BMS Advanced Course

Stochastic Partial Differential Equations

Michael Scheutzow

Lecture Notes

Technische Universität Berlin

Sommersemester 2019

PRELIMINARY VERSION
JULY 10TH 2019

Contents

Chapter 1

Introduction

Here are some basic examples of partial differential equations (PDEs).

- Heat equation:
$$\frac{\partial u}{\partial t}(t,x) = c\,\Delta u(t,x),\ x \in D \subseteq \mathbb{R}^d,\ t \geq 0,$$
- Wave equation:
$$\frac{\partial^2 u}{\partial t^2}(t,x) = c\,\Delta u(t,x),\ x \in D \subseteq \mathbb{R}^d,\ t \geq 0,$$
- Reaction diffusion equation:
$$\frac{\partial u}{\partial t}(t,x) = c\,\Delta u(t,x) - u^3(t,x),$$

where $c > 0$ and each equation is equipped with suitable initial and boundary conditions.

One way of obtaining a *stochastic partial differential equation* is to add noise to the right hand side of the equation (which may or may not depend on the solution u). In many cases (such as for the heat equation and the reaction diffusion equation) such equations can be represented in the following *evolutionary form*:

$$\mathrm{d}X_t = A(t, X_t)\,\mathrm{d}t + B(t, X_t)\,\mathrm{d}W_t,\ X_0 = x, \tag{1.0.1}$$

where X takes values in a suitable infinite dimensional Hilbert space H and where W is a U-valued Wiener process where U is another Hilbert space and where B takes values in the space $L(U, H)$ of linear operators from U to H. For this to make sense we need to

- define the concept of a U-valued Wiener process when U is infinite dimensional,
- introduce stochastic integrals in infinite dimensions,
- clarify what we mean by a solution of (1.0.1),
- find conditions on A and B which guarantee existence and uniqueness of solutions of (1.0.1).

There are three different approaches to SPDEs in the mathematical literature, namely:

- the martingale measures approach, see [W86],
- the semigroup approach, see [DZ92],
- the variational approach, see [PR07] or [LR15].

Our emphasis in this lecture will be on the variational approach and we will largely follow the corresponding references above but we will also discuss the two other approaches. In order to obtain some idea of the semigroup approach, consider the following SDE in finite dimensions:

$$\mathrm{d}X_t = A \cdot X_t \,\mathrm{d}t + F(X_t)\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t,\ X_0 = x \in \mathbb{R}^d,$$

where A is a $d \times d$-matrix, W is an m-dimensional Wiener process and B is a map from $\mathbb{R}^d$ to $\mathbb{R}^{d\times m}$. Under suitable assumptions (which we discussed in WT3), this SDE has a unique strong solution and it is easy to see that it can be represented in the form

$$X_t = \mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)}F(X_s)\,\mathrm{d}s + \int_0^t \mathrm{e}^{A(t-s)}B(X_s)\,\mathrm{d}W_s$$

which corresponds to the variation-of-constants formula known from the theory of ODEs. Such a formula also holds in the infinite dimensional case (under suitable assumptions) and the corresponding solution X is then called a *mild solution*. It turns out that the three approaches are not equivalent in the sense that they cannot be applied to exactly the same family of SPDEs. For example, to apply the semigroup approach, one needs a dominant linear part A which is not required in the variational approach.

Chapter 2

The Bochner integral

This chapter follows [LR15, Appendix A] closely (see also [PR07, Appendix A]).

2.1 Definition of the Bochner integral

In the whole chapter $(X, \|.\|)$ is a Banach space and $(\Omega, \mathcal{F}, \mu)$ is a finite measure space. We want to define the integral $\int f \,\mathrm{d}\mu$ for a sufficiently large class of functions $f : \Omega \to X$.

As usual, we start by defining the integral for simple functions. Set

$$\mathcal{E} := \Big\{ f : \Omega \to X \Big| \, f = \sum_{k=1}^{n} x_k 1_{A_k},\, x_k \in X,\, A_k \in \mathcal{F},\, 1 \le k \le n,\, n \in \mathbb{N} \Big\},$$

and define a semi-norm $\|.\|_{\mathcal{E}}$ on the linear space $\mathcal{E}$ by

$$\|f\|_{\mathcal{E}} := \int \|f\| \,\mathrm{d}\mu,\ f \in \mathcal{E}.$$

After taking equivalence classes with respect to $\|.\|_{\mathcal{E}}$ (which we will do from now on without changing notation), $(\mathcal{E}, \|.\|_{\mathcal{E}})$ becomes a normed linear space. For $f \in \mathcal{E}$ with representation $f = \sum_{k=1}^{n} x_k 1_{A_k}$, we define

$$\int f \,\mathrm{d}\mu := \sum_{k=1}^{n} x_k \mu(A_k).$$

It is not hard to see that this definition does not depend on the particular choice of the representation and that the map $\mathrm{int} : (\mathcal{E}, \|.\|_{\mathcal{E}}) \to (X, \|.\|)$ which maps f to $\int f \,\mathrm{d}\mu$ is linear. The map int is also Lipschitz (with constant 1), since

$$\| \int f \,\mathrm{d}\mu \| \le \int \|f\| \,\mathrm{d}\mu = \|f\|_{\mathcal{E}},$$

and therefore uniformly continuous and can hence be uniquely extended to a continuous linear map from the completion $\bar{\mathcal{E}}$ of $\mathcal{E}$ with respect to $\|.\|_{\mathcal{E}}$. Just like in the case of stochastic interals in WT3 it is desirable to have a more concrete representation of $\bar{\mathcal{E}}$.

Definition 2.1. A function $f : \Omega \to X$ is called *strongly measurable* if it is $\mathcal{F}-\mathcal{B}(X)$-measurable and $f(\Omega)$ is separable.

We will provide an example of a measurable function which is not strongly measurable in class. Don't miss it!

Definition 2.2. Let $1 \le p < \infty$. Then we define

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mu; X) := \Big\{ f : \Omega \to X \Big| \, f \text{ is strongly measurable and } \int \|f\|^p \, \mathrm{d}\mu < \infty \Big\}$$

and the semi-norm

$$\|f\|_{L^p} := \Big(\int \|f\|^p \, \mathrm{d}\mu \Big)^{1/p}, \; f \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu; X).$$

The corresponding space of equivalence classes is denoted by $L^p(\Omega, \mathcal{F}, \mu; X)$.

Theorem 2.3. *$L^1(\Omega, \mathcal{F}, \mu; X) = \bar{\mathcal{E}}$.*

The following proposition is a kind of replacement for the monotone (or dominated) convergence theorem (which do not make sense for X-valued functions).

Proposition 2.4. *For each $f \in L^1(\Omega, \mathcal{F}, \mu; X)$ there exists a sequence of $f_n \in \mathcal{E}$ such that $\|f_n(\omega) - f(\omega)\| \downarrow 0$ for all $\omega \in \Omega$ as $n \to \infty$.*

The result is an immediate consequence of the following lemma (see [DZ92, Lemma 1.1]).

Lemma 2.5. *Let (E, ρ) be a separable metric space and X an E-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. There exists a sequence of random variables X_n taking only a finite number of values each such that $\rho(X_n(\omega), X(\omega)) \downarrow 0$ for every $\omega \in \Omega$.*

Proof. Let $E_0 = \{e_1, e_2, ...\}$ be a countable dense set in E. For $n \in \mathbb{N}$ define

$$\begin{aligned} \rho_n(\omega) &= \min\{\rho(X(\omega), e_k), \, k = 1, ..., n\} \\ k_n(\omega) &= \min\{k \le n : \rho_n(\omega) = \rho(X(\omega), e_k)\} \\ X_n(\omega) &= e_{k_n(\omega)}. \end{aligned}$$

Then each X_n is a random variable with values in the finite set $\{e_1, ..., e_n\}$ and $\rho(X_n(\omega), X(\omega)) \downarrow 0$ for every $\omega \in \Omega$. □

2.2 Properties of the Bochner integral

We first state Bochner's inequality.

Proposition 2.6. *Let $f \in L^1(\Omega, \mathcal{F}, \mu; X)$. Then*

$$\Big\| \int f \, \mathrm{d}\mu \Big\| \le \int \|f\| \, \mathrm{d}\mu.$$

Proof. The claim holds for each $f \in \mathcal{E}$ and extends to $\bar{\mathcal{E}}$ by continuity of the extension $\overline{\mathrm{int}}$ of int. □

Proposition 2.7. *Let $f \in L^1(\Omega, \mathcal{F}, \mu; X)$. Then*

$$\int L \circ f \, \mathrm{d}\mu = L\Big(\int f \, \mathrm{d}\mu \Big)$$

for every $L \in L(X, Y)$, where Y is another Banach space.

Chapter 3

Hilbert space valued Wiener processes

3.1 Some background in functional analysis

Most of this section is taken from [PR07, Appendix B] (or [LR15, Appendix B]) and [DZ92, Appendix C]. Proofs or at least references to proofs of the propositions below can be found there.

Let $(U, \|.\|_U)$ and $(V, \|.\|_V)$ be real Banach spaces. As usual, $L(U, V)$ denotes the set of all continuous (or bounded) linear operators from U to V and we often write $L(U)$ instead of $L(U, U)$ when $U = V$. $U^* := L(U, \mathbb{R})$ is called the *dual space* of U. $L(U, V)$ is itself a Banach space with respect to the usual addition and multiplication by scalars and the norm

$$\|f\|_{U,V} := \sup_{u \in U : \|u\|_U = 1} \|f(u)\|_V.$$

We point out that $L(U, V)$ is generally *not* separable even if U and V are. If, for example, $U = l^2$ is the separable Hilbert space of square summable sequences, then $L(U)$ is not separable (nor is it a Hilbert space). Sometimes we write $\|.\|_{U^*}$ instead of $\|.\|_{U,\mathbb{R}}$. We encourage the reader to review the important Hahn-Banach theorem on the extension of bounded linear functionals. We will use it on several occasions.

Let us now consider the special case in which $(U, \langle ., . \rangle_U)$ and $(V, \langle ., . \rangle_V)$ are real separable Hilbert spaces.

Definition 3.1. $T \in L(U, V)$ is called a *nuclear operator* if there exist sequences $(a_j)_{j \in \mathbb{N}}$ in V and $(b_j)_{j \in \mathbb{N}}$ in U such that

$$Tx = \sum_{j=1}^{\infty} a_j \langle b_j, x \rangle_U \text{ for all } x \in U$$

and

$$\sum_{j=1}^{\infty} \|a_j\|_V \|b_j\|_U < \infty.$$

The space of all nuclear operators from U to H is denoted by $L_1(U, V)$. We write $L_1(U)$ instead of $L_1(U, U)$ in case $U = V$.

Proposition 3.2. *The space $L_1(U,V)$ endowed with the norm*

$$\|T\|_{L_1(U,V)} := \inf\Big\{\sum_{j\in\mathbb{N}} \|a_j\|_V \|b_j\|_U \Big| Tx = \sum_{j=1}^{\infty} a_j \langle b_j, x\rangle_U,\, x \in U\Big\}$$

is a Banach space. Further, $\|T\|_{U,V} \le \|T\|_{L_1(U,V)}$.

NB: The last claim in the previous proposition is very easy to show! Try it!

Definition 3.3. Let $T \in L_1(U)$ and let (e_k) be an ONB of U. Then $\operatorname{tr} T := \sum_k \langle Te_k, e_k\rangle$ is called the *trace* of T. $T \in L_1(U)$ is called *trace class (operator)* in case T is nonnegative definite (or positive semi-definite), [i.e. $\langle Tu, v\rangle = \langle u, Tv\rangle$ and $\langle Tu, u\rangle \ge 0$ for every $u, v \in U$].

Proposition 3.4. *If $T \in L_1(U)$, then $\operatorname{tr} T$ is a well-defined real number (the sum converges absolutely) which does not depend on the choice of the ONB (e_k).*

Proof. [DZ92, Proposition C.1]. □

Proposition 3.5. *A non-negative definite operator $T \in L(U)$ is nuclear iff there exists an ONB (e_k) of U such that*

$$\sum_k \langle Te_k, e_k\rangle < \infty.$$

In this case $\operatorname{tr} T = \|T\|_{L_1(U,V)}$.

Proof. [DZ92, Proposition C.3]. □

Definition 3.6. $T \in L(U,V)$ is called a *Hilbert-Schmidt operator* if

$$\sum_j \|Te_j\|_V^2 < \infty, \tag{3.1.1}$$

where (e_j) is an ONB of U. The space of all Hilbert-Schmidt operators from U to V is denoted by $L_2(U,V)$.

Proposition 3.7. *The number $\|T\|_{L_2(U,V)} := \big(\sum_j \|Te_j\|_V^2\big)^{1/2} < \infty$ does not depend on the choice of the ONB. Further, $L_2(U,V)$ equipped with this norm is a separable Hilbert space.*

Proposition 3.8. *We have*

$$L_1(U,V) \subset L_2(U,V) \subset \mathcal{K}(U,V),$$

where $\mathcal{K}(U,V)$ denotes the space of all compact *linear operators, i.e. the space of all $T \in L(U,V)$ such that the closure of the image of the unit ball in U under the map T is a compact subset of V.*

Remark 3.9. If $v \in U$, then the map $u \mapsto \langle v, u\rangle$, $u \in U$ defines a linear continuous map from U to $\mathbb{R}$, i.e. an element in U^*. The famous *Riesz representation theorem* states that the converse is also true: if $\bar{v} \in U^*$, then there exists a unique $v \in U$ such that $\bar{v}(u) = \langle v, u\rangle$ for every $u \in U$. Clearly, $\|\bar{v}\|_{U^*} = \|v\|_U$. In particular, separability of U implies separability of U^* (the same statement does not hold true for Banach spaces).

3.2 Gaussian measures on Hilbert spaces

In this section we will introduce Gaussian measures on Hilbert spaces (for the more general case of Gaussian measures on Banach spaces the reader is referred to [B98]). Until further notice, $(U, \langle ., . \rangle)$ will be a real separable Hilbert space and we will denote its Borel σ-algebra (with respect to the topology induced by the norm $\|.\|$ associated to the inner product $\langle ., . \rangle$) by $\mathcal{B}(U)$.

