Solutions of Stochastic Differential Equations in Infinite Dimensional Hilbert Spaces and their Dependence on Initial Data

Diplomarbeit
vorgelegt von
Claudia Knoche und Katja Frieler

Fakultät für Mathematik
Universität Bielefeld
Juni 2001

Contents

Introduction **1**

1 The Stochastic Integral in General Hilbert Spaces **9**

1.1 Infinite dimensional Wiener processes 9

1.2 Martingales in general Banach spaces 17

1.3 The definition of the stochastic integral 20

1.3.1 Scheme of the construction of the stochastic integral . 20

1.3.2 The construction of the stochastic integral in detail . . 21

1.4 Properties of the stochastic integral 34

1.4.1 The Itô formula 37

1.4.2 The Burkholder-Davis-Gundy inequality 38

1.4.3 Stochastic Fubini theorem 42

1.5 The stochastic integral for cylindrical Wiener processes 43

1.5.1 Cylindrical Wiener processes 43

1.5.2 The definition of the stochastic integral for cylindrical Wiener processes 45

2 Strong, Mild and Weak Solutions **49**

3 Existence, Uniqueness and Continuity of the Mild Solution **65**

3.1 Existence, uniqueness and continuity with respect to the initial data 65

3.2 Smoothing property of the semigroup: Pathwise continuity of the mild solution 83

4 Differentiable Dependence on Initial Data for the Mild Solution **91**

5 Second Order Differentiability of the Mild Solution **111**

6 Application: Feller Property of the Transition Semigroup and Kolmogorov Equation **129**
6.1 Feller property of the transition semigroup 130
6.2 Kolmogorov Equation 137

A The Bochner Integral **149**
A.1 Definition of the Bochner integral 149
A.2 Properties of the Bochner integral 151

B Nuclear and Hilbert-Schmidt Operators **153**

C Pseudo Inverse of Linear Operators **159**

D Continuity and Differentiability of Implicit Functions **163**
D.1 Continuity of the implicit function 163
D.2 Different concepts of differentiability in general Banach spaces 164
D.3 First order differentiability of the implicit function 168
D.4 Second order differentiability of the implicit function 174

Symbols **187**

Bibliography **189**

Introduction

One of the most studied equations in mathematics and physics is the following second order parabolic equation

$$\text{(1)} \qquad \begin{cases} \frac{\partial u}{\partial t}(t,x) & = Lu(t,x), \quad t > 0 \\ u(0,x) & = \varphi(x), \quad x \in \mathbb{R}^d \end{cases}$$

where

$$L\varphi(x) = \frac{1}{2} \sum_{i,j=1}^{d} q_{ij}(x) \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} g_i(x) \frac{\partial \varphi}{\partial x_i}(x), \quad x \in \mathbb{R}^d.$$

Therewith $Q(x) = (q_{ij}(x))$ is a nonnegative definite matrix and $G(x) = (g_1(x), \ldots, g_d(x))$ a vector from $\mathbb{R}^d$ for each $x \in \mathbb{R}^d$.
In the mid 1960s Gross [Gr 67] and Daletsky [Dal 66] were the first who considered this kind of equation in an infinite dimensional Hilbert space H instead of $\mathbb{R}^d$. Introducing Hilbert space-valued Wiener processes and the stochastic integral in infinite dimensions they presented the possibility to make use of infinite dimensional stochastic differential equations as an efficient tool to solve the originally deterministic problem (1). In the case of $\mathbb{R}^d$ this approach was proposed by Kolmogorov and Itô. That is why equation (1) is called *Kolmogorov equation*. One of the more recent works where the above equation is considered in the finite dimensional situation and where it is handled by purely probabilistic methods is for example [Ku 90, Chapter 6].
But in any case it is not only important to find a solution of the stochastic equation but it is crucial to analyze its dependence on initial data.
Apart from this application to parabolic equations, stochastic differential equations are also of independent interest for modeling processes in physics (e.g. stochastic equations of the free field or the stochastic quantization of Euclidean quantum fields), economics (e.g. development of the price of a share) and biology (e.g. population genetics). For a more detailed discussion we refer to [DaPrZa 92, p.1-p.11; Za 98, p.204-p.206].

Up to now, there are several papers which deal with the problem of existence of solutions of stochastic equations and their regularity with respect to initial data (see for example [Dal 67], [DaPrZa 92], [DaPrZa 96], [Ce 98], [Ce 2000], [DaPrElZa 95], [ElLi 93], [MasSei 2000]). But from a strict point of view there were some questions left open or at least unclear concerning some of the proofs, as already remarked by C. Zühlsdorff (see [Zü 95]). Since a large number of works are based on the corresponding results, it is the aim of this diploma thesis to provide detailed and complete proofs for these regularity results, with particularly emphasis on the respective weakest possible assumptions to be imposed.

At first we now want to present the framework in which we treat this problem and then we will summarize our main results. We shall follow largely the set-up in the fundamental monographs [DaPrZa 96] and also [DaPrZa 92], and try to stick to the notations therein.

We consider the following type of stochastic differential equations on a separable (infinite dimensional) Hilbert space H

$$\begin{cases} dX(t) &= [AX(t) + F(X(t))]\, dt + B(X(t))\, dW(t),\ t \in [0,T] \\ X(0) &= \xi \end{cases} \tag{2}$$

where $W(t)$, $t \in [0,T]$, is a cylindrical Wiener process on a probability space $(\Omega, \mathcal{F}, P)$ taking values in another Hilbert space U. A is the generator of a C_0-semigroup $S(t)$, $t \in [0,T]$.

Under rather general conditions on the drift term $F : H \to H$ and the diffusion term $B : H \to L(U,H) := \{L : U \to H \mid L \text{ is linear and bounded}\}$ there exists a mild solution of problem (2) that is a predictable process $X(\xi)(t)$, $t \in [0,T]$, such that

$$\begin{aligned} X(\xi)(t) = S(t)\xi &+ \int_0^t S(t-s)F(X(\xi)(s))\, ds \\ &+ \int_0^t S(t-s)B(X(\xi)(s))\, dW(s) \quad P\text{-a.s.}. \end{aligned}$$

As mentioned above the main part of this work deals with the problem which respective conditions imposed on the coefficients F and B and $q \geq p \geq 2$ imply that the mapping $\xi \mapsto X(\xi)$ is once or twice differentiable from L^q to $\mathcal{H}^p(T,H)$, where

$$\mathcal{H}^p(T,H) := \{Y : [0,T] \times \Omega \to H \mid Y \text{ is predictable and } \sup_{t\in[0,T]} \|Y(t)\|_{L^p} < \infty\},$$

equipped with the norm given by $\|Y\|_{\mathcal{H}^p} := \sup_{t\in[0,T]} \|Y(t)\|_{L^p}$. Since a mild solution of the stochastic differential equation (2) is defined implicitly by

