Analysis* 169.2 (1999), pp. 532–558.

[BM07] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice.* Springer, 2007.

[CT07] R. Carmona and M. R. Tehranchi. *Interest rate models: an infinite dimensional stochastic analysis perspective.* Springer Science & Business Media, 2007.

[DF92] A. Defant and K. Floret. *Tensor norms and operator ideals.* Vol. 176. Elsevier, 1992.

[DZ14] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions.* Cambridge university press, 2014.

[Els11] J. Elstrodt. *Maß-und Integrationstheorie.* 7th ed. Springer, 2011.

[EN99] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations.* Vol. 194. Springer Science & Business Media, 1999.

[Fil01] D. Filipović. *Consistency problems for Heath-Jarrow-Morton interest rate models.* 1760. Springer, 2001.

[Fil09] D. Filipović. *Term-structure models.* Springer Finance. A graduate course. Berlin: Springer-Verlag, 2009.

[Gra16] G. Grafendorfer. "Infinite-dimensional affine processes". PhD thesis. ETH Zürich, 2016.

[Hai09] M. Hairer. *An introduction to stochastic PDEs.* 2009. arXiv: `0907.4178`.

[Jag] E. M. de Jager. *On the origin of the Korteweg-de Vries equation.* arXiv: `math/0602661`.

[Jen15] A. Jentzen. *Stochastic Partial Differential Equations: Analysis and Numerical Approximations.* Lecture Notes, ETH Zürich. 2015.

[Jen16] A. Jentzen. *Stochastic Partial Differential Equations: Analysis and Numerical Approximations.* Lecture Notes, ETH Zürich. 2016.

[KV95] D. Korteweg and G. de Vries. "On the change of form of long waves advancing in a rectangular channel, and a new type of long stationary wave". In: *Philos. Mag* 39 (1895), pp. 422–443.

[Lan93] S. Lang. *Real and functional analysis.* 3rd ed. Vol. 142. Graduate Texts in Mathematics. Springer, 1993.

[Mil06] P. D. Miller. *Integrable Systems, Nonlinear Waves, and Soliton Theory.* Lecture Notes Math 651. University of Michigan, 2006. URL: `http://math.arizona.edu/~mcl/Miller/`.

[NJW15] L. J. de Naurois, A. Jentzen, and T. Welti. *Weak convergence rates for spatial spectral Galerkin approximations of semilinear stochastic wave equations with multiplicative noise.* 2015. arXiv: `1508.05168`.

[PR07] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations.* Vol. 1905. Springer, 2007.

[Roz90] B. L. Rozovskii. *Stochastic Evolution Systems. Linear Theory and Applications to Non-linear Filtering.* Vol. 35. Mathematics and Its Applications (Soviet Series). Springer, 1990.

[Rya02] R. A. Ryan. *Introduction to Tensor Products of Banach Spaces.* Springer Monographs in Mathematics. Springer, 2002.

[Van08] J. Van Neerven. *Stochastic evolution equations.* ISEM lecture notes. 2007–2008.