Definition 3.10. A probability measure μ on $(U, \mathcal{B}(U))$ is called *Gaussian* if for all $v \in U$ the (bounded, linear) mapping $u \mapsto \langle u, v \rangle$ from $(U, \mathcal{B}(U))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ has a Gaussian law.

Remark 3.11. If μ is Gaussian, then for each $m \in \mathbb{N}$ and $v_1, ..., v_m \in U$, the image $\mu^{(v_1,...,v_m)}$ of μ under the map $u \mapsto (\langle u, v_1 \rangle, ..., \langle u, v_m \rangle)$ is Gaussian. Recall that a probability measure on $\mathbb{R}^m$ is Gaussian iff the image under every linear combination of the coordinates is a Gaussian measure on $\mathbb{R}$, but for $\lambda_1, ..., \lambda_m \in \mathbb{R}$ we have $\lambda_1 \langle u, v_1 \rangle + ... + \lambda_m \langle u, v_m \rangle = \langle u, \lambda_1 v_1 + ... + \lambda_m v_m \rangle$ which has a Gaussian law by definition.

Theorem 3.12. *A probability measure μ on $(U, \mathcal{B}(U))$ is Gaussian iff there exists some $m \in U$ and some trace-class $Q \in L(U)$ such that*

$$\hat{\mu}(u) := \int_U \mathrm{e}^{\mathrm{i}\langle u,v \rangle} \, \mu(\mathrm{d}v) = \mathrm{e}^{\mathrm{i}\langle m,u \rangle - \frac{1}{2}\langle Qu,u \rangle}, \; u \in U. \tag{3.2.1}$$

In this case we write $\mathcal{N}(m, Q)$ instead of μ. m is called mean *of μ and Q is called* covariance (operator) *of μ. The Gaussian measure μ determines m and Q uniquely and vice versa.*

Further, for all $h, g \in U$, we have

(i) $\int_U \langle x, h \rangle \, \mu(\mathrm{d}x) = \langle m, h \rangle$,

(ii) $\int_U \big(\langle x, h \rangle - \langle m, h \rangle\big)\big(\langle x, g \rangle - \langle m, g \rangle\big) \, \mu(\mathrm{d}x) = \langle Qg, h \rangle$,

(iii) $\int_U \|x - m\|^2 \, \mu(\mathrm{d}x) = \operatorname{tr} Q$.

Proof. Let us first show "$\Leftarrow$". Assume that μ satisfies (3.2.1) and let $v \in U$. We have to show that the image $\mu^{(v)}$ of μ under the map $u \mapsto \langle u, v \rangle$ is Gaussian (on $\mathbb{R}$). For $t \in \mathbb{R}$ we have

$$\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}t\lambda} \, \mathrm{d}\mu^{(v)}(\lambda) = \int_U \mathrm{e}^{\mathrm{i}t\langle v,u \rangle} \, \mathrm{d}\mu(u) = \int_U \mathrm{e}^{\mathrm{i}\langle tv,u \rangle} \, \mathrm{d}\mu(u) = \mathrm{e}^{\mathrm{i}t\langle m,v \rangle - \frac{1}{2}t^2\langle Qv,v \rangle},$$

so $\mu^{(v)} = \mathcal{N}(\langle m, v \rangle, \langle Qv, v \rangle)$.

The proof of the converse statement "$\Rightarrow$" is more complicated (see [LR15, p10ff], or [PR07, p6ff], or [DZ92, p53ff]).

Property (i) is an immediate consequence of the fact that $\mu^{(h)}$ has mean $\langle m, h \rangle$ and (ii) with $g = h$ is an immediate consequence of the fact that $\mu^{(g)}$ has variance $\langle Qg, g \rangle$. For the general case, we use the equality $ab = \frac{1}{4}\big((a + b)^2 - (a - b)^2\big)$ for $a, b \in \mathbb{R}$ and obtain that the left hand side of the equality (ii) equals

$$\frac{1}{4}\big(\langle Q(g + h), g + h \rangle - \big(\langle Q(g - h), g - h \rangle\big) = \langle Qg, h \rangle$$

using the fact that Q is symmetric.

Let us show (iii): let (e_i) be an ONB of U and observe that for any $x \in U$, we have (using Parseval's identity):

$$\|x-m\|^2 = \|\sum_i \langle x-m, e_i\rangle e_i\|^2 = \sum_i |\langle x-m, e_i\rangle|^2 = \sum_i |\langle x, e_i\rangle - \langle m, e_i\rangle|^2.$$

Then, by (ii),

$$\operatorname{tr} Q = \sum_i \langle Qe_i, e_i\rangle = \sum_i \int_U \big(\langle x, e_i\rangle - \langle m, e_i\rangle\big)^2 \,\mathrm{d}\mu(x) = \int_U \|x-m\|^2 \,\mathrm{d}\mu(x),$$

so property (iii) follows.

It remains to prove the uniqueness claims.

First assume that $\mu = \mathcal{N}(m,Q) = \mathcal{N}(\bar{m},\bar{Q})$. Then, for $v \in U$, $\mu^{(v)} = \mathcal{N}(\langle m, v\rangle, \langle Qv, v\rangle) = \mathcal{N}(\langle \bar{m}, v\rangle, \langle \bar{Q}v, v\rangle)$, so $\langle m, v\rangle = \langle \bar{m}, v\rangle$ and $\langle Qv, v\rangle = \langle \bar{Q}v, v\rangle$ for all $v \in U$, so $m = \bar{m}$. To see that this also implies $Q = \bar{Q}$ note that, by (ii), we see that $\langle Qu, v\rangle$ only depends on m and μ but not explicitly on Q, and therefore equals $\langle \bar{Q}u, v\rangle$ for all $u, v \in U$. Hence, $\langle (Q-\bar{Q})u, v\rangle = 0$ for all $u, v \in U$. Letting $v := (Q-\bar{Q})u$, we get $\|(Q-\bar{Q})u\|^2 = 0$ for all $u \in U$, so $Q = \bar{Q}$.

Finally, assume that $\mu = \mathcal{N}(m,Q)$ and $\bar{\mu} = \mathcal{N}(m,Q)$. Then $\mu^{(v)} = \bar{\mu}^{(v)}$ since they have the same characteristic function. This implies $\mu = \bar{\mu}$, but why??? We will provide the answer in class. □

At this point it is not yet clear if for any $m \in U$ and trace class $Q \in L(U)$ there *exists* a Gaussian measure with mean m and covariance Q. The positive answer to this question will be given in the next section.

Definition 3.13. A U-valued random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ is called *Gaussian* (with mean m and covariance Q) if its law is a Gaussian measure (with mean m and covariance Q).

Remark 3.14. Theorem 3.12 shows that if X is Gaussian with mean m and covariance Q, then

(i) $\mathbb{E}\langle X, h\rangle = \langle m, h\rangle$,

(ii) $\operatorname{cov}\big(\langle X, h\rangle, \langle X, g\rangle\big) = \langle Qg, h\rangle$,

(iii) $\mathbb{E}\|X-m\|^2 = \operatorname{tr} Q$.

3.3 Representation of Gaussian random variables

The following theorem shows in particular, that for every $m \in U$ and trace class $Q \in L(U)$ there exists a Gaussian random variable (and hence a Gaussian measure) with mean m and covariance Q.

Theorem 3.15. *Let $m \in U$ and $Q \in L(U)$ trace class. Let (e_k) be an ONB of eigenvectors of Q with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$. A U-valued random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ has law $\mathcal{N}(m,Q)$ iff $X = m + \sum_{k\in\mathcal{K}} \sqrt{\lambda_k}\beta_k e_k$ a.s., where $\mathcal{K} = \{k : \lambda_k > 0\}$ and β_k, $k \in \mathcal{K}$ are independent $\mathcal{N}(0,1)$ variables. The series converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$.*

Proof. "$\Rightarrow$": Assume that X has law $\mathcal{N}(m,Q)$. Then, for each $\omega \in \Omega$, we have $X(\omega) = \sum_k \langle X(\omega), e_k\rangle e_k$. Define

$$\beta_k(\omega) := \begin{cases} \frac{\langle X(\omega),e_k\rangle - \langle m,e_k\rangle}{\sqrt{\lambda_k}} & k \in \mathcal{K} \\ 0 & k \notin \mathcal{K}. \end{cases}$$

Then

$$X(\omega) = m + \sum_k \langle X(\omega), e_k\rangle e_k - \sum_k \langle m, e_k\rangle e_k = m + \sum_k \sqrt{\lambda_k}\beta_k(\omega)e_k.$$

Clearly, β_k has law $\mathcal{N}(0,1)$ for each $k \in \mathcal{K}$. In order to see that they are independent, it suffices to show that they are jointly Gaussian and uncorrelated. The first claim is easy to show (see Remark 3.11): just take some finite linear combination of the (β_k) and show that it has a normal law. Then compute the covariance $\mathbb{E}\beta_i\beta_j$ and show that it equals δ_{ij} in case $i,j \in \mathcal{K}$. We will do this in class.

It remains to show that the series converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$. There is nothing to prove when $\mathcal{K}$ is a finite set, so we can and will assume that $\mathcal{K} = \mathbb{N}$. Since the space $L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$ is complete it suffices to show that the series is Cauchy. Indeed, for $n \geq m$ we have

$$\mathbb{E}\Big(\Big\|\sum_{k=m}^n \sqrt{\lambda_k}\beta_k e_k\Big\|^2\Big) = \sum_{k=m}^n \lambda_k \mathbb{E}\big(|\beta_k|^2\big) = \sum_{k=m}^n \lambda_k.$$

Since the sum of the λ_k equals the trace of Q which is finite, the series is Cauchy.

"$\Leftarrow$": Define $X(\omega) := m + \sum_{k\in\mathcal{K}} \sqrt{\lambda_k}\beta_k(\omega)e_k$. We showed above that the series converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$. We check that X has law $\mathcal{N}(m,Q)$. If $\mathcal{K}$ is a finite set, then for $u \in U$, we have

$$\langle X(\omega), u\rangle = \langle m, u\rangle + \sum_{k\in\mathcal{K}} \sqrt{\lambda_k}\beta_k(\omega)\langle e_k, u\rangle,$$

which is normally distributed with mean $\langle m, u\rangle$ and variance $\sum_{k\in\mathcal{K}} \lambda_k \langle u, e_k\rangle^2$. If $\mathcal{K}$ is an infinite set, then we can and will again assume that $\mathcal{K} = \mathbb{N}$. Let $X_n := m + \sum_{k=1}^n \sqrt{\lambda_k}\beta_k e_k$. Then $\langle X_n, u\rangle$ is normally distributed with mean $\langle m, u\rangle$ and variance $\sum_{k=1}^n \lambda_k\langle u, e_k\rangle^2$ and

$$\mathbb{E}\big(\big|\langle X_n, u\rangle - \langle X, u\rangle\big|^2\big) \leq \|u\|^2 \mathbb{E}\big(\|X_n - X\|^2\big)$$

which converges to 0 as $n \to \infty$. Since L^2 limits of Gaussian random variables are Gaussian it follows that $\langle X, u\rangle$ is normally distributed with mean $\langle m, u\rangle$ and variance $\sum_{k\in\mathcal{K}} \lambda_k\langle u, e_k\rangle^2$ no matter if $\mathcal{K}$ is finite or not. Now

$$\langle Qu, u\rangle = \langle Q\sum_j \langle u, e_j\rangle e_j, \sum_k \langle u, e_k\rangle e_k\rangle = \sum_j \langle u, e_j\rangle^2 \lambda_j,$$

so the law of X equals $\mathcal{N}(m,Q)$ by Theorem 3.12. □

3.4 Infinite dimensional Wiener processes

Definition 3.16. Let $Q \in L(U)$ be trace class and let $T > 0$. A U-valued stochastic process $W(t)$, $t \in [0,T]$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *Q-Wiener process* if:

- $W(0,\omega) = 0$ for all $\omega \in \Omega$,

- W has a.s. continuous trajectories,
- W has independent increments,
- for every $0 \le s \le t \le T$ the law of $W(t) - W(s)$ is $\mathcal{N}(0, (t-s)Q)$.

Theorem 3.17. *Let (e_k) be an ONB of U consisting of eigenvectors of the trace class operator Q with corresponding eigenvalues λ_k. Then a U-valued stochastic process $W(t)$, $t \in [0,T]$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is a Q-Wiener process iff*

$$W(t) = \sum_k \sqrt{\lambda_k}\beta_k(t)e_k,\ t \in [0,T],$$

where β_k, $k \in \mathcal{K} := \{k : \lambda_k > 0\}$ are independent real-valued standard Wiener processes on $(\Omega, \mathcal{F}, \mathbb{P})$. The series converges in $L^2\big(\Omega, \mathcal{F}, \mathbb{P}; C([0,T], U)\big)$.

Proof. "$\Rightarrow$": Let W be a Q-Wiener process. For fixed $t \in [0,T]$, $W(t)$ has law $\mathcal{N}(0, tQ)$ and Theorem 3.15 shows that $W(t) = \sum_k \sqrt{\lambda_k}\beta_k(t)e_k$, where

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

Clearly, $\beta_k(t)$ has law $\mathcal{N}(0,t)$ for each $t \in [0,T]$ and $k \in \mathcal{K} := \{k : \lambda_k > 0\}$. We need to check that β_k is a 1-dimensional Wiener process for each $k \in \mathcal{K}$ and that the β_k are independent. This is not hard and we will do this in class.