$X(\xi) = \mathcal{F}(\xi, X(\xi))$, where $\mathcal{F} : L_0^p \times \mathcal{H}^p(T,H) \to \mathcal{H}^p(T,H)$ is given by

$$\mathcal{F}(\xi, Y)(t) = S(t)\xi + \int_0^t S(t-s)F(Y(s))\,ds + \int_0^t S(t-s)B(Y(s))\,dW(s), \quad t \in [0,T],$$

differentiability properties of X can be deduced from properties of the mapping $\mathcal{F}$. Questions concerning differentiability of implicit functions, however, can be treated on the very abstract level of a general contracting mapping $G : \Lambda \times E \to E$ on arbitrary Banach spaces Λ and E.
There already is an implicit function theorem [Za 98, Theorem 10.2, p.207; Theorem 10.4, p.208] which makes statements about first and second order directional derivatives and which can be applied to the mapping $\mathcal{F}$. But so far, there has been no version of an abstract implicit function theorem which provides explicit statements about first and second order Fréchet differentiability and which works with regard to $\mathcal{F}$.
One of the contributions of our work is to present such a version (see Theorem D.8 and Theorem D.13). Using this we are able to give a complete proof for the first and second order *Fréchet* differentiability of the mild solution $X : L^q \to \mathcal{H}^p(T,H)$ with respect to the initial condition ξ which may even be random. This is our main result (see Theorem 4.3 and Theorem 5.3). There are several papers which have already used this (see [AlKoRö 95, Theorem 1, p.107], [DaPrZa 92, Theorem 9.16, p.258], [DaPrZa 96, Theorem 5.4.2, p.71], [Za 98, Theorem 6.7, p.179]) since, as said before, various versions of such a result have been stated previously but without complete proofs.
This strong kind of differentiability is of great importance with regard to the transition semigroup $p_t\varphi(x) := E(\varphi(X(x)(t)))$, $t \in [0,T]$, corresponding to the mild solution X. In fact, as a consequence we can verify that $p_t : C_b^2 \to C_b^2$, $t \in [0,T]$ (see Theorem 6.1). This plays an important role with respect to the solvability of the Kolmogorov equation (see Theorem 6.4).

Before we describe our main results more precisely, we comment on the history of the problem trying at the same time to identify the motivation and contributions of our work, in more detail.

Daletsky and Belopolskaja belong to the first who deal with this question (see [Dal 67, Theorem 2.1, p.33; Theorem 2.2, p.34] and [BeDal 80, § 2, p.125-p.129]). They consider the case that $A = 0$ and claim that $X : L^{2k} \to \mathcal{H}^2(T,H)$ has continuous derivatives of order $k = 1,2$ under certain regularity assumptions on F and B. But it is not made explicit which kind of differentiability is meant and the proof is not completely presented. Especially, the exact formulation of the abstract implicit function theorem (as well as its proof) the authors refer to, is missing.

The problem is taken up again by Da Prato and Zabczyk in [DaPrZa 92]. There, A is no longer equal to zero. But as in [Dal 67] and [BeDal 80] B is still a mapping from H to the space of all Hilbert-Schmidt operators and not only to the space $L(U, H)$ of all linear and bounded operators from U to H as in this diploma thesis.
Da Prato and Zabczyk state that the mild solution is twice differentiable with respect to a deterministic initial condition under the assumption that the coefficients F and B are twice continuously differentiable with bounded derivatives of order $k = 1, 2$ [DaPrZa 92, Theorem 9.4, p.245]. In this case the implicit function theorem, which the proof is based on, is presented ([DaPrZa 92, Lemma 9.2, p.244]). But in the form given there it is not adequate if one is interested in Fréchet differentiability of the mild solution X. From a strict point of view it turns out that even in the case that one only asks for the directional derivatives of second order, an application is not possible without modification, i.e. without adding a further Banach space E_0 (as done later in [DaPrZa 96]).

As just mentioned, with regard to the second order directional differentiability an adapted version of the implicit function theorem is presented in the book *Ergodicity for Infinite Dimensional Systems* by Da Prato and Zabczyk [DaPrZa 96, Proposition C.1.3, p.319] or in [Za 98, 10.Appendix, p.206-209] where a very detailed proof is given. By virtue of this theorem it is possible to get the existence of the directional derivatives of second order (with respect to deterministic initial conditions) even in the case that only the composition $S(t)B(x)$, $x \in H$, is a Hilbert-Schmidt operator ([DaPrZa 96, Theorem 5.4.1, p.69]). For that F and $B : H \to L(U, H)$ are required to have bounded and continuous Fréchet derivatives of order $k = 1, 2$ and further assumptions are made concerning the Lipschitz continuity of $S(t)B$ and the boundedness of $S(t)D^2B$ in the space of Hilbert-Schmidt operators as functions of t.
For us, however, there still remains a problem to apply this version of the implicit function theorem without making a slightly stronger assumption on $S(t)DB$ and $S(t)D^2B$ [Hypothesis H.1, p.91; Hypothesis H.2, p.111]. From a strict point of view, despite our stronger conditions, the question on the Fréchet differentiability (treated in this diploma thesis) of the mild solution remains open, even concerning the first order. That is, because the implicit function theorem in [DaPrZa 96] does not make any statement about it.

Now we want to go into the particulars of the structure of our work summarizing the contents and results chapterwise. As said before, our setting is taken from the book *Ergodicity of Infinite Dimensional Systems* by Da Prato and Zabczyk [DaPrZa 96].

In Chapter 1 a detailed introduction to the theory of stochastic integration on infinite dimensional separable Hilbert spaces is given. In this context the notion of a (cylindrical) Wiener process is presented. The book *Stochastic Equations in Infinite Dimensions* by Da Prato and Zabczyk [DaPrZa 92] is the work we mainly refer to in this chapter.

In Chapter 2 different kinds of solutions of stochastic differential equations are compared. In addition to the notion of mild solution which we are interested in, there are at least two other concepts of solutions of stochastic differential equations called weak and strong solution. They are presented and the (essentially known) relations between these three different kinds of solutions are worked out.

In Chapter 3 we are now able to show that Lipschitz assumptions on F and $S(t)B$ imply Lipschitz properties of $\mathcal{F}$ which by the preparations of Chapter 1 can now be defined. So, we can finally use the first part of the implicit function theorem (Theorem D.1 (ii)) to get that there is a unique mild solution X of problem (2) which is Lipschitz continuous with respect to the initial condition.

In Chapter 4 we analyze the first order differentiability of the mapping $\xi \mapsto X(\xi)$ in the case where $F : H \to H$ and $B : H \to L(U, H)$ are continuously Fréchet differentiable with bounded derivatives. First, we prove the Gâteaux differentiability, i.e. the existence of the directional derivatives as linear and bounded operators with respect to the directions (see Theorem 4.3 (i)). For the proof one can refer to the version of the implicit function theorem given by Zabczyk [Za 98, Theorem 10.2, p. 307] (see also Theorem D.8). But in comparison to Da Prato and Zabczyk [DaPrZa 96, Theorem 5.4.1, p.69] we have to supplement the assumptions by a condition concerning the continuity of $S(t)DB$ to justify its application.

To get the existence of the Fréchet derivative of $\xi \mapsto X(\xi)$ we have to modify the abstract implicit function theorem by introducing a further Banach space $E_0 \subset E$ (see Theorem D.8 (ii)) following the idea presented in [DaPrZa 96], to get the second order directional derivatives. That becomes necessary, since the mapping $\mathcal{F}$ is not Fréchet differentiable in the second variable as a mapping from $L_0^p \times \mathcal{H}^p(T, H)$ to $\mathcal{H}^p(T, H)$ but only as a mapping from $L_0^q \times \mathcal{H}^q(T, H)$ to $\mathcal{H}^p(T, H)$ with $q > p$. (see Theorem 4.3 (iv))

Chapter 5 is devoted to the question of the second order differentiability of the mild solution. Analogously to [DaPrZa 96, Theorem 5.4.1 (ii), p.69], we require $F : H \to H$ and $B : H \to L(U, H)$ to be twice continuously Fréchet

differentiable with bounded derivatives of first and second order. In addition, we also need their assumptions concerning the boundedness of $S(t)D^2B$ in the space of all Hilbert-Schmidt operators. But in contrast to [DaPrZa 96, Theorem 5.4.1 (ii), p.69] we think that, additionally, one has to require the continuity of $S(t)D^2B$ (see Hypothesis H.2, p.111) similar to the change of the Hypothesis H.1 in Chapter 4.
Since we are interested in the second order Fréchet differentiability, the abstract theorem in the given form [Za 98, Theorem 10.4, p.208] is not sufficient and has to be supplemented by a second part (Theorem D.13 (ii), (iii)). In our concrete setting that means that we have to deal with the question under which conditions $\mathcal{F}$ is twice continuously Fréchet differentiable in the second variable instead of only having second order directional derivatives. To prove this, one has to consider $\mathcal{F}$ as a mapping from $L_0^q \times \mathcal{H}^q(T, H)$ to $\mathcal{H}^p(T, H)$ with $q > 2p \geq 4$. At first, by Theorem D.13 (ii) we only obtain the existence of the Gâteaux derivative ∂DX in this way. To verify that it is even the Fréchet derivative D^2X the initial condition has to be restricted again such that the conditions of Theorem D.13 (iii) are fulfilled (see Theorem 5.3 (i), (ii)).
Each of the Chapters 3-5 includes an additional part about the pathwise continuity of the mild solution $X(\xi)(t)$, $t \in [0, T]$, (see [DaPrZa 96, Theorem 5.3.1, p.66]) and its derivatives respectively (see Proposition 4.8 and Proposition 5.7). The proofs are not based on the implicit function theorem, but on the well-known factorization formula for stochastic convolution integrals by Da Prato and Zabczyk (see [DaPrZa 96, Theorem 5.2.5, p.58]). To apply it, one needs additional assumptions on the function $K : [0, T] \to [0, \infty[$ which describes the Lipschitz property of $S(t)B$ (see [DaPrZa 96, Theorem 5.3.1, p.66]) and on the function $K_1 : [0, T] \to [0, \infty[$ respectively, which describes the boundedness of $S(t)D^2B$ in terms of t (see Hypothesis H.2 and Proposition 5.7).
Chapter 6 is devoted to applications, in particular with regard to the Kolmogorov equation (1). In its first subsection we show that the transition semigroup given by $p_t\varphi := E(\varphi(X(\cdot)(t)))$ preserves regularity properties of φ in x if one makes corresponding regularity assumptions on the mild solution X with respect to the deterministic initial data $x \in H$ (see Theorem 6.1). For example, Theorem 3.2 provides the conditions which imply the Feller property of the semigroup p_t, $t \in [0, T]$, and Corollary 5.6 states conditions under which we prove that

$$p_t\varphi \in UC_b^2 := \{f \in C_b^2 \mid f \text{ has uniformly continuous derivatives}\},$$

if $\varphi \in UC_b^2$. This result plays an important role in the second section of Chapter 6 where we prove that $p_t\varphi$ is a strict solution of the Kolmogorov equation

$$
(3) \qquad \begin{cases} \frac{\partial u}{\partial t}(t,x) & = \frac{1}{2}\,\mathrm{tr}\,\left[D^2u(t,x)B(x)(B(x))^*\right] \\ & \quad +\langle Ax + F(x), Du(t,x)\rangle, \quad t\in[0,T], x\in D(A) \\ u(0,x) & = \varphi(x), \quad x\in H \end{cases}
$$

associated to our stochastic equation (2) (see Theorem 6.4 (i)). The proof also uses the pathwise continuity of the mild solution and its derivatives, especially these of second order. From our point of view a respective assumption concerning the second order derivative seems to be missing in [DaPrZa 96, Theorem 5.4.2, p.71]. In comparison to [DaPrZa 96, Theorem 5.4.2, p.71], according to our analysis, it seems that a given strict solution is not unique in general, but only under further assumptions (see Theorem 6.4 (ii)).

First of all we want to thank Prof. Dr. Michael Röckner who led us to the study of stochastic differential equations. We are grateful for his guidance and support in this project. He always gave us the opportunity to get an insight into the latest research in and outside the University of Bielefeld. We want to thank him for providing us the possibility to work under almost optimal conditions.
Furthermore, we would like to thank Dr. Wilhelm Stannat for his encouragement and the patience in answering our questions. The door of his office seems to be always open. Moreover we are grateful to Prof. Dr. Sandra Cerrai (Università degli Studi di Firenze) who was in Bielefeld on invitation of Prof. Dr. Michael Röckner within the research project BiBoS. During her visit she gave lectures on stochastic differential equations which were very useful and interesting for us. We thank her for her willingness to discuss our problems.
Finally, we thank God that our friendship survived the ups and downs of the last months.

Chapter 1

The Stochastic Integral in General Hilbert Spaces

We fix two separable Hilbert spaces $(U, \langle\ ,\ \rangle_U)$ and $(H, \langle\ ,\ \rangle)$. The first part of this chapter is devoted to the construction of the stochastic Itô integral

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

where $W(t)$, $t \in [0,T]$, is a Wiener process on U and Φ is a process with values that are linear but not necessarily bounded operators from U to H.

For that we first will have to introduce the notation of the standard Wiener process in infinite dimensions. Then there will be a short section about martingales in general Hilbert spaces. These two concepts are important for the construction of the stochastic integral which will be explained in the following section.
In the second part of this chapter we will present the Itô formula and the stochastic Fubini Theorem and establish basic properties of the stochastic integral, including the Burkholder-Davis-Gundy inequality.
Finally, we will describe how to transmit the definition of the stochastic integral to the case that $W(t)$, $t \in [0,T]$, is a cylindrical Wiener process.

1.1 Infinite dimensional Wiener processes

Definition 1.1. A probability measure μ on $(U, \mathcal{B}(U))$ is called *Gaussian* if all bounded linear mappings

$$\begin{aligned} v' : U &\to \mathbb{R} \\ u &\mapsto \langle u, v \rangle_U, \quad u \in U \end{aligned}$$

have Gaussian laws, i.e. for all $v \in U$ there exist $m := m(v) \in \mathbb{R}$ and $\sigma := \sigma(v) > 0$ such that

$$\mu(v' \in A) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_A e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx \quad \text{for all } A \in \mathcal{B}(U)$$

or

$$\mu = \delta_u \quad \text{for one } u \in U \text{ where } \delta_u \text{ is the Dirac measure in } u.$$

Theorem 1.2. *A measure μ on $(U, \mathcal{B}(U))$ is Gaussian if and only if*

$$\hat{\mu}(u) := \int e^{i\langle u,v\rangle_U}\, \mu(dv) = e^{i\langle m,u\rangle_U - \frac{1}{2}\langle Qu,u\rangle_U}, \quad u \in U,$$

where $m \in U$ and $Q \in L(U)$ is nonnegative, symmetric, with finite trace (see Appendix B.3).
In this case μ will be denoted by $N(m, Q)$ where m is called mean and Q is called covariance. The measure μ is uniquely determined by m and Q.

Proof. [DaPrZa 92, p.55] □

The following result can be found for example in [DaPrZa 92, p.54].