"$\Leftarrow$": Let $W(t) := \sum_k \sqrt{\lambda_k}\beta_k(t)e_k$, $t \in [0,T]$. The series converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$ for each fixed $t \in [0,T]$. Further, $W(0) = 0$ almost surely, and, by Theorem 3.15, $W(t)$ has law $\mathcal{N}(0, tQ)$ and, moreover, $W(t) - W(s)$ has law $\mathcal{N}(0, (t-s)Q)$ whenever $0 \le s \le t \le T$. The fact that W has independent increments is easy to see (we will do that in class). It remains to show that the series converges in $L^2\big(\Omega, \mathcal{F}, \mathbb{P}); C([0,T], U)\big)$ (and hence W has almost surely continuous trajectories). This is clear in case $\mathcal{K}$ is finite, so we assume that $\mathcal{K} = \mathbb{N}$. Set

$$W^N(t,\omega) := \sum_{k=1}^N \sqrt{\lambda_k}\beta_k(t)e_k,\ t \in [0,T].$$

Then $t \mapsto W^N(t,\omega)$ is almost surely continuous and for $N \ge M$ we have

$$\mathbb{E} \sup_{t\in[0,T]} \big(\|W^N(t) - W^M(t)\|^2\big) = \mathbb{E}\Big(\sup_{t\in[0,T]} \sum_{k=M+1}^N \lambda_k\beta_k^2(t)\Big) \le \sum_{k=M+1}^N \lambda_k\mathbb{E}\Big(\sup_{t\in[0,T]} \beta_k^2(t)\Big) = c\sum_{k=M+1}^N \lambda_k,$$

where $c := \mathbb{E}\Big(\sup_{t\in[0,T]} \beta_1^2(t)\Big)$, so the sequence is Cauchy. The claim follows since $L^2\big(\Omega, \mathcal{F}, \mathbb{P}; C([0,T], U)\big)$ is complete. $\square$

Let us recall the following definition.

Definition 3.18. Fix $T > 0$. A filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *normal* (or *satisfies the usual conditions*) if

- $\mathcal{F}_0$ contains all $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$ and
- $\mathcal{F}_t = \mathcal{F}_t^+ := \bigcap_{s>t} \mathcal{F}_s$ for all $t \in [0, T)$.

Definition 3.19. A Q-Wiener process $W(t)$, $t \in [0, T]$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called a Q-Wiener process with respect to $\mathbb{F}$ if

- W is $\mathbb{F}$-adapted and
- $W(t) - W(s)$ is independent of $\mathcal{F}_s$ for all $0 \le s \le t \le T$.

Remark 3.20. Let W be a Q-Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$. Then it is a Q-Wiener process with respect to some normal filtration $\mathbb{F}$. One can choose the smallest normal filtration $\mathbb{F}$ containing the filtration which is generated by W. It is then clear by construction that W is $\mathbb{F}$-adapted and it is not hard to see that the independence property in the previous definition also holds (see e.g. [PR07, Proposition 2.1.13.]).

Chapter 4

Stochastic integrals

4.1 Conditional expectation and martingales in Banach spaces

Most of this section is taken from Section 2.2 of [LR15]. In this section we assume that $(E, \|.\|)$ is a real separable Banach space. We start with a proposition stating (in particular) that the Borel-σ-algebra (with respect to the norm topology) $\mathcal{B}(E)$ coincides with the Borel-σ-algebra of the *weak topology* on E.

Proposition 4.1. *(i) There exists a countable set $E_0^* \subset E^*$ such that $\|\varphi\|_{E^*} = 1$ for all $\varphi \in E_0^*$ and $\|x\| = \sup_{\varphi \in E_0^*} |\varphi(x)| = \sup_{\varphi \in E_0^*} \varphi(x)$ for all $x \in E$.*

(ii) For every set E_0^* as in (i) we have $\mathcal{B}(E) = \sigma(E^*) = \sigma(E_0^*)$.

Proof. Let $E_0 = \{z_1, z_2, ...\}$ be countable and dense in E. For each n there exists some $\varphi_n \in E^*$ such that $\|\varphi_n\|_{E^*} = 1$ and $\varphi_n(z_n) = \|z_n\|$ (by Hahn-Banach). Define $E_0^* := \{\varphi_1, \varphi_2, ...\}$. Fix $x \in E$ and let $z_{n_1}, z_{n_2}, ...$ be a subsequence of $z_1, z_2, ...$ such that $\|z_{n_k} - x\| \to 0$. Then

$$\|z_{n_k}\| - \varphi_{n_k}(x) = \varphi_{n_k}(z_{n_k} - x) \le \|\varphi_{n_k}\|_{E^*} \|z_{n_k} - x\| = \|z_{n_k} - x\| \to 0,$$

so $\liminf_{k \to \infty} \varphi_{n_k}(x) \ge \|x\|$. Since $\|\varphi_{n_k}\|_{E^*} = 1$ (i) follows.

Let us show (ii). Obviously, $\sigma(E_0^*) \subseteq \sigma(E^*) \subseteq \mathcal{B}(E)$, since every continuous map is measurable. By (i) we have, for each $x \in E$,

$$\|x\| = \sup_{\varphi \in E_0^*} |\varphi(x)|.$$

Therefore, for each $a \in E$ and $r > 0$, we have

$$B(a, r) := \{y : \|y - a\| \le r\} = \bigcap_{\varphi \in E_0^*} \{y : |\varphi(y - a)| \le r\},$$

so $B(a, r) \in \sigma(E_0^*)$. Since the family of closed balls generates $\mathcal{B}(E)$ the proof is complete. $\square$

We continue with the definition of a conditional expectation.

Theorem 4.2. *Let X be a Bochner-integrable E-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then there exists a unique (up to sets of $\mathbb{P}$-measure 0) Bochner-integrable E-valued random variable Z on $(\Omega, \mathcal{F}, \mathbb{P})$ which is $\mathcal{G}$-measurable such that*

$$\int_A X \,\mathrm{d}\mathbb{P} = \int_A Z \,\mathrm{d}\mathbb{P}, \text{ for all } A \in \mathcal{G}.$$

Z is called conditional expectation of X. *We write $\mathbb{E}(X|\mathcal{G})$ instead of Z.*

Further, we have $\|\mathbb{E}(X|\mathcal{G})\| \leq \mathbb{E}(\|X\||\mathcal{G})$.

Proof. We begin with the existence proof of Z in case $X \in \mathcal{E}$. Then X has a representation $X = \sum_{k=1}^n x_k 1_{A_k}$ with disjoint sets $A_1, ..., A_n \in \mathcal{F}$. Define

$$Z := \sum_{k=1}^n x_k \mathbb{E}(1_{A_k}|\mathcal{G}).$$

Then Z is $\mathcal{G}$-measurable and for $A \in \mathcal{G}$ we have

$$\int_A X \mathrm{d}\mathbb{P} = \sum_{k=1}^n x_k \mathbb{P}(A_k \cap A) \text{ and } \int_A Z \mathrm{d}\mathbb{P} = \sum_{k=1}^n x_k \int_A \mathbb{E}(1_{A_k}|\mathcal{G}) \,\mathrm{d}\mathbb{P} = \sum_{k=1}^n x_k \mathbb{P}(A_k \cap A),$$

so $Z = \mathbb{E}(X|\mathcal{G})$. Further,

$$\|Z\| \leq \sum_{k=1}^n \|x_k\| \mathbb{E}(1_{A_k}|\mathcal{G}) = \mathbb{E}\Big(\sum_{k=1}^n \|x_k\| 1_{A_k} \Big| \mathcal{G} \Big) = \mathbb{E}\big(\|X\| \big| \mathcal{G}\big).$$

For the general case and the uniqueness proof the reader is referred to [PR07, Proposition 2.2.1]. The uniqueness proof is slightly different from the usual one due to the lack of an order relation on E. The existence proof uses Proposition 2.4. □

Proposition 4.3. *If X is E-valued Bochner-integrable on $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$ and $Y : \Omega \to E$ is some map, then $Y = \mathbb{E}(X|\mathcal{G})$ iff $l(Y) = \mathbb{E}(l(X)|\mathcal{G})$ for every $l \in E^*$.*

Proof. We first show "$\Rightarrow$". Assume that $Y = \mathbb{E}(X|\mathcal{G})$ and $l \in E^*$. Then $l(Y)$ is $\mathcal{G}$-measurable since it is the composition of measurable maps. Further, $l(Y)$ is integrable since $|l(Y)| \leq \|l\|_{E^*} \|Y\|$ and Y is Bochner-integrable. Fix $A \in \mathcal{G}$. We show that $\int_A l(Y) \,\mathrm{d}\mathbb{P} = \int_A l(X) \,\mathrm{d}\mathbb{P}$. Since this is clear when $\mathbb{P}(A) = 0$ we now assume that $\mathbb{P}(A) > 0$ and define $Q(B) := \mathbb{P}(B|A)$, $B \in \mathcal{F}$. Since $Y = \mathbb{E}(X|\mathcal{G})$, we have $\int Y \,\mathrm{d}Q = \frac{1}{\mathbb{P}(A)} \int_A Y \,\mathrm{d}P = \frac{1}{\mathbb{P}(A)} \int_A X \,\mathrm{d}P = \int X \,\mathrm{d}Q$. Proposition 2.7 implies

$$\int_A l(Y) \,\mathrm{d}\mathbb{P} = \mathbb{P}(A) \int l(Y) \,\mathrm{d}Q = \mathbb{P}(A) l\Big(\int Y \,\mathrm{d}Q \Big) = l\Big(\mathbb{P}(A) \int X \,\mathrm{d}Q \Big) = \int_A l(X) \,\mathrm{d}\mathbb{P},$$

so $l(Y) = \mathbb{E}(l(X)|\mathcal{G})$.

Now we show "$\Leftarrow$". Let $\bar{Y} := \mathbb{E}(X|\mathcal{G})$ ($\bar{Y}$ is well-defined and Bochner-integrable since X is Bochner-integrable). The previous step shows that for each $l \in E^*$ we have

$$l(\bar{Y}) = \mathbb{E}(l(X)|\mathcal{G}) = l(Y), \text{ a.s..}$$

Define E_0^* as in Proposition 4.1. Since E_0^* is countable, there exists a set $N \in \mathcal{F}$ of measure 0 such that $l(\bar{Y}(\omega)) = l(Y(\omega))$ for all $l \in E_0^*$ and all $\omega \notin N$. Hence Proposition 4.1 (i) implies $Y(\omega) = \bar{Y}(\omega)$ for all $\omega \notin N$. Further, arguing like in the remark below, we see that Y is $\mathcal{G}$-measurable and the proof is complete. □

Remark 4.4. If $X : \Omega \to E$ is some map, then the reader may wonder whether or not the assumption that $l(X)$ is integrable for every $l \in E^*$ already implies that X is Bochner-integrable. Well, measurability of X is easy since $\mathcal{B}(E)$ is generated by the family $\{l^{-1}(A);\, l \in E^*,\, A \in \mathcal{B}(\mathbb{R})\}$, but $\|X\|$ is not necessarily integrable. As an example take the separable Hilbert space $E = l^2$ of square summable sequences of real numbers (with the standard inner product). Let X take the value $(0, ..., 0, 2^i/i, 0, ...)$ with probability 2^{-i}, $i = 1, 2, ...$ Clearly, $\|X\|$ is not integrable. Now consider any $l \in E^*$. According to the Riesz representation theorem, we can identify l with an element in U. I claim that $\mathbb{E}|l(X)| < \infty$. We have

$$\mathbb{E}\big|l(X)\big| = \mathbb{E}\big|\langle l, X\rangle\big| \le \mathbb{E}\sum_{i=1}^{\infty}\big(|l_i||X_i|\big) = \sum_{i=1}^{\infty}|l_i|2^{-i}\frac{2^i}{i} \le \|l\|a,$$

where a denotes the norm of the vector $(1, 1/2, 1/3, ...)$ (which is finite).

In the following, we will always assume that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a normal FPS.

Definition 4.5. An E-valued process $M(t)$, $t \ge 0$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is an $\mathbb{F}$-*martingale* if

- M is $\mathbb{F}$-adapted,
- $\mathbb{E}\|M(t)\| < \infty$ for all $t \ge 0$,
- $\mathbb{E}(M(t)|\mathcal{F}_s) = M(s)$ almost surely whenever $0 \le s \le t$.

The following proposition is a simple consequence of Proposition 4.3.

Proposition 4.6. *If M satisfies conditions (i) and (ii) of the previous definition, then M is an $\mathbb{F}$-martingale iff $l(M)$ is an $\mathbb{F}$-martingale for every $l \in E^*$.*

Proposition 4.7. *If M is an $\mathbb{F}$-martingale, $p \in [1, \infty)$ and $\mathbb{E}\|M(t)\|^p < \infty$ for all $t \ge 0$, then $\|M(t)\|^p$, $t \ge 0$ is an $\mathbb{F}$-submartingale.*

Proof. Adaptedness and integrability is clear. First we show the submartingale property for $p = 1$. Define the countable set E_0^* as in the proof of Proposition 4.1. and let $t > s \ge 0$. Then

$$\mathbb{E}\big(\|M(t)\|\big|\mathcal{F}_s\big) \ge \sup_{\varphi \in E_0^*}\mathbb{E}\big(\varphi(M(t))|\mathcal{F}_s\big) = \sup_{\varphi \in E_0^*}\varphi\big(\mathbb{E}(M(t)|\mathcal{F}_s)\big) = \sup_{\varphi \in E_0^*}\varphi(M(s)) = \|M(s)\|.$$

This proves the claim for $p = 1$. The case $p > 1$ follows using Jensen's inequality. □

Theorem 4.8. *(Maximal inequality) If M is a right-continuous $\mathbb{F}$-martingale, then, for $p > 1$ and $T \in (0, \infty)$*

$$\Big(\mathbb{E}\big(\sup_{t \in [0,T]}\|M(t)\|^p\big)\Big)^{1/p} \le \frac{p}{p-1}\sup_{t \in [0,T]}\Big(\mathbb{E}\big(\|M(t)\|^p\big)\Big)^{1/p} = \frac{p}{p-1}\Big(\mathbb{E}\big(\|M(T)\|^p\big)\Big)^{1/p}.$$

Proof. This is an easy consequence of the previous proposition and Doob's maximal inequality for real-valued non-negative submartingales. □

Now we fix $T > 0$ and denote by $\mathcal{M}_T^2(E)$ the space of all E-valued continuous square integrable martingales.

Proposition 4.9. *The space $\mathcal{M}_T^2(E)$ equipped with the norm*

$$\|M\|_{\mathcal{M}_T^2} := \sup_{t\in[0,T]} \Big(\mathbb{E}\big(\|M(t)\|^2\big)\Big)^{1/2} = \Big(\mathbb{E}\big(\|M(T)\|^2\big)\Big)^{1/2} \le \Big(\mathbb{E}\big(\sup_{t\in[0,T]} \|M(t)\|^2\big)\Big)^{1/2} \le 2\Big(\mathbb{E}\big(\|M(T)\|^2\big)\Big)^{1/2}$$

is a Banach space.

Proof. [LR15, p26] □

Proposition 4.10. *Let $T > 0$ and let $W(t)$, $t \in [0,T]$ be a U-valued Q-Wiener process with respect to $\mathbb{F}$. Then $W \in \mathcal{M}_T^2(U)$.*

Proof. Clearly, W is adapted and has continuous paths. Further, $\mathbb{E}\|W(t)\|_U^2 = t \cdot \operatorname{tr} Q < \infty$, so W is square integrable. It remains to check the martingale property which is easy (we will do this in class). □

4.2 Definition of the stochastic integral

Fix $T > 0$, real separable Hilbert spaces $(H, \langle .,. \rangle_H)$ and $(U, \langle .,. \rangle_U)$, a trace class operator $Q \in L(U)$, a normal filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and a Q-Wiener process W with respect to $\mathbb{F}$. Define $\Omega_T := [0,T] \times \Omega$ and $\mathbb{P}_T := \lambda_T \otimes \mathbb{P}$ where λ_T denotes the Lebesgue measure on $[0,T]$. We want to define the stochastic integral $\int_0^t \Phi(s)\, \mathrm{d}W(s)$, $t \in [0,T]$ for a reasonably large class of processes Φ with values in $L(U,H)$. The basic steps of the construction are (similarly as in the finite dimensional case)

Step 1: For a certain class $\mathcal{E}$ of "elementary" $L(U,H)$-valued processes we define the stochastic integral $\operatorname{Int} : \mathcal{E} \to \mathcal{M}_T^2(H) =: \mathcal{M}_T^2$ in a natural way.

Step 2: For a certain norm on (equivalence classes of) $\mathcal{E}$ we show that $\operatorname{Int} : \mathcal{E} \to \mathcal{M}_T^2$ is a linear isometry, so we can extend Int uniquely to a linear isometry on the (abstract) completion $\bar{\mathcal{E}}$ of $\mathcal{E}$.

Step 3: Find a more explicit representation of $\bar{\mathcal{E}}$.

Step 4: Extend the map Int further by localization.

Let us go into more detail:

Step 1:

Definition 4.11. An $L := L(U,H)$-valued process $\Phi(t)$, $t \in [0,T]$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called *elementary* if there exist (deterministic) $0 = t_0 < ... < t_k \le T$, $k \in \mathbb{N}$ such that

$$\Phi(t) = \sum_{m=0}^{k-1} \Phi_m \, 1_{(t_m, t_{m+1}]}(t),\ t \in [0,T], \tag{4.2.1}$$

where, for each $m \in \{0, ..., k-1\}$, Φ_m is an $\mathcal{F}_{t_m}$-measurable map from Ω to L and $\Phi_m(\Omega)$ is a finite subset of L. We denote the set of all elementary processes by $\mathcal{E}$.

For $\Phi \in \mathcal{E}$ with representation (4.2.1) we define

$$\mathrm{Int}(\Phi)(t) := \int_0^t \Phi(s)\,\mathrm{d}W(s) := \sum_{m=0}^{k-1} \Phi_m\big(W(t_{m+1}\wedge t) - W(t_m \wedge t)\big),\ t \in [0,T].$$

As usual, one has to ensure that this definition does not depend on the representation of Φ. It remains to verify that $\mathrm{Int}(\Phi) \in \mathcal{M}_T^2$.

Continuity of the map $t \mapsto \mathrm{Int}(\Phi)(t)$ is clear from the definition.

To see that $\mathrm{Int}(\Phi)(t)$ is square integrable for a given $t \in [0,T]$ note that

$$\|\Phi_m\big(W(t_{m+1}\wedge t) - W(t_m\wedge t)\big)\|_H \le \|\Phi_m\|_{L(U,H)}\|W(t_{m+1}\wedge t) - W(t_m\wedge t)\|_U.$$

Since Φ_m takes only finitely many values and $W(t_{m+1}\wedge t) - W(t_m\wedge t)$ is Gaussian (hence square integrable), our claim follows.

We skip the proof of the martingale property (see e.g. [LR15, p29]).

Step 2: Recall the well-known fact that for a positive semi-definite $Q \in L(U)$ there exists a unique positive semi-definite $R \in L(U)$ such that $R \circ R = Q$. We call R the square root of Q and denote it by $Q^{1/2}$. If Q is trace class, then, obviously, $Q^{1/2} \in L_2(U)$ (see Section 3.1 for the definition) and $\|Q^{1/2}\|^2_{L_2(U)} = \mathrm{tr}\,(Q)$. Let $L_2 := L_2(U,H)$. Recall that $S \in L(U,H)$ and $\tilde S \in L_2(U)$ imply $S \circ \tilde S \in L_2(U,H)$. We can now state the Itô isometry property of Int on $\mathcal{E}$.

Theorem 4.12. *If $\Phi = \sum_{m=0}^{k-1} \Phi_m\, 1_{(t_m,t_{m+1}]} \in \mathcal{E}$, then*

$$\big\|\mathrm{Int}(\Phi)\big\|^2_{\mathcal{M}_T^2} = \Big\|\int_0^{\cdot} \Phi(s)\,\mathrm{d}W(s)\Big\|^2_{\mathcal{M}_T^2} = \mathbb{E}\Big(\int_0^T \|\Phi(s)\circ Q^{1/2}\|^2_{L_2}\,\mathrm{d}s\Big) =: \|\Phi\|_T^2.$$

Proof.

$$\big\|\mathrm{Int}(\Phi)\big\|^2_{\mathcal{M}_T^2} = \mathbb{E}\Big(\Big\|\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big\|^2_H\Big) = \mathbb{E}\Big(\Big\|\sum_{m=0}^{k-1}\Phi_m\big(W(t_{m+1}) - W(t_m)\big)\Big\|^2_H\Big) = \sum_{m=0}^{k-1}\mathbb{E}\big\|\Phi_m\Delta_m\big\|^2_H,$$

where $\Delta_m := W(t_{m+1}) - W(t_m)$. It remains to show that $\mathbb{E}\big\|\Phi_m\Delta_m\big\|^2_H = (t_{m+1}-t_m)\mathbb{E}\big\|\Phi_m \circ Q^{1/2}\big\|^2_{L_2}$. The proof is not difficult but a bit lengthy (see [LR15, p31]) and we skip it. □

Remark 4.13. Note that, strictly speaking, both $\|.\|_{\mathcal{M}_T^2}$ and $\|.\|_T$ are semi-norms and become norms only after defining corresponding equivalence classes (which we will assume without changing notation). Note that if two elementary processes Φ and $\tilde\Phi$ belong to the same equivalence class then it does *not* follow that they agree almost surely on $[0,T]$! If, for example, $Q \equiv 0$, then *all* elementary processes are equivalent.

Before entering into Step 3, we introduce the pseudo inverse of a linear operator $S \in L(U,H)$ (for more details and some proofs see [LR15, Appendix C]).

Definition 4.14. Let $S \in L(U,H)$ and define its null space $\mathrm{Ker}(S) := \{x \in U : Sx = 0\}$ as usual. The pseudo inverse of S is defined as

$$S^{-1} := \big(S|_{\mathrm{Ker}(S)^\perp}\big)^{-1} : S\big(\mathrm{Ker}(S)^\perp\big) = S(U) \to \mathrm{Ker}(S)^\perp.$$

Note that the map S^{-1} is linear and one-to-one.

Proposition 4.15. *Let $S \in L(U)$ and define $\langle x, y\rangle_{S(U)} := \langle S^{-1}x, S^{-1}y\rangle_U$. Then $\big(S(U), \langle .,.\rangle_{S(U)}\big)$ is a separable Hilbert space.*

Step 3: After this short excursion we return to our set-up. It turns out that the space $L(U,H)$ is not very convenient to work with as a space where the integrand Φ takes values in. The space is neither Hilbert nor is it separable when U and H are infinite dimensional. On the other hand, in many cases, $W(t)$ takes values in a proper subspace U_0 of U and hence there is no point in defining the integrand Φ outside that subspace. Define

$$U_0 := Q^{1/2}(U),\ \langle u_0, v_0\rangle_0 := \langle Q^{-1/2}u_0, Q^{-1/2}v_0\rangle_U,\ u_0, v_0 \in U_0.$$

Then $(U_0, \langle .,.\rangle_0)$ is a separable Hilbert space by Proposition 4.15.

Abbreviate $L_2^0 := L_2(U_0, H)$ and let (g_k) be an ONB of $\big(\mathrm{Ker}\big(Q^{1/2}\big)\big)^{\perp}$. Then $\big(Q^{1/2}g_k\big)$ is an ONB of $\big(U_0, \langle .,.\rangle_0\big)$. We extend (g_k) to an ONB of U. Note that $\|S\|_{L_2^0} = \|S \circ Q^{1/2}\|_{L_2(U,H)} = \sum_k \langle SQ^{1/2}g_k, SQ^{1/2}g_k\rangle_H$ for each $S \in L_2^0$. Let

$$L(U,H)_0 := \{S|_{U_0} : S \in L(U,H)\}.$$

Then $L(U,H)_0 \subset L_2^0$ since $S \in L(U,H)$ implies

$$\|S|_{U_0}\|^2_{L_2^0} = \sum_k \|S \circ Q^{1/2}g_k\|_H^2 \le \|S\|^2_{L(U,H)} \operatorname{tr} Q < \infty.$$

Further, for $\Phi \in \mathcal{E}$, we have

$$\|\Phi\|_T = \Big(\mathbb{E}\Big(\int_0^T \|\Phi(s) \circ Q^{1/2}\|^2_{L_2}\, \mathrm{d}s\Big)\Big)^{1/2} = \Big(\mathbb{E}\int_0^T \|\Phi(s)\|^2_{L_2^0}\, \mathrm{d}s\Big)^{1/2}.$$

Definition 4.16. The *predictable σ-algebra* $\mathcal{P}_T$ is the sub-σ-algebra of $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$ defined as

$$\mathcal{P}_T := \sigma\big(\{(s,t] \times F_s : 0 \le s < t \le T,\ F_s \in \mathcal{F}_s\} \cup \{\{0\} \times F_0 : F_0 \in \mathcal{F}_0\}\big).$$

A process Y defined on Ω_T taking values in a measurable space is called predictable if Y is measurable with respect to $\mathcal{P}_T$.

Proposition 4.17.

$$\mathcal{P}_T = \sigma\big(Y : \Omega_T \to \mathbb{R}\big|\, Y \text{ adapted and left-continuous}\big).$$

Proof. Let $\tilde{\mathcal{P}} := \sigma\big(Y : \Omega_T \to \mathbb{R}\big|\, Y \text{ adapted and left-continuous}\big)$. To see that $\mathcal{P}_T \subseteq \tilde{\mathcal{P}}$, take $(s,t] \times F_s$ as in the previous definition and define $Y(u,\omega) := 1_{(s,t]}(u)\, 1_{F_s}(\omega)$, $u \in [0,T]$. Then Y is adapted and left-continuous, so $(s,t] \times F_s \in \tilde{\mathcal{P}}$. Obviously, we also have $\{0\} \times F_0 \in \tilde{\mathcal{P}}$ if $F_0 \in \mathcal{F}_0$.

The opposite direction is proved by approximation. □

Now we are ready to identify the completion $\bar{\mathcal{E}}$ of $\mathcal{E}$.

Proposition 4.18. *We have*

$$\bar{\mathcal{E}} = \mathcal{N}_W^2(0,T;H) := \big\{\Phi : \Omega_T \to L_2^0\big|\, \Phi \text{ predictable and } \|\Phi\|_T < \infty\big\} = L^2\big(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; L_2^0\big).$$

You can find the proof in [LR15, p33ff]. One has to check that $\mathcal{E} \subset \mathcal{N}_W^2(0,T;H)$, that $\mathcal{N}_W^2(0,T;H)$ is complete and that $\mathcal{E}$ is dense in $\mathcal{N}_W^2(0,T;H)$.

Step 4: Now we extend the definition of the stochastic integral by localization.

Definition 4.19. The set

$$\mathcal{N}_W(0,T;H) := \Big\{\Phi : \Omega_T \to L_2^0 \Big| \, \Phi \text{ predictable with } \mathbb{P}\Big(\int_0^T \|\Phi(s)\|_{L_2^0}^2 \, \mathrm{d}s < \infty\Big) = 1\Big\}$$

is called the space of *stochastically integrable processes* on $[0,T]$.

For $\Phi \in \mathcal{N}_W(0,T;H)$ and $n \in \mathbb{N}$ define $\tau_n := \inf\big\{t \in [0,T] \big| \int_0^t \|\Phi(s)\|_{L_2^0}^2 \, \mathrm{d}s > n\big\} \wedge T$. Then, for each $n \in \mathbb{N}$, τ_n is an $\mathbb{F}$-stopping time, $1_{(0,\tau_n]}\Phi$ is L_2^0-predictable, and $\mathbb{E}\int_0^T \|1_{(0,\tau_n]}(s)\,\Phi(s)\|_{L_2^0}^2 \, \mathrm{d}s \leq n < \infty$, so $t \mapsto \int_0^t 1_{(0,\tau_n]}(s)\,\Phi(s)\,\mathrm{d}W(s)$ is well-defined and in $\mathcal{M}_T^2$. One has to ensure that letting $n \to \infty$, we obtain a limit and that this limit does not depend on the particular choice of the stopping times. The limit is a continuous local martingale (but generally not a martingale). For details, see [LR15].

4.3 Cylindrical Wiener processes

As before, we assume that $(U, \langle .,. \rangle_U)$ is a real separable Hilbert space. In this section we will show how one can define a Q-Wiener process for positive semi-definite $Q \in L(U)$ without assuming that Q is trace class. Define $U_0 := Q^{1/2}(U)$ and $\langle u_0, v_0 \rangle_0$ as before, where $Q^{1/2}$ denotes the unique positive semi-definite $R \in L(U)$ such that $R^2 = Q$ (this fact does not depend on the trace class property of Q).