Proposition 1.3. *Let X be a U-valued Gaussian random variable on a probability space $(\Omega, \mathcal{F}, P)$, i.e. there exist $m \in U$ and $Q \in L(U)$ nonnegative, symmetric, with finite trace such that $P \circ X^{-1} = N(m, Q)$.*
Then $\langle X, u\rangle_U$ is normal distributed for all $u \in U$ and the following statements hold

- $E(\langle X, u\rangle_U) = \langle m, u\rangle_U$ *for all* $u \in U$
- $E(\langle X - m, u\rangle_U \langle X - m, v\rangle_U) = \langle Qu, v\rangle_U$ *for all* $u, v \in U$
- $E(\|X - m\|_U^2) = \operatorname{tr} Q$

The following Proposition will lead to a representation of a U-valued Gaussian random variable by the help of real valued Gaussian random variables.

Proposition 1.4. *If $Q \in L(U)$ is nonnegative, symmetric, with finite trace then there exists an orthonormal basis e_k, $k \in \mathbb{N}$, of U such that*

$$Qe_k = \lambda_k e_k, \quad \lambda_k \ge 0,\ k \in \mathbb{N}$$

Proof. [ReSi 72, Theorem VI.21; Theorem VI.16 (Hilbert-Schmidt theorem)]

Proposition 1.5 (Representation of a Gaussian random variable). *Let $m \in U$ and $Q \in L(U)$ be nonnegative, symmetric, with tr $Q < \infty$. In addition, we assume that e_k, $k \in \mathbb{N}$, is an orthonormal basis of U consisting of eigenvectors of Q with corresponding eigenvalues λ_k, $k \in \mathbb{N}$.*
Then a U-valued random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is Gaussian with $P \circ X^{-1} = N(m, Q)$ if and only if

$$X = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k e_k + m$$

where β_k, $k \in \{n \in \mathbb{N} \mid \lambda_n > 0\}$, are independent real valued random variables with $P \circ {\beta_k}^{-1} = N(0,1)$ for all $k \in \mathbb{N}$ with $\lambda_k > 0$. The series converges in $L^2(\Omega, \mathcal{F}, P; U)$.

Proof. 1. Let X be a Gaussian random variable with mean m and covariance Q. Without loss of generality we assume that $\lambda_k > 0$, $k \in \mathbb{N}$.
Then $X = \sum_{k \in \mathbb{N}} \langle X, e_k \rangle e_k$ in U where $\langle X, e_k \rangle$ is normal distributed with mean $\langle m, e_k \rangle$ and variance λ_k, $k \in \mathbb{N}$, by Proposition 1.3. If we define now

$$\beta_k := \frac{\langle X, e_k \rangle - \langle m, e_k \rangle}{\sqrt{\lambda_k}}, \quad k \in \mathbb{N},$$

then we get that $P \circ \beta_k^{-1} = N(0,1)$ and $X = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k e_k + m$. To prove the independence of β_k, $k \in \mathbb{N}$, we take an arbitrary $n \in \mathbb{N}$ and $a_k \in \mathbb{R}$, $1 \leq k \leq n$, and obtain that

$$\sum_{k=1}^{n} a_k \beta_k = \sum_{k=1}^{n} \frac{a_k}{\sqrt{\lambda_k}} \langle X, e_k \rangle + c = \langle X, \sum_{k=1}^{n} \frac{a_k}{\sqrt{\lambda_k}} e_k \rangle + c$$

which is normal distributed since X is a Gaussian random variable. Therefore we have that β_k, $k \in \mathbb{N}$, are a Gaussian family. Hence, to get the independence, we only have to check that the covariance of β_i and β_j, $i, j \in \mathbb{N}$, $i \neq j$, is equal to zero. But this is clear since

$$E(\beta_i \beta_j) = \frac{1}{\sqrt{\lambda_i \lambda_j}} E(\langle X - m, e_i \rangle \langle X - m, e_j \rangle) = \frac{1}{\sqrt{\lambda_i \lambda_j}} \langle Q e_i, e_j \rangle = 0$$

for $i \neq j$.
Besides it is easy to see that the series $\sum_{k=1}^{n} \sqrt{\lambda_k} \beta_k e_k$, $n \in \mathbb{N}$, converges in $L^2(\Omega, \mathcal{F}, P; U)$ since the space is complete and

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

Since $\sum_{k\in\mathbb{N}}\lambda_k = \operatorname{tr} Q < \infty$ this expression becomes arbitrarily small for m and n large enough.
2. Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U such that $Qe_k = \lambda_k e_k$, $k \in \mathbb{N}$, and let β_k, $k \in \mathbb{N}$, be a family of independent real valued Gaussian random variables with mean 0 and variance 1. Then it is clear that the series $\sum_{k=1}^n \sqrt{\lambda_k}\beta_k e_k + m$, $n \in \mathbb{N}$, converges to $X := \sum_{k\in\mathbb{N}} \sqrt{\lambda_k}\beta_k e_k + m$ in $L^2(\Omega, \mathcal{F}, P; U)$ (see 1.). Now we fix $u \in U$ and get that

$$\langle \sum_{k=1}^n \sqrt{\lambda_k}\beta_k e_k + m, u\rangle = \sum_{k=1}^n \sqrt{\lambda_k}\beta_k \langle e_k, u\rangle + \langle m, u\rangle$$

is normal distributed for all $n \in \mathbb{N}$ and the sequence converges in $L^2(\Omega, \mathcal{F}, P)$. This implies that the limit $\langle X, u\rangle$ is also normal distributed where

$$\begin{aligned} E(\langle X, u\rangle) &= E(\sum_{k\in\mathbb{N}} \sqrt{\lambda_k}\beta_k\langle e_k, u\rangle + \langle m, u\rangle) \\ &= \lim_{n\to\infty} E(\sum_{k=1}^n \sqrt{\lambda_k}\beta_k\langle e_k, u\rangle) + \langle m, u\rangle = \langle m, u\rangle \end{aligned}$$

and concerning the covariance we obtain that

$$\begin{aligned} E(\langle X, u\rangle\langle X, v\rangle) &= \lim_{n\to\infty} E(\sum_{k=1}^n \sqrt{\lambda_k}\beta_k\langle e_k, u\rangle \sum_{k=1}^n \sqrt{\lambda_k}\beta_k\langle e_k, v\rangle) \\ &= \sum_{k\in\mathbb{N}} \lambda_k\langle e_k, u\rangle\langle e_k, v\rangle \\ &= \sum_{k\in\mathbb{N}} \langle Qe_k, u\rangle\langle e_k, v\rangle \\ &= \sum_{k\in\mathbb{N}} \langle e_k, Qu\rangle\langle e_k, v\rangle = \langle Qu, v\rangle \end{aligned}$$

for all $u, v \in U$. □

By part 2. of this proof we finally get the following existence result.

Corollary 1.6. *Let Q be a nonnegative and symmetric operator in $L(U)$ with finite trace and let $m \in U$. Then there exists a Gaussian measure $\mu = N(m, Q)$ on $(U, \mathcal{B}(U))$.*

After these preparations we will give the definition of the standard Q-Wiener process. To this end we fix an element $Q \in L(U)$, nonnegative, symmetric and with finite trace and a positive real number T.