Since $U_0 \subseteq U$ we can define the inclusion map $J : U_0 \to U$. Is $J \in L_2(U_0, U)$? Take an ONB (g_k) of $(\operatorname{Ker} Q^{1/2})^\perp$ and define $e_k := Q^{1/2} g_k$ as before. Then (e_k) is an ONB of U_0 and

$$\sum_k \langle Je_k, Je_k \rangle_U = \sum_k \langle JQ^{1/2}g_k, JQ^{1/2}g_k \rangle_U = \sum_k \langle Q^{1/2}g_k, Q^{1/2}g_k \rangle_U = \sum_k \langle Qg_k, g_k \rangle_U$$

is finite iff Q is trace class. We now replace U by a different space U_1 and a different map (which we again call J) so that J becomes Hilbert-Schmidt.

Proposition 4.20. *There exists a separable Hilbert space $(U_1, \langle .,. \rangle_1)$ and $J : (U_0, \langle .,. \rangle_0) \to (U_1, \langle .,. \rangle_1)$ which is Hilbert-Schmidt and one-to-one.*

Proof. Define $U_1 := U$ and let $\alpha_k > 0$ be real numbers which are square summable. For (e_k) and (g_k) as above let

$$J(u) := \sum_k \alpha_k \langle u, e_k \rangle_0 e_k, \; u \in U_0.$$

Then J is one-to-one and

$$\sum_l \langle Je_l, Je_l \rangle_U = \sum_l \alpha_l^2 \langle e_l, e_l \rangle_U = \sum_l \alpha_l^2 \langle Q^{1/2}g_l, Q^{1/2}g_l \rangle_U \leq \|Q^{1/2}\|_{L(U)}^2 \sum_l \alpha_l^2 < \infty,$$

so J is Hilbert-Schmidt. □

We now fix $(U_1, \langle ., . \rangle_1)$ and J as in the previous definition (but not necessarily as in its proof) and define $Q_1 := J \circ J^* \in L(U_1)$. We claim that Q_1 is trace class. Clearly, Q_1 is positive semi-definite and for an ONB (f_k) of U_1 we have

$$\sum_k \langle Q_1 f_k, f_k \rangle_1 = \sum_k \langle J^* f_k, J^* f_k \rangle_0 < \infty$$

due to the fact that J is Hilbert-Schmidt implies that J^* is also Hilbert-Schmidt.

Proposition 4.21. *Let (e_k) be an ONB of $U_0 := Q^{1/2}(U)$ and let (β_k) be independent real-valued Brownian motions. Then the sum*

$$W(t) := \sum_k \beta_k(t) J e_k, \ t \in [0, T]$$

converges in $\mathcal{M}_T^2(U_1)$ and defines a Q_1-Wiener process on U_1. Further, $Q_1^{1/2}(U_1) = J(U_0)$ and for $u_0 \in U_0$ we have

$$\|u_0\|_0 = \|Q_1^{-1/2} J u_0\|_1 = \|J u_0\|_{Q_1^{1/2}(U_1)},$$

so $J : U_0 \to Q_1^{1/2}(U_1)$ is an isometry.

The process W is called a cylindrical Q-Wiener process *in U.*

Proof. [LR15, p50f] □

Next, we define the stochastic integral with respect to a cylindrical Q-Wiener process (where $Q \in L(U)$ is positive semi-definite but not necessarily trace class). Let W be a cylindrical Q-Wiener process with Q_1 and J as above. Define

$$\mathcal{N}_W := \Big\{ \Phi : \Omega_T \to L_2^0 \Big| \, \Phi \text{ predictable and } \mathbb{P}\Big(\int_0^T \|\Phi(s)\|_{L_2^0}^2 \, \mathrm{d}s < \infty \Big) = 1 \Big\}.$$

Note that this definition does not depend on the choice of Q_1 and J. We want to define the stochastic integral of $\Phi \in \mathcal{N}_W$. For $\phi \in L_2^0$ we have

$$\|\phi\|_{L_2^0}^2 = \sum_k \langle \phi e_k, \phi e_k \rangle_H = \sum_k \langle \phi J^{-1}\big(J e_k\big), \phi J^{-1}\big(J e_k\big) \rangle_H = \|\phi \circ J^{-1}\|_{L_2(Q_1^{1/2}(U_1), H)}^2,$$

so, for $\Phi \in \mathcal{N}_W$, we define

$$\int_0^t \Phi(s) \, \mathrm{d}W(s) := \int_0^t \Phi(s) \circ J^{-1} \, \mathrm{d}W(s),$$

where the second integral should be interpreted as stochastic integral with respect to the Q_1-Wiener process W which we defined before (one has to check that $\Phi \circ J^{-1}$ is really stochastically integrable with respect to the Q_1-Wiener process W).

Remark 4.22. The stochastic integral is independent of the choice of $(U_1, \langle ., . \rangle_1)$ and J. To see this consider $\Phi \in \mathcal{E}$ with representation (4.2.1). Then

$$\begin{aligned}\Phi_m \circ J^{-1}\big(W(t_{m+1} \wedge t) - W(t_m \wedge t)\big) &= \Phi_m \circ J^{-1}\Big(\sum_k \beta_k(t_{m+1} \wedge t) J e_k - \sum_k \beta_k(t_m \wedge t) J e_k \Big) \\ &= \Phi_m\Big(\sum_k \beta_k(t_{m+1} \wedge t) e_k - \sum_k \beta_k(t_m \wedge t) e_k \Big)\end{aligned}$$

does not depend on U_1 and J, so for $\Phi \in \mathcal{E}$ the stochastic integral does not depend on U_1 and J and this immediately implies that the same is true for general $\Phi \in \mathcal{N}_W$. Note however that $\mathcal{E}$ is in general *not* contained in L_2^0!

Chapter 5

Stochastic differential equations in infinite dimensional spaces

In this chapter we largely follow Chapter 4 of [LR15] but we will impose more restrictive assumptions in order to keep things reasonably simple. In particular, we will assume that the coefficients of an infinite dimensional stochastic differential equation are neither explicitly time-dependent nor random.

5.1 Gelfand triples and conditions on the coefficients

Let $(H, \langle .,. \rangle_H)$ be a real separable Hilbert space with dual H^* and let $(V, \|.\|_V)$ be a reflexive Banach space (please take a look at basic monographs in functional analysis in case you don't know what *reflexive* means). We assume that $V \subset H$ continuously and densely (meaning that there is a one-to-one linear continuous map L from V to H such that $L(V)$ is dense in H). Then it follows that $H^* \subset V^*$ continuously and densely and identifying H and H^* we get

$$V \subset H \subset V^*.$$

Defining ${}_{V^*}\langle z, v \rangle_V := z(v)$ for $v \in V$, $z \in V^*$, we have

$${}_{V^*}\langle z, v \rangle_V = \langle z, v \rangle_H \text{ for all } z \in H,\ v \in V.$$

(V, H, V^*) is called a *Gelfand triple*. The fact that $H^* \subset V^*$ continuously and densely and that H^* is separable implies that V^* is separable (since every dense set in H^* is also a dense set in V^* with respect to the corresponding topologies) which in turn implies that V is separable.

Recall that $\mathcal{B}(V) = \sigma(V^*)$ and $\mathcal{B}(H) = \sigma(H^*)$ by Proposition 4.1. Without proof, we state *Kuratowski's theorem* (see [J78, p420] or [C13, Theorem 8.3.7]).

Theorem 5.1. *Let (E_1, d_1) and (E_2, d_2) be Polish spaces and assume that $f : E_1 \to E_2$ is one-to-one and Borel-measurable. Then $f(\mathcal{B}(E_1)) \subset \mathcal{B}(E_2)$.*

Since V, H, V^* are Polish, Kuratowski's theorem implies immediately the following: $V \in \mathcal{B}(H)$, $H \in \mathcal{B}(V^*)$, $\mathcal{B}(V) = \mathcal{B}(H) \cap V$, and $\mathcal{B}(H) = \mathcal{B}(V^*) \cap H$.

We will study stochastic differential equations of the following type

$$\mathrm{d}X(t) = A(X(t))\,\mathrm{d}t + B(X(t))\,\mathrm{d}W(t),\ X(0) = x \in H, \tag{5.1.1}$$

where

- $A : V \to V^*$ is measurable,
- $B : V \to L_2(U, H)$ is measurable, where U is a real separable Hilbert space,
- W is a cylindrical Q-Wiener process with $Q = \mathrm{id} \in L(U)$,
- X takes values in H.

We impose the following conditions on A and B:

(H1) (*Hemicontinuity*) For all $u, v, x \in V$ the map $\lambda \mapsto {}_{V^*}\langle A(u + \lambda v), x\rangle_V$ is in $C(\mathbb{R}, \mathbb{R})$.

(H2) (*Weak monotonicity*) There exists some $c \in \mathbb{R}$ such that for all $u, v \in V$ we have

$$2\,{}_{V^*}\langle A(u) - A(v), u - v\rangle_V + \|B(u) - B(v)\|^2_{L_2(U,H)} \le c\|u - v\|^2_H.$$

(H3) (*Coercivity*) There exist $\alpha \in (1, \infty)$, $c_1 \in \mathbb{R}$, $c_2 \in (0, \infty)$ and $c_4 \in \mathbb{R}$ such that for all $v \in V$

$$2\,{}_{V^*}\langle A(v), v\rangle_V + \|B(v)\|^2_{L_2(U,H)} \le c_1\|v\|^2_H - c_2\|v\|^\alpha_V + c_4.$$

(H4) (*Boundedness*) There exist $c_3, c_5 \ge 0$ such that for all $v \in V$ we have

$$\|A(v)\|_{V^*} \le c_3\|v\|^{\alpha-1}_V + c_5,$$

where α is as in (H3).

Remark 5.2. Note that, even in the finite dimensional case in which $A : \mathbb{R}^d \to \mathbb{R}^d$, property (H1) alone does not imply continuity of A (try to find a counterexample when $d = 2$!). Interestingly, (H1) and (H2) together do imply continuity even in the general case. More precisely, $u_n \to u$ in the norm topology of V implies $A(u_n) \to A(u)$ weakly in V^*, see [LR15, p71fff].

Remark 5.3. It may be interesting to investigate whether measurability of A and B actually follow from (H1)-(H4).

In order to get some feel for conditions (H1)-(H4), we look at two examples.

Example 5.4. Let us consider the finite dimensional case $H = V = V^* = \mathbb{R}^d$, $U = \mathbb{R}^m$. Then (H1) holds if the *drift* $A : \mathbb{R}^d \to \mathbb{R}^d$ is continuous (by the previous remark continuity of A is even necessary for (H1) and (H2) to hold). (H2) is the usual *one-sided (global) Lipschitz condition* as in WT3. (H3) with $\alpha = 2$ is essentially the one-sided linear growth conditions from WT3 while (H4) imposes a linear growth condition on $A(v)$ which we did not need to impose in the finite dimensional case.

The next example is truly infinite dimensional.

Example 5.5. Let $\Lambda \subseteq \mathbb{R}^d$ be non-empty and open. Does $A = \Delta$ (the Laplace operator on Λ) satisfy (H1)-(H4) for a suitable choice of (V, H, V^*) at least in case $B \equiv 0$? If yes, then this means that the usual (deterministic) heat equation is covered by our approach. We first consider the Laplacian Δ as an operator on $C_0^\infty(\Lambda)$ which is the space of infinitely differentiable real functions on Λ with compact support. A reasonable choice for the spaces are $V = H_0^{1,p}(\Lambda)$

for $p \in [2, \infty)$ and $H = L^2(\Lambda)$ (the space of functions on Λ which are square-integrable with respect to Lebesgue measure). We explain what $H_0^{1,p}(\Lambda)$ is and show that A satisfies (H1)-(H4) for $B \equiv 0$.

For $p \in [1, \infty)$ and $u \in C_0^\infty(\Lambda)$ define

$$\|u\|_{1,p} := \Big(\int_\Lambda |u(x)|^p + |\nabla u(x)|^p \,\mathrm{d}x \Big)^{1/p}$$

and denote the completion of $C_0^\infty(\Lambda) \subset L^p(\Lambda)$ with respect to the norm $\|.\|_{1,p}$ by $H_0^{1,p}(\Lambda)$. This space is called *Sobolev space of order 1 in L^p with Dirichlet boundary conditions*. Observe that the embedding $i : C_0^\infty(\Lambda) \to L^p(\Lambda)$ is continuous since for $u \in C_0^\infty(\Lambda)$, we have $\|u\|_p \le \|u\|_{1,p}$, so $\|i\|_{L(C_0^\infty(\Lambda), L^p(\Lambda))} \le 1$. Let $\bar{i} : H_0^{1,p}(\Lambda) \to L^p(\Lambda)$ be the unique continuous linear extension of i. Note that the map $\bar{i}$ is one-to-one and that $H_0^{1,p}(\Lambda) \subset L^p(\Lambda)$.

Next, we want to extend the gradient operator $\nabla : C_0^\infty(\Lambda) \to L^p(\Lambda; \mathbb{R}^d)$ to $H_0^{1,p}(\Lambda)$. Let $u \in H_0^{1,p}(\Lambda)$ and take a sequence (u_n) in $C_0^\infty(\Lambda)$ which converges to u in $H_0^{1,p}(\Lambda)$. Then, in particular, (∇u_n) is Cauchy in $L^p(\Lambda; \mathbb{R}^d)$. Define $\nabla u := \lim_{n\to\infty} \nabla u_n$ in $L^p(\Lambda; \mathbb{R}^d)$. This limit only depends on u and not on the approximating sequence (u_n). We mention without proof that $H_0^{1,p}(\Lambda)$ is reflexive whenever $p \in (1, \infty)$ ([LR15, p79]).

Now, we assume $p \in [2, \infty)$ and define $V := H_0^{1,p}(\Lambda)$, $H = L^2(\Lambda)$, $V^* := \big(H_0^{1,p}(\Lambda)\big)^*$. If $p > 2$, then we assume, in addition, that $|\Lambda| < \infty$, where $|\Lambda|$ denotes the d-dimensional Lebesgue measure of the set Λ. Then $V \subset L^p(\Lambda) \subset H$ where both inclusions are continuous and dense. We want to extend Δ with domain $C_0^\infty(\Lambda)$ to a bounded linear operator $A : V \to V^*$. First note that A maps $C_0^\infty(\Lambda)$ to itself which is a subspace of $L^2(\Lambda) \subset V^*$, so A maps $C_0^\infty(\Lambda)$ to V^*.

For $u, v \in C_0^\infty(\Lambda)$, we have, using integration by parts and Hölder's inequality,

$$\begin{aligned} \Big| {}_{V^*}\langle \Delta u, v\rangle_V \Big| &= \Big| \langle \Delta u, v\rangle_H \Big| = \Big| \int_\Lambda (\Delta u) v \,\mathrm{d}x \Big| = \Big| - \int_\Lambda \big(\nabla u\big)\big(\nabla v\big) \,\mathrm{d}x \Big| \\ &\le \Big(\int_\Lambda \big|\nabla u(x)\big|^{\frac{p}{p-1}} \,\mathrm{d}x \Big)^{\frac{p-1}{p}} \Big(\int_\Lambda \big|\nabla v(x)\big|^p \,\mathrm{d}x \Big)^{1/p} \le \| |\nabla u| \|_{\frac{p}{p-1}} \|v\|_{1,p}, \end{aligned}$$

so, using Hölder's inequality again, we obtain

$$\|\Delta u\|_{V^*} \le \| |\nabla u| \|_{\frac{p}{p-1}} \le \Big(\int_\Lambda \big|\nabla u(x)\big|^p \,\mathrm{d}x \Big)^{1/p} \Big(\int_\Lambda 1 \,\mathrm{d}x \Big)^{\frac{p-2}{p-1}} \le \|u\|_{1,p} |\Lambda|^{\frac{p-2}{p-1}},$$

so $\|\Delta u\|_{V^*} \le \|u\|_{1,p} |\Lambda|^{\frac{p-2}{p-1}}$. So the map

$$\Delta : \big(C_0^\infty(\Lambda), \|.\|_{1,p}\big) \to V^*$$

is continuous and linear and can therefore be uniquely extended to a continuous linear operator $A : V \to V^*$ (with norm $\|A\|_{L(V,V^*)} \le |\Lambda|^{\frac{p-2}{p-1}}$).

We claim that A satisfies conditions (H1), (H2), (H4) and that A satisfies (H3) if, in addition, $p = 2$.

(H1) By linearity, ${}_{V^*}\langle A(u + \lambda v), x\rangle_V = {}_{V^*}\langle Au, x\rangle_V + \lambda\, {}_{V^*}\langle Av, x\rangle_V$ which is continuous in λ.

(H2) Let $u, v \in V$ and choose sequences (u_n) and (v_n) in $C_0^\infty(\Lambda)$ such that $u_n \to u$ and $v_n \to v$ in V. Then

$$_{V^*}\langle A(u) - A(v), u - v\rangle_V = \lim_{n\to\infty} {}_{V^*}\langle \Delta\big(u_n - v_n\big), u_n - v_n\rangle_V \le 0.$$

(H3) For $p=2$, $\alpha=2$, $v \in V$, and a sequence (v_n) in C_0^∞ such that $\|v_n - v\|_V \to 0$, we have

$$2\,{}_{V^*}\langle A(v), v\rangle_V = 2\lim_{n\to\infty} {}_{V^*}\langle \Delta v_n, v_n\rangle_V = -2\lim_{n\to\infty}\int_\Lambda \left|\nabla v_n(x)\right|^2 \mathrm{d}x$$
$$= -2\int_\Lambda \left|\nabla v(x)\right|^2 \mathrm{d}x = 2\big(\|v\|_H^2 - \|v\|_{1,2}^2\big).$$

(H4) For $v \in V$, $\|A(v)\|_{V^*} \le \|v\|_{1,p}\,|\Lambda|^{\frac{p-2}{p-1}}$, so (H4) holds.

Now assume that $B : V \to L_2(U,H)$ is not necessarily 0 but satisfies $\|B(u)-B(v)\|^2_{L_2(U,H)} \le c\|u-v\|_H^2$ and $\|B(v)\|^2_{L_2(U,H)} \le \tilde{c}\big(\|v\|_H^2+1\big)$ for some $c, \tilde{c} \ge 0$ and all $u, v \in V$ (in fact the second property follows from the first for some suitable $\tilde{c}$). Then, for A as above, (A,B) satisfies (H1), (H2), (H4) for every $p \in [2,\infty)$. If $p=2$, then (A,B) also satisfies (H3). In this case, the results in the next section imply that the sde

$$\mathrm{d}X(t) = \Delta X(t)\,\mathrm{d}t + B(X(t))\,\mathrm{d}W(t)$$

has a unique solution.

5.2 Existence and uniqueness of solutions and an Itô formula

We assume that A and B satisfy assumptions (H1)-(H4). Fix $T>0$ and denote Lebesgue measure on $[0,T]$ by λ_T.

Definition 5.6. A continuous H-valued $\mathbb{F}$-adapted process $(X(t))_{t\in[0,T]}$ is called a *solution of* (5.1.1) if for its $\lambda_T \otimes \mathbb{P}$-equivalence class $\hat{X}$ of V^*-valued processes, we have

$$\hat{X} \in L^\alpha([0,T]\times\Omega, \lambda_T\otimes\mathbb{P}; V) \cap L^2([0,T]\times\Omega, \lambda_T\otimes\mathbb{P}; H)$$

for some $\alpha > 1$ and, $\mathbb{P}$-almost surely,

$$X(t) = X(0) + \int_0^t A(\bar{X}(s))\,\mathrm{d}s + \int_0^t B(\bar{X}(s))\,\mathrm{d}W(s),\ t\in[0,T],$$

where $\bar{X}$ is any V-valued predictable modification of $\hat{X}$.

Remark 5.7. Any continuous H-valued $\mathbb{F}$-adapted process $(X(t))_{t\in[0,T]}$ whose $\lambda_T\otimes\mathbb{P}$-equivalence class $\hat{X}$ is in $L^\alpha([0,T]\times\Omega, \lambda_T\otimes\mathbb{P}; V)$ has a V-valued predictable modification $\bar{X}$. To see this, define $\mathcal{R} := \{(s,\omega)\in[0,T]\times\Omega :\ X(s,\omega)\in V\}$. Then $\mathcal{R}\in\mathcal{P}_T$ since it is the inverse image of the set $V\in\mathcal{B}(H)$ under the (left-)continuous H-valued process X (which is therefore predictable).

Define

$$\bar{X}(s,\omega) := \begin{cases} X(s,\omega) & (s,\omega)\in\mathcal{R}, \\ 0 & \text{else.}\end{cases}$$

Then $\bar{X}$ has the asserted property.

Remark 5.8. In the definition of a solution in [LR15], the word *predictable* is replaced by *progressively measurable*. This is really a matter of taste. Our definition has the advantage that we do not need to introduce the *progressive* σ-algebra at all and that our integrands in stochastic integrals will be predictable as required and we do not have to argue that in fact we could also allow progressive integrands. Another advantage of our definition is that it generalizes to more general integrators (possibly with jumps) in which case one cannot allow general progressive integrands. On the other hand, [LR15] allow the coefficients to be time-dependent and random (which we do not) and in this case requiring them to be progressive is a weaker assumption than predictability.

Before we state and prove the main result on existence and uniqueness of solutions to (5.1.1) we state a kind of Itô's formula without proof (see [LR15, p91fff]; note that in the finite dimensional case this is a special case of the usual Itô's formula).

Theorem 5.9. *Let $\alpha > 1$ and $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Let $Y \in L^{\frac{\alpha}{\alpha-1}}([0,T] \times \Omega, \lambda_T \otimes \mathbb{P}; V^*)$ and $Z \in L^2([0,T] \times \Omega, \lambda_T \otimes \mathbb{P}; L_2(U,H))$ be predictable. Define the continuous V^*-valued process*

$$X(t) := X_0 + \int_0^t Y(s)\,\mathrm{d}s + \int_0^t Z(s)\,\mathrm{d}W(s), \quad t \in [0,T].$$

If the $\lambda_T \otimes \mathbb{P}$ equivalence class $\hat{X}$ of X satisfies $\hat{X} \in L^\alpha([0,T] \times \Omega, \lambda_T \otimes \mathbb{P}; V)$, and if $\mathbb{E}\big(\|X(t)\|_H^2 < \infty$ for λ_T-almost all $t \in [0,T]$, then X is adapted and continuous H-valued and

$$\|X(t)\|_H^2 = \|X_0\|_H^2 + \int_0^t \Big(2_{V^*}\langle Y(s), \bar{X}(s)\rangle_V + \|Z(s)\|^2_{L_2(U,H)}\Big)\,\mathrm{d}s + 2\int_0^t \langle \bar{X}(s), Z(s)\,\mathrm{d}W(s)\rangle_H,\ t \in [0,T],$$

where $\bar{X}$ is any V-valued predictable modification of X.

Theorem 5.10. *Let A and B satisfy (H1)-(H4) and let $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Then there exists a unique solution of (5.1.1) with $X(0) = X_0$ in the sense of the previous definition. Moreover,*

$$\mathbb{E}\Big(\sup_{t \in [0,T]} \|X(t)\|_H^2\Big) < \infty.$$

Proof. We can and will assume that H (and hence V) is infinite dimensional. The basic strategy of the proof is as follows:

a) define finite dimensional "projected solutions" $X^{(n)}$,

b) show that there exists a subsequence $X^{(n_k)}$ which converges to some process $\bar{X}$ in an appropriate sense,

c) show that $\bar{X}$ solves (5.1.1),

d) show uniqueness.

Let us now provide more details about these steps (for full details, see [LR15, Theorem 4.2.4]:

a) Let $(e_n)_{n \in \mathbb{N}}$ be an ONB of H such that $e_i \in V$ for all $i \in \mathbb{N}$ and $\mathrm{span}\{e_i,\ i \in \mathbb{N}\}$ is dense in V. Define $H_n := \mathrm{span}\{e_1, ..., e_n\}$, $n \in \mathbb{N}_0$. Define the projection map $P_n : V^* \to H_n$ by $P_n y := \sum_{i=1}^n {}_{V^*}\langle y, e_i\rangle_V e_i$. For ease of notation we assume that U is infinite dimensional and let

$(g_n)_{n\in\mathbb{N}}$ be an ONB of U (the necessary change in case U is finite dimensional will be obvious). Let

$$W(t)=\sum_{i=1}^{\infty}\beta_i(t)Jg_i$$

be a representation of W with independent standard Wiener processes β_i, $i\in\mathbb{N}$ and define

$$W^{(n)}(t):=\sum_{i=1}^{n}\beta_i(t)g_i.$$

For $n\in\mathbb{N}$ consider the SDE

$$\mathrm{d}X^{(n)}(t)=P_nA(X^{(n)}(t))\,\mathrm{d}t+P_nB(X^{(n)}(t))\,\mathrm{d}W^{(n)}(t),\ X^{(n)}(0)=P_nX_0.$$

We will show in class that the conditions for existence and uniqueness of a strong solution (with values in the finite dimensional space H_n) from WT3 are satisfied.

b) Define

$$K:=L^{\alpha}([0,T]\times\Omega,\lambda_T\otimes\mathbb{P};V).$$

We claim that there exists some $C>0$ such that

$$\|X^{(n)}\|_K+\|A(X^{(n)})\|_{K^*}+\sup_{t\in[0,T]}\mathbb{E}\Big(\|X^{(n)}(t)\|_H^2\Big)\le C \tag{5.2.1}$$

for all $n\in\mathbb{N}$. By the usual Itô formula (from WT3), we have

$$\begin{aligned}\|X^{(n)}(t)\|_H^2&=\|X_0^{(n)}\|_H^2+\int_0^t\Big(2_{V^*}\langle P_nA(X^{(n)}(s)),X^{(n)}(s)\rangle_V+\|P_nB(X^{(n)}(s))\|^2_{L_2(U,H)}\Big)\,\mathrm{d}s+M_n(t)\\&=\|X_0^{(n)}\|_H^2+\int_0^t\Big(2_{V^*}\langle A(X^{(n)}(s)),X^{(n)}(s)\rangle_V+\|P_nB(X^{(n)}(s))\|^2_{L_2(U,H)}\Big)\,\mathrm{d}s+M_n(t)\\&\le\|X_0^{(n)}\|_H^2+\int_0^t\Big(c_1\|X^{(n)}(s)\|_H^2-c_2\|X^{(n)}(s)\|_V^{\alpha}+c_4\Big)\,\mathrm{d}s+M_n(t),\end{aligned}$$

where M_n is a continuous local martingale and where we used (H3) and $\|P_n\|_{L(H)}\le 1$. This implies

$$Z(t):=\|X^{(n)}(t)\|_H^2+c_2\int_0^t\|X^{(n)}(s)\|_V^{\alpha}\,\mathrm{d}s\le c_1\int_0^t Z(s)\,\mathrm{d}s+c_4t+\|X_0^{(n)}\|_H^2+M_n(t).$$

By a straightforward localization argument we see that $\mathbb{E}Z(t)<\infty$ for all t and Gronwall's inequality applied to $\mathbb{E}Z(t)$ implies

$$\mathbb{E}Z(t)\le \mathrm{e}^{c_1t}\Big(c_4t+\mathbb{E}\|X_0^{(n)}\|_H^2\Big)\le\mathrm{e}^{c_1t}\Big(c_4t+\mathbb{E}\|X_0\|_H^2\Big),$$

so boundedness of the first and the third term in (5.2.1) follow.