Definition 1.7. A U-valued stochastic process $W(t)$, $t \in [0,T]$, on a probability space $(\Omega, \mathcal{F}, P)$ is called (standard) Q-Wiener process if

- $W(0) = 0$
- W has P-a.s. continuous trajectories.
- The increments of W are independent, i.e. that the random variables $W(t_1)$, $W(t_2) - W(t_1)$, $\dots$, $W(t_n) - W(t_{n-1})$ are independent for all $0 \leq t_1 < \dots < t_n \leq T$, $n \in \mathbb{N}$.
- The increments have Gaussian laws, i.e. that $P \circ (W(t) - W(s))^{-1} = N(0, (t-s)Q)$ for all $0 \leq s \leq t \leq T$.

Similar to the existence of Gaussian measures the existence of a Q-Wiener process in U can be traced back to the real valued case. It will be done by the following Proposition.

Proposition 1.8 (Representation of the Q-Wiener process). *Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U consisting of eigenvectors of Q with corresponding eigenvalues λ_k, $k \in \mathbb{N}$. Then we get that a U-valued stochastic process $W(t)$, $t \in [0,T]$, is a Q-Wiener process if and only if*

$$W(t) = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k(t) e_k, \quad t \in [0,T],$$

where β_k, $k \in \{n \in \mathbb{N} \mid \lambda_n > 0\}$, are independent real valued Brownian motions on a probability space $(\Omega, \mathcal{F}, P)$. The series converges in $L^2(\Omega, \mathcal{F}, P; U)$.

Proof. Without loss of generality we assume that $\lambda_k > 0$ for all $k \in \mathbb{N}$.
1. Let $W(t)$, $t \in [0,T]$, be a Q-Wiener process in U.
Since $P \circ W(t)^{-1} = N(0, tQ)$ we know by Proposition 1.5 that

$$W(t) = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k(t) e_k, \quad t \in [0,T],$$

where $P \circ \beta_k^{-1}(t) = N(0,t)$, $k \in \mathbb{N}$, and $\beta_k(t)$, $k \in \mathbb{N}$, are independent for each $t \in [0,T]$.
Now we fix $k \in \mathbb{N}$ and first we show that $\beta_k(t)$, $t \in [0,T]$, is a Brownian motion:
If we take an arbitrary intersection $0 = t_0 \leq t_1 < \dots < t_n \leq T$, $n \in \mathbb{N}$, of $[0,T]$ we get that

$$\beta_k(t_1),\ \beta_k(t_2)-\beta_k(t_1),\dots,\beta_k(t_n)-\beta_k(t_{n-1})$$

are independent for each $k\in\mathbb{N}$ since

$$\beta_k(t_j)-\beta_k(t_{j-1})=\frac{1}{\sqrt{\lambda_k}}\langle W(t_j)-W(t_{j-1}),e_k\rangle,\quad 1\le j\le n.$$

Moreover we obtain for the same reason that $P\circ(\beta_k(t)-\beta_k(s))^{-1}=N(0,t-s)$ for $0\le s\le t\le T$.
In addition

$$t\mapsto\frac{1}{\sqrt{\lambda_k}}\langle W(t),e_k\rangle=\beta_k(t)$$

is P-a.s. continuous for all $k\in\mathbb{N}$.
Secondly it remains to prove that β_k, $k\in\mathbb{N}$, are independent.
We take $k_1,\dots,k_n\in\mathbb{N}$, $n\in\mathbb{N}$, $k_i\neq k_j$ if $i\neq j$ and an arbitrary intersection $0=t_0\le t_1\le\dots\le t_m\le T$, $m\in\mathbb{N}$.
Then we have to show that

$$\sigma(\beta_{k_1}(t_1),\dots,\beta_{k_1}(t_m)),\dots,\sigma(\beta_{k_n}(t_1),\dots,\beta_{k_n}(t_m))$$

are independent.
We will prove this by induction with respect to m:
If $m=1$ it is clear that $\beta_{k_1}(t_1),\dots,\beta_{k_n}(t_1)$ are independent because of Proposition 1.5. Thus we take now an intersection $0=t_0\le t_1\le\dots\le t_{m+1}\le T$ and assume that

$$\sigma\big(\beta_{k_1}(t_1),\dots,\beta_{k_1}(t_m)\big),\dots,\sigma\big(\beta_{k_n}(t_1),\dots,\beta_{k_n}(t_m)\big)$$

are independent. We notice that

$$\begin{aligned}&\sigma(\beta_{k_i}(t_1),\dots,\beta_{k_i}(t_m),\beta_{k_i}(t_{m+1}))\\ &=\sigma(\beta_{k_i}(t_1),\dots,\beta_{k_i}(t_m),\beta_{k_i}(t_{m+1})-\beta_{k_i}(t_m)),\quad 1\le i\le n,\end{aligned}$$

and that $\beta_{k_i}(t_{m+1})-\beta_{k_i}(t_m)=\frac{1}{\sqrt{\lambda_{k_i}}}\langle W(t_{m+1})-W(t_m),e_{k_i}\rangle_U$, $1\le i\le n$, are independent by Proposition 1.5 since $W(t_{m+1})-W(t_m)$ is a Gaussian random variable. If we take $A_{i,j}\in\mathcal{B}(\mathbb{R})$, $1\le i\le n$, $1\le j\le m+1$, then we get because of the independence of $\sigma(W(s)\mid s\le t_m)$ and $\sigma(W(t_{m+1})-W(t_m))$ that

$$P(\underbrace{\bigcap_{i=1}^{n}\bigcap_{j=1}^{m}\{\beta_{k_i}(t_j)\in A_{i,j}\}}_{\in\,\sigma(W(s)\mid s\le t_m)}\cap\underbrace{\bigcap_{i=1}^{n}\{\beta_{k_i}(t_{m+1})-\beta_{k_i}(t_m)\in A_{i,m+1}\}}_{\in\,\sigma(W(t_{m+1})-W(t_m))})$$

$$= P(\bigcap_{i=1}^{n}\bigcap_{j=1}^{m}\{\beta_{k_i}(t_j) \in A_{i,j}\})P(\bigcap_{i=1}^{n}\{\beta_{k_i}(t_{m+1}) - \beta_{k_i}(t_m) \in A_{i,m+1}\})$$

$$= \Big(\prod_{i=1}^{n} P(\bigcap_{j=1}^{m}\{\beta_{k_i}(t_j) \in A_{i,j}\})\Big)\Big(\prod_{i=1}^{n} P(\beta_{k_i}(t_{m+1}) - \beta_{k_i}(t_m) \in A_{i,m+1})\Big)$$

$$= \prod_{i=1}^{n} P(\bigcap_{j=1}^{m}\{\beta_{k_i}(t_j) \in A_{i,j}\} \cap \{\beta_{k_i}(t_{m+1}) - \beta_{k_i}(t_m) \in A_{i,m+1}\})$$

and therefore the assertion follows.
2. If we define

$$W(t) := \sum_{k \in \mathbb{N}} \sqrt{\lambda_k}\beta_k(t)e_k, \quad t \in [0,T]$$

where β_k, $k \in \mathbb{N}$, are independent real valued continuous Brownian motions then it is clear that $W(t)$, $t \in [0,T]$, is well defined in $L^2(\Omega, \mathcal{F}, P; U)$. Besides it is obvious that the process $W(t)$, $t \in [0,T]$, starts in zero and that $P \circ (W(t) - W(s))^{-1} = N(0, (t-s)Q)$, $0 \leq s < t \leq T$, by Proposition 1.5. It is also clear that the increments are independent.
Thus it remains to show that the trajectories of $W(t)$, $t \in [0,T]$, are P-a.s. continuous. For this end we set