It remains to show boundedness of the second term in (5.2.1). Using Hölder's inequality and (H4), we obtain

$$\begin{aligned}\|A\big(X^{(n)}(.)\big)\|_{K^*}&=\sup_{f\in K,\|f\|_K=1}\mathbb{E}\int_0^T\big(A(X^{(n)}(s)\big)f(s)\,\mathrm{d}s\le\sup_{f\in K,\|f\|_K=1}\mathbb{E}\int_0^T\|A\big(X^{(n)}(s)\big)\|_{V^*}\|f(s)\|_V\,\mathrm{d}s\\&\le\mathbb{E}\Big(\int_0^T\|A\big(X^{(n)}(s)\big)\|_{V^*}^{\frac{\alpha}{\alpha-1}}\mathrm{d}s\Big)^{\frac{\alpha-1}{\alpha}}\le\mathbb{E}\Big(\int_0^T\Big(c_5+\|X^{(n)}(s)\|_V^{\alpha-1}\Big)^{\frac{\alpha}{\alpha-1}}\mathrm{d}s\Big)^{\frac{\alpha-1}{\alpha}}\end{aligned}$$

which is bounded with respect to n by the previous step, so the proof of (5.2.1) is complete.

The fact that K is reflexive implies that every bounded sequence in K has a weakly convergent subsequence. By (5.2.1) this means that there is a subsequence $X^{(n_k)}$ of $X^{(n)}$ which converges to some $\tilde{X}$ weakly in K, i.e. $T(X^{(n_k)}) \to T(\tilde{X})$ for every $T \in K^*$. Since the approximating processes are predictable, we can assume that so is $\tilde{X}$.

c) We skip the proof that $\bar{X}$ is a solution.

d) To show uniqueness, assume that X and Y are two solutions with the same initial condition with corresponding $\bar{X}$ and $\bar{Y}$ as in Definition 5.6. Applying Theorem 5.9 to the difference $X-Y$, we get, using (H2),

$$\begin{aligned}
&\|X(t) - Y(t)\|_H^2 \\
&\quad = \int_0^t \Big(2 {}_{V^*}\langle A\big(\bar{X}(s)\big) - A\big(\bar{Y}(s)\big), \bar{X}(s) - \bar{Y}(s)\rangle_V + \|B(\bar{X}(s)) - B(\bar{Y}(s))\|_{L_2(U,H)}^2 \Big) \,\mathrm{d}s + M(t) \\
&\quad \le (c \vee 0) \int_0^t \|\bar{X}(s) - \bar{Y}(s)\|_H^2 \,\mathrm{d}s + M(t) \\
&\quad = (c \vee 0) \int_0^t \|X(s) - Y(s)\|_H^2 \,\mathrm{d}s + M(t),
\end{aligned}$$

where M is a continuous local martingale and the last equality holds for all $t \in [0, T]$ and almost all $\omega \in \Omega$, so uniqueness follows from the stochastic Gronwall lemma.

$\square$

Chapter 6

The martingale measure approach to SPDEs

This chapter follows [W86] rather closely. Another reference is [DKMNX09], in particular the first article.

The random wave equation in the one-dimensional case can be written as

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \dot{W}(t,x),$$

where $\dot{W}(t,x)$ is space-time white noise. Our previous results do not apply to this equation. We first introduce the Brownian sheet and white noise.

6.1 White noise and the Brownian sheet

Let $(E,\mathcal{E},\nu)$ be a σ-finite measure space and $\tilde{\mathcal{E}} := \{A \in \mathcal{E} : \nu(A) < \infty\}$.

Definition 6.1. A *white noise based on* ν is a map $W : \tilde{\mathcal{E}} \times \Omega \to \mathbb{R}$ such that

(i) $(\Omega,\mathcal{F},\mathbb{P})$ is a probability space,

(ii) $W(A)$ is measurable and $\mathcal{N}(0,\nu(A))$-distributed for every $A \in \tilde{\mathcal{E}}$,

(iii) if $A \cap B = \emptyset$, $A, B \in \tilde{\mathcal{E}}$, then $W(A)$ and $W(B)$ are independent and $W(A \cup B) = W(A) + W(B)$ a.s.

Does there exist a white noise associated to ν? We show that the answer is *yes*.

If W is white noise, then W is a Gaussian process with index set $\tilde{E}$, i.e. the distribution of $\big(W(A_1),...,W(A_n)\big)$ is Gaussian (why?) and for $A, B \in \tilde{E}$ we have

$$\operatorname{cov}\big(W(A),W(B)\big) = \operatorname{cov}\big(W(A),W(A\cap B)+W(B\backslash A)\big) = \nu(A\cap B).$$

Conversely, if ν is given, we define $C(A,B) := \nu(A\cap B)$ for $A, B \in \tilde{E}$. It is easy to see that the function C is nonnegative definite and therefore there exists a centered Gaussian process W with index set $\tilde{E}$ with covariance function C on some probability space $(\Omega,\mathcal{F},\mathbb{P})$ and W is easily seen to be white noise based on ν.

Remark 6.2. If W is white noise based on ν then we can define $\int f\,\mathrm{d}W$ for $f \in L^2(E, \mathcal{E}, \nu)$ in a straightforward way by first defining the integral for linear combinations of indicators of sets in $\tilde{E}$ and then for general f by approximation. Then

$$\mathbb{E}\Big(\Big(\int f\,\mathrm{d}W\Big)\Big(\int g\,\mathrm{d}W\Big)\Big) = \int fg\,\mathrm{d}\nu.$$

Note that here we restrict our attention to deterministic functions f which is quite a bit easier than the case of random functions f.

An important special case is covered by the following definition.

Definition 6.3. Let $E = \mathbb{R}_+^n := \{(t_1, ..., t_n) : t_i \geq 0,\ i = 1, ...n\}$ be equipped with the Borel σ-algebra and let $\lambda = \nu$ be Lebesgue measure on E. For $t = (t_1, ..., t_n) \in E$ define $(0, t] := (0, t_1] \times ... \times (0, t_n]$. If W is white noise based on λ, then

$$W_t := W\big((0, t]\big),\ t \in E$$

is called (d-dimensional) *Brownian sheet.*

Remark 6.4.

- W_t, $t \in E$ is centered Gaussian with covariance

$$\mathbb{E}\big(W_s W_t\big) = \mathbb{E}\big(W((0, s])W((0, t])\big) = \lambda\big([0, s] \cap [0, t]\big) = \big(s_1 \wedge t_1\big) \cdots \big(s_n \wedge t_n\big) =: s \wedge t.$$

  In particular W_t, $t \in E$ is a standard Brownian motion in case $n = 1$.

- For fixed $t_2, ..., t_n \geq 0$, the process $t_1 \mapsto W_{(t_1,...,t_n)}$ is (one-dimensional) Brownian motion with "volatility" $t_2 \cdots t_n$.

Proposition 6.5. *The trajectories of the Brownian sheet have a continuous modification which is even Hölder $\frac{1}{2} - \varepsilon$ for any $\varepsilon > 0$.*

Proof. This follows from a straightforward application of Kolmogorov's continuity theorem (as in case $n = 1$). □

Remark 6.6. Let W be an $n = d + 1$-dimensional Brownian sheet (with $d \geq 1$). Denote the coordinates by $(t, x_1, ..., x_d)$. Then $\mu := \mathcal{L}\big(W_{1,.}\big)$ is a centered Gaussian measure on the Hilbert space $H = L^2\big([0, \infty)^d, \mathcal{B}\big([0, \infty)^d\big), \mathrm{e}^{-(x_1 + ... + x_d)}\,\mathrm{d}x\big)$. Let us compute the covariance operator Q. For $f, g \in H$ and $V(x) := W_{1,x}$, $x \in \mathbb{R}_+^d$ we have

$$\begin{aligned}\mathbb{E}\big(\langle V, f\rangle_H \langle V, g\rangle_H\big) &= \int_{\mathbb{R}_+^d} \int_{\mathbb{R}_+^d} f(x)g(y)\mathbb{E}\big(V(x)V(y)\big)\mathrm{e}^{-|x|-|y|}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}_+^d} \int_{\mathbb{R}_+^d} f(x)g(y)(x \wedge y)\mathrm{e}^{-|x|-|y|}\,\mathrm{d}x\mathrm{d}y = \langle Qf, g\rangle_H,\end{aligned}$$

where

$$Qf(y) = \int_{\mathbb{R}_+^d} f(x)(x \wedge y)\mathrm{e}^{-|x|}\,\mathrm{d}x.$$

It follows that W_t, $t \geq 0$ is a Q-Wiener process.

Let us explain the interesting phenomenon of *propagation of singularities*. For ease of exposition we assume $n = 2$. For a proof see [W86, p208]. We will give a heuristic explanation in class.

Proposition 6.7. *Fix $t_0 > 0$ and a random variable $S \geq 0$ which is measurable with respect to $\sigma(W_{s,t},\ s \geq 0,\ t \in [0, t_0])$ and for which*

$$\limsup_{h \downarrow 0} \frac{W_{S+h,t} - W_{S,t}}{\sqrt{2h \log\log 1/h}} = \infty \ \textit{a.s.} \tag{6.1.1}$$

for $t = t_0$ (such random variables exist!). Then, (6.1.1) *holds for all $t \geq t_0$.*

6.2 The Brownian sheet and the vibrating string

We present the following simple example (for details see [W86, p281ff]).

Let $n = 2$ and define

$$D := \{(s,t) \in \mathbb{R}^2 :\ s + t \geq 0\},\ \hat{R}_{st} := D \cap \big((-\infty, s] \times (\infty, t]\big),\ s + t \geq 0.$$

Let W be white noise on $\mathbb{R}^2$ based on Lebesgue measure and define $\hat{W}_{st} = W(\hat{R}_{st})$. Note that $W_{st} := \hat{W}_{st} - \hat{W}_{s0} - \hat{W}_{0t},\ s, t \geq 0$ is the two-dimensional Brownian sheet (draw a picture!).

Let $V(x,t)$ be the vertical position of a guitar string at time $t \geq 0$ and location $x \in \mathbb{R}$. We model the motion by the following stochastic wave equation.

$$\frac{\partial^2 V}{\partial t^2} = \frac{\partial^2 V}{\partial x^2} + \dot{W}(x,t)$$

with initial conditions $V(x,0) = \frac{\partial V}{\partial t}(x,0) = 0,\ x \in \mathbb{R}$, where $\dot{W}(t,x)$ is space-time white noise. So far we have not yet defined what we mean by a solution but let us solve it anyway.

If $\dot{W}(x,t)$ is replaced by some smooth $f(x,t)$, then one easily checks that

$$V(x,t) = \frac{1}{2} \int_0^t \int_{x+s-t}^{x+t-s} f(y,s)\,\mathrm{d}y\,\mathrm{d}s,\ t \geq 0,\ x \in \mathbb{R}$$

solves the equation.

Rotate the coordinates by 45^0: $u := \frac{s-y}{\sqrt{2}},\ v := \frac{s+y}{\sqrt{2}},\ \hat{V}(u,v) = V(y,s),\ \hat{f}(u,v) = f(y,s)$. Then

$$\hat{V}(u,v) = \frac{1}{2} \int_{\hat{R}_{uv}} \hat{f}(u',v')\,\mathrm{d}u'\,\mathrm{d}v'.$$

If we (formally) change $f \mathrm{d}y \mathrm{d}s$ to $W(\mathrm{d}y, \mathrm{d}s)$, then we obtain the explicit formula

$$\hat{V}(u,v) = \frac{1}{2} \hat{W}_{uv},\ u + v \geq 0.$$

This consideration can be made rigorous by interpreting the SPDE in the distributional sense (in a distributional sense, the Brownian sheet is infinitely differentiable and white noise is the space-time derivative of the Brownian sheet).

6.3 Martingale measures

To keep things simple we will not present the general definition of a martingale measure but only that of an *orthogonal martingale measure* with time index set $[0,T]$ for fixed $T>0$ (for the general case of "worthy" martingale measures, see [W86]).

In the whole section, let $(E,\mathcal{E},\nu)$ and $\tilde{\mathcal{E}}$ be as before. Further, from now on, we assume that $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ is a filtered probability space with normal filtration.

Definition 6.8. $M = M_t(A)$, $t\in[0,T]$, $A\in\tilde{\mathcal{E}}$ is called an *orthogonal martingale measure* (based on ν) if, almost surely,

- $M_0(A)=0$ for all $A\in\tilde{\mathcal{E}}$,
- $t\mapsto M_t(A)$ is an $\mathbb{F}$-martingale with càdlàg paths for all $A\in\tilde{\mathcal{E}}$,
- for disjoint sets $A,B\in\tilde{\mathcal{E}}$ we have $M_T(A\cup B)=M_T(A)+M_T(B)$ almost surely,
- $\mathbb{E}M_T^2(A)=\nu(A)$ for all $A\in\tilde{\mathcal{E}}$,
- for any disjoint sets $A,B\in\tilde{\mathcal{E}}$ the martingales $M_t(A)$ and $M_t(B)$ are orthogonal, i.e. $M_t(A)M_t(B)$ is a martingale (or, equivalently, $\mathbb{E}\Big(\big(M_t(A)-M_s(A)\big)\big(M_t(B)-M_s(B)\big)|\mathcal{F}_s\Big)=0$ a.s. whenever $0\le s\le t\le T$).

Remark 6.9. The third property in the previous definition is also true for $t\in[0,T]$ instead of T (why?).

Remark 6.10. If M is an orthogonal martingale measure based on ν, then for any $t\in[0,T]$, $\nu_t(A):=\mathbb{E}M_t^2(A)$, $A\in\tilde{\mathcal{E}}$ is σ-additive (and can be uniquely extended to a σ-finite measure on $(E,\mathcal{E})$).

Example 6.11. Let Π be a Poisson point process on $[0,T]\times E$ with intensity $\lambda\otimes\mu$ (I will explain this in class). Define

$$P_t(A):=\Pi([0,t]\times A),\ t\in[0,T],\ A\in\tilde{\mathcal{E}}.$$

Then $M_t(A):=P_t(A)-t\mu(A)$ defines an orthogonal martingale measure (based on $\nu=T\mu$). In fact, for disjoint sets $A,B\in\tilde{\mathcal{E}}$, the martingales $M_t(A)$ and $M_t(B)$ are not only orthogonal but even independent and $t\mapsto M_t(A)$ is a compensated Poisson process (with intensity $\mu(A)$).