$$W^N(t,\omega) := \sum_{k=1}^{N} \sqrt{\lambda_k}\beta_k(t,\omega)e_k$$

for all $(t,\omega) \in \Omega_T := [0,T] \times \Omega$ and $N \in \mathbb{N}$. Then W^N, $N \in \mathbb{N}$, is P-a.s. continuous and we get that

$$\begin{aligned} E(\sup_{t\in[0,T]} \|W^N(t) - W^M(t)\|_U^2) &= E(\sup_{t\in[0,T]} \sum_{k=M+1}^{N} \lambda_k \beta_k^2(t)) \\ &\leq \sum_{k=M+1}^{N} \lambda_k E(\sup_{t\in[0,T]} \beta_k^2(t)) \\ &\leq c \sum_{k=M+1}^{N} \lambda_k \end{aligned}$$

where $c = E(\sup_{t\in[0,T]} \beta_1^2(t)) < \infty$ because of the maximal inequality for real valued martingales. As $\sum_{k\in\mathbb{N}} \lambda_k = \operatorname{tr} Q < \infty$ it follows from the Tchebychev inequality that there is an increasing sequence $(N_m)_{m\in\mathbb{N}}$ such that

$$P(\sup_{t\in[0,T]} \|W^{N_{m+1}}(t) - W^{N_m}(t)\|_U \geq \frac{1}{2^m}) \leq \frac{1}{2^m}, \quad m \in \mathbb{N}.$$

Hence the lemma of Borel-Cantelli provides that W^{N_m}, $m \in \mathbb{N}$, is P-a.s. uniformly convergent and this implies that there is a continuous version of $W(t)$, $t \in [0,T]$, i.e. that there exists a continuous process $\tilde{W}(t)$, $t \in [0,T]$, such that $P(W(t) = \tilde{W}(t)) = 1$ for all $t \in [0,T]$ (see [Za 98, Proposition 2.10, p.128]). □

Definition 1.9 (Normal filtration). A filtration $\mathcal{F}_t$, $t \in [0,T]$, on a probability space $(\Omega, \mathcal{F}, P)$ is called normal if

- $\mathcal{F}_0$ contains all elements $A \in \mathcal{F}$ with $P(A) = 0$ and
- $\mathcal{F}_t = \mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$ for all $t \in [0,T]$.

Definition 1.10 (Q-Wiener process with respect to a filtration). A Q-Wiener process $W(t)$, $t \in [0,T]$, is called Q-Wiener process with respect to a filtration $\mathcal{F}_t$, $t \in [0,T]$, if

- $W(t)$, $t \in [0,T]$, is adapted to $\mathcal{F}_t$, $t \in [0,T]$, and
- $W(t) - W(s)$ is independent of $\mathcal{F}_s$ for all $0 \le s \le t \le T$.

In fact it is possible to show that any U-valued Q-Wiener process $W(t)$, $t \in [0,T]$, is a Q-Wiener process with respect to a normal filtration:
We define

$$\mathcal{N} := \{A \in \mathcal{F} \mid P(A) = 0\}, \quad \tilde{\mathcal{F}}_t := \sigma(W(s) \mid s \le t)$$
$$\text{and} \quad \tilde{\mathcal{F}}_t^0 := \sigma(\tilde{\mathcal{F}}_t \cup \mathcal{N})$$

Then it is clear that

$$\mathcal{F}_t := \bigcap_{s>t} \tilde{\mathcal{F}}_s^0, \quad t \in [0,T], \tag{1.1}$$

is a normal filtration and we get that

Proposition 1.11. *Let $W(t)$, $t \in [0,T]$, be an arbitrary U-valued Q-Wiener process on a probability space $(\Omega, \mathcal{F}, P)$. Then it is a Q-Wiener process with respect to the normal filtration $\mathcal{F}_t$, $t \in [0,T]$, given by (1.1).*

Proof. It is clear that $W(t)$, $t \in [0,T]$, is adapted to $\mathcal{F}_t$, $t \in [0,T]$. Hence we only have to verify that $W(t) - W(s)$ is independent from $\mathcal{F}_s$, $0 \le s < t \le T$. But if we fix $0 \le s < t \le T$ it is clear that $W(t) - W(s)$ is independent of $\tilde{\mathcal{F}}_s$ since

$$\begin{aligned}&\sigma(W(t_1),\ W(t_2),\dots,W(t_n))\\&=\sigma(W(t_1),\ W(t_2)-W(t_1),\dots,W(t_n)-W(t_{n-1}))\end{aligned}$$

for all $0 \le t_1 < t_2 < \cdots < t_n \le s$. Of course $W(t) - W(s)$ is then also independent of $\tilde{\mathcal{F}}_s^0$. To prove now that $W(t) - W(s)$ is independent of $\mathcal{F}_s$ it is enough to show that

$$P(\{W(t)-W(s)\in A\}\cap B) = P(W(t)-W(s)\in A)P(B)$$

for any $B \in \mathcal{F}_s$ and any closed subset $A \subset U$ as $\mathcal{E} := \{A \subset U \mid A \text{ closed}\}$ generates $\mathcal{B}(U)$ and is stable under finite intersections. We get this result in the following way

$$\begin{aligned}&P(\{W(t)-W(s)\in A\}\cap B)\\&=E\big(1_A\circ(W(t)-W(s))1_B\big)\\&=\lim_{n\to\infty}E\Big(\big[(1-n\ \mathrm{dist}(W(t)-W(s),A))\vee 0\big]1_B\Big)\\&=\lim_{n\to\infty}\lim_{m\to\infty}E\Big(\big[(1-n\ \mathrm{dist}(W(t)-W(s+\frac{1}{m}),A))\vee 0\big]1_B\Big)\\&=\lim_{n\to\infty}\lim_{m\to\infty}E\big((1-n\ \mathrm{dist}(W(t)-W(s+\frac{1}{m}),A))\vee 0\big)P(B)\\&=P(W(t)-W(s)\in A)P(B)\end{aligned}$$

since $W(t)-W(s+\frac{1}{m})$ is independent of $\tilde{\mathcal{F}}^0_{s+\frac{1}{m}} \supset \mathcal{F}_s$ if m is large enough. □

1.2 Martingales in general Banach spaces

Analogous to the real-valued case it is possible to define the conditional expectation of any Bochner integrable random variable with values in an arbitrary separable Banach space $(E, \|\ \|)$. This result is formulated in the following Proposition.

Proposition 1.12 (Existence of the conditional expectation). *Assume that E is a separable Banach space. Let X be a Bochner integrable E-valued random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$.*
Then there exists a unique, up to a set of P-probability zero, integrable E-valued random variable Z, measurable with respect to $\mathcal{G}$ such that

$$\int_A X\,dP = \int_A Z\,dP \qquad \text{for all } A \in \mathcal{G}.$$

The random variable Z will be denoted as $E(X|\mathcal{G})$ and called the conditional expectation *of X given $\mathcal{G}$.*

Proof. [DaPrZa 92, Proposition 1.10, p.27] □

Later we will need the following result

Proposition 1.13. *Let $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ be two measurable spaces and $\Psi : E_1 \times E_2 \to \mathbb{R}$ a bounded measurable function. Let X_1 and X_2 be two random variables on $(\Omega, \mathcal{F}, P)$ with values in $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ respectively, and let $\mathcal{G} \subset \mathcal{F}$ be a fixed σ-field.*
Assume that X_1 in $\mathcal{G}$-measurable and X_2 is independent of $\mathcal{G}$, then

$$E(\Psi(X_1, X_2)|\mathcal{G}) = \hat{\Psi}(X_1)$$

where

$$\hat{\Psi}(x_1) = E(\Psi(x_1, X_2)), \quad x_1 \in E_1.$$

Proof. [DaPraZa 92, Proposition 1.12, p.29] □

Remark 1.14. The previous Proposition can be easily extended to the case that the function Ψ is not necessarily bounded but nonnegative.

Proposition 1.12 is the basis for the generalization of the definition of the martingale:

Definition 1.15. Let $M(t)$, $t \geq 0$, be a stochastic process on $(\Omega, \mathcal{F}, P)$ with values in a separable Banach space E. Besides we consider a filtration $\mathcal{F}_t$, $t \geq 0$, on $(\Omega, \mathcal{F}, P)$.
The process M is called *$\mathcal{F}_t$-martingale*, if

- $E(\|M(t)\|) < \infty$ for all $t \geq 0$
- $M(t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$
- $E(M(t)|\mathcal{F}_s) = M(s)$ P-a.s. for all $0 \leq s \leq t < \infty$

There is the following connection to real valued submartingales.

Proposition 1.16. *If $M(t)$, $t \geq 0$, is an E-valued $\mathcal{F}_t$-martingale then $\|M(t)\|$, $t \geq 0$, is a real valued $\mathcal{F}_t$-submartingale.*

Proof. [DaPrZa 92, Proposition 3.7, p.78] □

Theorem 1.17 (Maximal inequality). *Let $p > 1$ and let E be a separable Banach space.*
If $M(t)$, $t \in [0,T]$, is a continuous E-valued $\mathcal{F}_t$-martingale then

$$E(\sup_{t\in[0,T]} \|M(t)\|^p) \leq (\frac{p}{p-1})^p \sup_{t\in[0,T]} E(\|M(t)\|^p)$$

Proof. The inequality is a consequence of the previous Proposition and the well known maximal inequality for real valued submartingales. □

Now we fix $0 < T < \infty$ and denote by $\mathcal{M}_T^2(E)$ the space of all E-valued continuous, square integrable martingales $M(t)$, $t \in [0,T]$. This space will play an important role with regard to the definition of the stochastic integral. We will use especially the following fact.

Proposition 1.18. *The space $\mathcal{M}_T^2(E)$ equipped with the norm*

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

is a Banach space.

Proof. [DaPrZa 92, Proposition 3.9, p.79] □

Lemma 1.19. *If M_n, $n \in \mathbb{N}$, is a sequence in $\mathcal{M}_T^2(E)$ which converges to M in $\mathcal{M}_T^2$ then there is a subsequence n_k, $k \in \mathbb{N}$, such that M_{n_k}, $k \in \mathbb{N}$, converges to M P-a.s. uniformly on $[0,T]$.*

Proof. As M_n, $n \in \mathbb{N}$, is a Cauchy sequence in $\mathcal{M}_T^2$ we get by Theorem 1.17 that

$$E(\sup_{t\in[0,T]} \|M_n(t) - M_m(t)\|^2) \longrightarrow 0 \quad \text{as } n, m \to \infty$$

Hence we can find a subsequence n_k, $k \in \mathbb{N}$, such that

$$P(\sup_{t\in[0,T]} \|M_{n_{k+1}}(t) - M_{n_k}(t)\| \geq 2^{-k}) \leq 2^{-k}$$

and the assertion follows by the lemma of Borel-Cantelli. □

Proposition 1.20. *Let $T > 0$ and $W(t)$, $t \in [0,T]$, be a U-valued Q-Wiener process with respect to a normal filtration $\mathcal{F}_t$, $t \in [0,T]$, on a probability space $(\Omega, \mathcal{F}, P)$. Then $W(t)$, $t \in [0,T]$, is a continuous square integrable martingale, i.e. $W \in \mathcal{M}_T^2(U)$.*

Proof. The continuity is clear by definition and for each $t \in [0,T]$ we have that $E(\|W(t)\|_U^2) = t \operatorname{tr} Q < \infty$ (see Proposition 1.3). Hence let $0 \le s \le t \le T$ and $A \in \mathcal{F}_s$. Then we get by Proposition A.6 that

$$\begin{aligned}\langle \int_A W(t) - W(s)\, dP, u\rangle_U &= \int_A \langle W(t) - W(s), u\rangle_U\, dP \\ &= P(A) \int \langle W(t) - W(s), u\rangle_U\, dP = 0\end{aligned}$$

for all $u \in U$ as $\mathcal{F}_s$ is independent of $W(t) - W(s)$ and $E(\langle W(t) - W(s), u\rangle_U) = 0$ for all $u \in U$. Therefore

$$\begin{aligned}\int_A W(t)\, dP &= \int_A W(s) + (W(t) - W(s))\, dP \\ &= \int_A W(s)\, dP + \int_A W(t) - W(s)\, dP \\ &= \int_A W(s)\, dP\end{aligned}$$

for all $A \in \mathcal{F}_s$. □

1.3 The definition of the stochastic integral

For the whole section we fix a positive real number T and a probability space $(\Omega, \mathcal{F}, P)$ and we define $\Omega_T := [0,T] \times \Omega$ and $P_T := dx \otimes P$ where dx is the Lebesgue measure.
Moreover let $Q \in L(U)$ be symmetric, nonnegative and with finite trace and we consider a Q-Wiener process $W(t)$, $t \in [0,T]$, with respect to a normal filtration $\mathcal{F}_t$, $t \in [0,T]$.

1.3.1 Scheme of the construction of the stochastic integral

Step 1: At first we consider a certain class $\mathcal{E}$ of elementary $L(U,H)$-valued processes and define the mapping

$$\begin{aligned}\mathrm{Int} : \mathcal{E} &\to \mathcal{M}_T^2(H) =: \mathcal{M}_T^2 \\ \Phi &\mapsto \int_0^t \Phi(s)\, dW(s), \quad t \in [0,T].\end{aligned}$$

Step 2: We prove that there is a certain norm on $\mathcal{E}$ such that

$$\mathrm{Int} : \mathcal{E} \to \mathcal{M}_T^2$$

is an isometry. Since $\mathcal{M}_T^2$ is a Banach space this implies that Int can be extended to the abstract completion $\bar{\mathcal{E}}$ of $\mathcal{E}$. This extension remains isometric and it is unique.
Step 3: We give an explicit representation of $\bar{\mathcal{E}}$.
Step 4: We show how the definition of the stochastic integral can be extended by localization.

1.3.2 The construction of the stochastic integral in detail

Step 1: The class $\mathcal{E}$ of all elementary processes is determined by the following definition.

Definition 1.21 (Elementary process). A $L = L(U,H)$-valued process $\Phi(t)$, $t \in [0,T]$, on $(\Omega, \mathcal{F}, P)$ with normal filtration $\mathcal{F}_t$, $t \in [0,T]$, is said to be elementary if there exist $0 = t_0 < \cdots < t_k = T$, $k \in \mathbb{N}$, such that

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

where

- $\Phi_m : \Omega \to L(U,H)$ is $\mathcal{F}_{t_m}$-measurable, $0 \le m \le k-1$,
- Φ_m takes only a finite number of values in $L(U,H)$, $1 \le m \le k-1$.