Example 6.12. Let $(E,\mathcal{E},\nu)$ be as above and let W be white noise on $\bar{E}:=E\times[0,T]$ based on $\mu\otimes\lambda_T$ where λ_T is Lebesgue measure on $[0,T]$ and $\mu=\nu/T$. Then $M=M_t(A)$, $t\in[0,T]$, $A\in\tilde{\mathcal{E}}$ is an orthogonal martingale measure based on ν. In fact, for disjoint sets $A,B\in\tilde{\mathcal{E}}$, the martingales $M_t(A)$ and $M_t(B)$ are not only orthogonal but even independent and $t\mapsto M_t(A)$ is Brownian motion (up to a deterministic factor which can be computed by the formula $\mathbb{E}M_1(A)=\mu(A)$ (or $\mathbb{E}M_T(A)=\nu(A)$).

For simplicity we will stick to the set-up in the previous example in the following and even assume that the measure μ (or ν) is finite (so $\tilde{\mathcal{E}}=\mathcal{E}$). We continue to use the letter M for the martingale measure defined in the previous example. We want to define the stochastic integral with respect to M. We proceed as usual by first defining the stochastic integral for elementary processes.

Definition 6.13. Fix $T > 0$. If $f : E\times[0,T]\times\Omega \to \mathbb{R}$ is of the form $f(x,s,\omega) = X(\omega)1_{(a,b]}(s)1_A(x)$, $A \in \mathcal{E}$, $0 \le a < b \le T$, $x \in E$, and X is bounded and $\mathcal{F}_a$-measurable, then we define

$$f \cdot M_t(B) := X(\omega)\big(M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)\big).$$

It is not hard to see that $f \cdot M$ is again an orthogonal martingale measure but it is generally untrue that the martingales $f \cdot M_t(A)$ and $f \cdot M_t(B)$ are independent when A and B are disjoint (try to find a simple example!).

It is clear how to extend the definition of the stochastic integral to elementary processes, i.e. linear combinations of simple processes as above.

As usual, we need to define and compare suitable norms on the spaces of integrands f and a space containing $f \cdot M$. For $f \in L^2(E \times [0,T] \times \Omega, \mu \otimes \lambda \otimes \mathbb{P})$, we define the usual L^2-norm by

$$\|f\|_M^2 := \mathbb{E}\Big(\int_0^T \int_E f^2(x,s,\omega)\,\mathrm{d}\mu(x)\,\mathrm{d}s\Big).$$

For elementary f, $B \in \mathcal{E}$, $t \in [0,T]$, one easily checks that

$$\mathbb{E}\big((f\cdot M)_t^2(B)\big) \le \|f\|_M^2 \tag{6.3.1}$$

(with equality if $t = T$ and $B = E$). Then one defines a suitable class of predictable processes $\mathcal{P}_M$ in $L^2(E \times [0,T] \times \Omega, \mu \otimes \lambda \otimes \mathbb{P})$ in which the elementary processes are dense. This allows us to define $(f \cdot M)_t^2(B)$ for $f \in \mathcal{P}_M$ by approximation. We will skip the proof that $(f \cdot M)_t^2(B)$ has a modification with the desired properties (in particular that it is a martingale measure). We will denote such a modification by $\int_0^t \int_A f(x,s)\,M(\mathrm{d}s,\mathrm{d}x) := f \cdot M_t(A)$. If M is the martingale measure associated to white noise W as in Example 6.12 (which is the only case we are considering) then we also write

$$\int_0^t \int_A f(x,s)\,\dot{W}(s,x)\,\mathrm{d}\mu(x)\,\mathrm{d}s \text{ instead of } (f\cdot M)_t(A).$$

We mention that for $\psi \in \mathcal{P}_M$, $t \in [0,T]$ and $B \in \mathcal{E}$, we have

$$\mathbb{E}\big((\psi\cdot M)_t^2(B)\big) = \mathbb{E}\int_0^t \int_B \psi^2(s,x)\,\mathrm{d}\mu(x)\,\mathrm{d}s.$$

Note that the usual Itô integral (with respect to Brownian motion) is a special case of this integral by letting E be a set of cardinality 1. It is natural to ask how this more general stochastic integral compares to that defined in Chapter 4. Here is at least a partial answer. For more details see [DZ92, p99ff].

Let $U := L^2(E,\mathcal{E},\mu)$ with ONB e_k, $k \in \mathbb{N}$ and let $W(t) := \sum_{k=1}^\infty \beta_k(t)Je_k$, $t \in [0,T]$ be a cylindrical Q-Wiener process with $Q = I$ in U and $J : U \to U_1$ Hilbert-Schmidt and one-to-one as in Chapter 4. Again, we assume that μ is a finite measure. W is white noise on $E \times [0,T]$ based on $\mu \otimes \lambda$ in the following sense: for $A \in \mathcal{E}$ and $0 \le s \le t \le T$ define

$$W(A \times (s,t]) := \sum_{k=1}^\infty \big(\beta_k(t) - \beta_k(s)\big) \int_A e_k(x)\,\mathrm{d}\mu(x).$$

This *is* white noise as in Example 6.12 (check it!). For example, we have, for $0 \le s \le t \le T$ and $A, B \in \mathcal{E}$)

$$\begin{aligned}\mathbb{E}\big(W(A \times (s,t])W(B \times (s,t])\big) &= \sum_{k=1}^{\infty}(t-s)\int_A e_k(x)\,\mathrm{d}\mu(x)\int_B e_k(y)\,\mathrm{d}\mu(y)\\ &= (t-s)\langle 1_A, 1_B\rangle_U = (t-s)\mu(A \cap B).\end{aligned}$$

Now let $M_t(A) := W(A \times (0,t])$ be as before. For $f(x,s,\omega) = X(\omega)1_{(a,b]}(s)1_A(x)$ as in Definition 6.13 we have

$$(f\cdot M)_t(B) = X(\omega)\big(M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)\big) = X(\omega)\sum_{k=1}^{\infty}\Big(\big(\beta_k(t\wedge b) - \beta_k(t\wedge a)\big)\int_{A\cap B} e_k(x)\,\mathrm{d}\mu(x)\Big). \tag{6.3.2}$$

Given $A, B \in \mathcal{E}$ and $X(\omega)$, can we find some predictable $\Phi^{(B)} \in L_2(U, \mathbb{R})$ such that $\int_0^t \Phi(s)\,\mathrm{d}W(s) = (f \cdot M)_t(B)$ for $\Phi(s) = 1_{(a,b]}(s)\Phi^{(B)}$? Indeed we can. Define

$$\Phi^{(B)}(g) := \langle X(\omega)1_{A\cap B}(.), g\rangle_U.$$

Then (at least if X takes only finitely many values)

$$\begin{aligned}\int_0^t \Phi(s)\,\mathrm{d}W(s) &= \Big(X(\omega)1_{A\cap B}(.)J^{-1}\Big)\Big(\sum_{k=1}^{\infty}\beta_k(t\wedge b)Je_k - \sum_{k=1}^{\infty}\beta_k(t\wedge a)Je_k\Big)\\ &= X(\omega)\sum_{k=1}^{\infty}\big(\beta_k(t\wedge b) - \beta_k(t\wedge a)\big)\langle 1_{A\cap B}, e_k\rangle_U,\end{aligned}$$

which equals (6.3.2), so the two integrals coincide.

We point out that stochastic integrals can be defined also for martingale measures which are not white noise but not for completely general martingale measures. In these cases there is no such analogy between the different ways of defining stochastic integrals.

6.4 A parabolic SPDE in one space dimension

This section is based on [W86, p311ff] and Lecture notes by Pardoux which are no longer available. We will not provide full proofs. Consider the SPDE

$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial^2 u}{\partial x^2}(t,x) + f(t,x;u(t,x)) + g(t,x;u(t,x))\dot{W}(t,x),\ t \ge 0,\ 0 \le x \le 1,$$

with initial condition $u(0,x) = u_0(x)$, $x \in [0,1]$ and Dirichlet boundary condition $u(t,0) = u(t,1) = 0$, $t \ge 0$. Here, $\dot{W}$ denotes "space-time white noise", i.e. white noise on $[0,T] \times [0,1]$ based on Lebesgue measure. There are (at least) two ways to define what we mean by a solution of the equation, namely the *weak* formulation (weak is meant in the pde sense and not in the sense of weak solutions as defined in WT4 in the context of martingale problems) and the *mild* formulation. For the weak formulation, we first integrate both sides from time 0 to time t, multiply by a sufficiently smooth function φ from $[0,1]$ to $\mathbb{R}$ and then perform integration by parts with respect to x:

$$\int_0^t u(t,x)\varphi(x)\,\mathrm{d}x = \int_0^t u_0(x)\varphi(x)\,\mathrm{d}x + \int_0^t \int_0^1 u(s,x)\varphi''(x)\,\mathrm{d}x\,\mathrm{d}s$$
$$+ \int_0^t \int_0^1 f(s,x;u(s,x))\varphi(x)\,\mathrm{d}x\,\mathrm{d}s + \int_0^t \int_0^1 g(s,x;u(s,x))\varphi(x)\,W(\mathrm{d}s,\,\mathrm{d}x).$$

A predictable process u with continuous paths is called a *weak solution* if the previous equation holds for every function φ which is continuous on $[0,1]$, and satisfies $\varphi(0) = \varphi(1) = 0$ and $\varphi \in C^2(0,1)$. Note that the stochastic integral has been defined before (provided the integrand is in $\mathcal{P}_M$ where M is the orthogonal martingale measure associated to W).

Let us now explain the *mild* formulation. First, let $p(t,x,y)$, $t \geq 0$, $x,y \in (0,1)$ be the *fundamental solution* of the one-dimensional heat equation on $[0,1]$ with Dirichlet boundary conditions, i.e. the solution of the SPDE above with $f = g = 0$ and with initial condition $p(0,x,.) = \delta_x$ (a unit point mass at x). Then p is given by

$$p(t,x,y) = \frac{1}{\sqrt{4\pi t}} \sum_{n\in\mathbb{Z}} \Big[\exp\Big(-\frac{(2n+y-x)^2}{4t} \Big) - \exp\Big(-\frac{(2n+y+x)^2}{4t} \Big) \Big]$$

as one can easily check. The mild formulation is now given by the following variation of constants type formula:

$$u(t,x) = \int_0^t p(t,x,y)u_0(y)\,\mathrm{d}y + \int_0^t \int_0^1 p(t-s,x,y)f(s,y;u(s,y))\,\mathrm{d}y\,\mathrm{d}s$$
$$+ \int_0^t \int_0^1 p(t-s,x,y)f(s,y;u(s,y))\,W(\mathrm{d}s,\,\mathrm{d}y)$$

and a solution to this equation is called a *mild* solution. It is natural to ask whether weak and mild solutions exist and if so whether or not they coincide. Fortunately, under natural conditions on the coefficients, the answer is positive. We impose the following assumptions (which are stronger than necessary) and skip the proof that under these conditions a continuous predictable process u is a weak solution of the SPDE iff it is a mild solution.

(H1) f, g are bounded and measurable.

(H2) There exists some $k \geq 0$ such that $|f(t,x;r) - f(t,x;\tilde{r})| \leq k|r - \tilde{r}|$ for all $t, r, \tilde{r}$ and the same for g.

Theorem 6.14. *If (H1), (H2) hold and if u_0 is continuous and satisfies $u_0(0) = u_0(1) = 0$, then there exists a unique continuous and predictable mild solution of the SPDE. Moreover,*

$$\sup_{0\leq x\leq 1} \sup_{0\leq t\leq T} \mathbb{E}|u(t,x)|^p < \infty \textit{ for every } p \geq 2.$$

Proof. We first show uniqueness. Assume that u and v are two (continuous) solutions and define $\bar{u} := u - v$. We will continue the proof in class.

Next, we sketch the existence proof. We perform a Picard iteration starting with $u^0(t,x) = 0$. As usual, we then define $u^{n+1}(t,x)$ by replacing u by u^n in the right-hand side of the mild equation and define

$$H_n(t) := \sup_{0\leq x\leq 1} \mathbb{E}\Big(\big| u^{n+1}(t,x) - u^n(t,x) \big|^2 \Big).$$

One shows by induction that H_n is bounded and measurable. We will continue the proof in class.

$\square$

It is of interest to investigate the Hölder-regularity of the solution u.

Theorem 6.15. *Under the assumptions above, the solution u has a modification which is Hölder of order $1/4 - \varepsilon$ jointly in (t, x) for any $\varepsilon > 0$.*

Proof. We check the conditions of Kolmogorov's continuity theorem. Clearly, it suffices to show this for the stochastic integral term

$$v(t,x) = \int_0^t \int_0^1 p(t-s,x,y) g(s,y;u(s,y))\, W(\mathrm{d}s, \mathrm{d}y).$$

For $h, k \geq 0$ and $p \geq 1$, we have

$$\Big(\mathbb{E}|v(t+k,x+h)-v(t,x)|^p\Big)^{1/p} \leq \Big(\mathbb{E}|v(t+k,x+h)-v(t+k,x)|^p\Big)^{1/p} + \Big(\mathbb{E}|v(t+k,x)-v(t,x)|^p\Big)^{1/p}.$$

Both summands on the right hand side can be estimated from above in a way which allows to apply Kolmogorov's theorem. We might do this in class. $\square$

Bibliography

[B98] V. Bogachev, *Gaussian Measures*, AMS, Providence, 1998.

[C13] D. Cohn, *Measure Theory*, Second edition, Birkhäuser/Springer, New York, 2013.

[DKMNX09] R. Dalang, Robert, D. Khoshnevisan, C. Mueller, D. Nualart, Y. Xiao, *A Minicourse on Stochastic Partial Differential Equations*, edited by Khoshnevisan and Firas Rassoul-Agha. Lecture Notes in Mathematics 1962, Springer, Berlin, 2009.

[DZ92] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge Univ. Press, Cambridge, 1992.

[DZ96] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge Univ. Press, Cambridge, 1996.

[J78] K. Jacobs, *Measure and Integral*, Academic Press, New York, 1978.

[K02] O. Kallenberg, *Foundations of Modern Probability*, second edition, Springer, New York, 2002.

[LR15] W. Liu and M. Röckner, *Stochastic Partial Differential Equations*, Springer, 2015.

[PR07] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer, 2007.

[W86] J. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, XIV-1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 265–439.