If we define now

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

for all $\Phi \in \mathcal{E}$ we have the following important result.

Proposition 1.22. *Let $\Phi \in \mathcal{E}$. Then the stochastic integral $\int_0^t \Phi(s)\, dW(s)$, $t \in [0,T]$, defined in the previous way, is a continuous square integrable martingale with respect to $\mathcal{F}_t$, $t \in [0,T]$, i.e.*

$$\operatorname{Int} : \mathcal{E} \to \mathcal{M}_T^2$$

Proof. Let $\Phi \in \mathcal{E}$ be given by

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

as in Definition 1.21. Then it is clear that

$$t \mapsto \int_0^t \Phi(s)\, dW(s) = \sum_{m=0}^{k-1} \Phi_m(W(t_{m+1} \wedge t) - W(t_m \wedge t))$$

is P-a.s. continuous because of the continuity of the Wiener process and the continuity of $\Phi_m(\omega) : U \to H$, $0 \le m \le k-1$, $\omega \in \Omega$. In addition, we get for each summand that

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

Since $W(t)$, $t \in [0,T]$, is square integrable this implies that $\int_0^t \Phi(s)\, dW(s)$ is square integrable for each $t \in [0,T]$.
To prove the martingale property we take $0 \le s \le t \le T$ and a set A from $\mathcal{F}_s$. If $\{\Phi_m(\omega) \mid \omega \in \Omega\} := \{L_1^m, \ldots, L_{k_m}^m\}$ we obtain by Lemma A.6 and the martingale property of the Wiener process that

$$\begin{aligned}
&\int_A \sum_{m=0}^{k-1} \Phi_m(W(t_{m+1} \wedge t) - W(t_m \wedge t))\, dP \\
&= \sum_{\substack{0 \le m \le k-1 \\ t_{m+1} < s}} \int_A \Phi_m(W(t_{m+1} \wedge s) - W(t_m \wedge s))\, dP \\
&\quad + \sum_{\substack{0 \le m \le k-1 \\ s \le t_{m+1}}} \sum_{j=1}^{k_m} \int_{A \cap \{\Phi_m = L_j^m\}} L_j^m (W(t_{m+1} \wedge t) - W(t_m \wedge t))\, dP \\
&= \sum_{\substack{0 \le m \le k-1 \\ t_{m+1} < s}} \int_A \Phi_m(W(t_{m+1} \wedge s) - W(t_m \wedge s))\, dP \\
&\quad + \sum_{\substack{0 \le m \le k-1 \\ s \le t_{m+1}}} \sum_{j=1}^{k_m} L_j^m \int_{\underbrace{A \cap \{\Phi_m = L_j^m\}}_{\in \mathcal{F}_{s \vee t_m}}} W(t_{m+1} \wedge t) - W(t_m \wedge t)\, dP \\
&= \sum_{\substack{0 \le m \le k-1 \\ t_{m+1} < s}} \int_A \Phi_m(W(t_{m+1} \wedge s) - W(t_m \wedge s))\, dP \\
&\quad + \sum_{\substack{0 \le m \le k-1 \\ t_m < s \le t_{m+1}}} \sum_{j=1}^{k_m} L_j^m \int_{A \cap \{\Phi_m = L_j^m\}} W(t_{m+1} \wedge s) - W(t_m \wedge s)\, dP \\
&= \int_A \sum_{m=0}^{k-1} \Phi_m(W(t_{m+1} \wedge s) - W(t_m \wedge s))\, dP
\end{aligned}$$

□

Step 2: To verify the assertion that there is a norm on $\mathcal{E}$ such that Int: $\mathcal{E} \to \mathcal{M}_T^2$ is an isometry we have to introduce the following notion.

Definition 1.23 (Hilbert-Schmidt operator). Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U. An operator $A \in L(U, H)$ is called Hilbert-Schmidt if

$$\sum_{k \in \mathbb{N}} \langle Ae_k, Ae_k \rangle < \infty$$

In Appendix B we take a close look at this notion. So here we only summerize the results which are important for the construction of the stochastic integral. The definition of Hilbert-Schmidt operator and the number

$$\|A\|_{L_2} := (\sum_{k \in \mathbb{N}} \|Ae_k\|^2)^{\frac{1}{2}}$$

are independent of the choice of the basis (see Remark B.6 (i)). Moreover the space $L_2(U, H)$ of all Hilbert-Schmidt operators from U to H equipped with the inner product

$$\langle A, B \rangle_{L_2} := \sum_{k \in \mathbb{N}} \langle Ae_k, Be_k \rangle$$

is a separable Hilbert space (see Proposition B.7). Later we will use the fact that $\|A\|_{L^2(U,H)} = \|A^*\|_{L^2(H,U)}$ where A^* is the adjoint operator of A (see Remark B.6 (i)).
Besides we have to note the following fact.

Proposition 1.24. *If $Q \in L(U)$ is nonnegative and symmetric then there exists exactly one element $Q^{\frac{1}{2}} \in L(U)$ nonnegative and symmetric such that $Q^{\frac{1}{2}} \circ Q^{\frac{1}{2}} = Q$.*
If, in addition, $\operatorname{tr} Q < \infty$ *we have that $Q^{\frac{1}{2}} \in L_2(U)$ where $\|Q^{\frac{1}{2}}\|_{L_2} = \operatorname{tr} Q$ and of course $L \circ Q^{\frac{1}{2}} \in L_2(U, H)$ for all $L \in L(U, H)$.*

Proof. [ReSi 72, Theorem VI.9, p.196] □

After these preparations we simply calculate the $\mathcal{M}_T^2$-norm of $\int_0^t \Phi(s)\, dW(s)$, $t \in [0, T]$, and get the following result.

Proposition 1.25. *If* $\Phi = \sum_{m=0}^{k-1} \Phi_m 1_{]t_m,t_{m+1}]}$ *is an elementary* $L(U,H)$*-valued process then*

$$\|\int_0^{\cdot} \Phi(s)\, dW(s)\|^2_{\mathcal{M}_T^2} = E(\int_0^T \|\Phi(s) \circ Q^{\frac{1}{2}}\|^2_{L_2}\, ds) =: \|\Phi\|_T^2$$

Proof. If we set $\Delta_m := W(t_{m+1}) - W(t_m)$ then we get that

$$\begin{aligned} &\|\int_0^{\cdot} \Phi(s)\, dW(s)\|^2_{\mathcal{M}_T^2} \\ &= E(\|\int_0^T \Phi(s)\, dW(s)\|^2) \\ &= E(\|\sum_{m=0}^{k-1} \Phi_m \Delta_m\|^2) \\ &= E(\sum_{m=0}^{k-1} \|\Phi_m \Delta_m\|^2) + 2E(\sum_{0 \le m < n \le k-1} \langle \Phi_m \Delta_m, \Phi_n \Delta_n \rangle) \end{aligned}$$

Claim 1:

$$E(\sum_{m=0}^{k-1} \|\Phi_m \Delta_m\|^2) = \sum_{m=0}^{k-1} (t_{m+1} - t_m) E(\|\Phi_m \circ Q^{\frac{1}{2}}\|^2_{L_2})$$

To prove this we take an orthonormal basis f_k, $k \in \mathbb{N}$, of H and get by the Parseval identity and Levi's monotone convergence theorem that

$$E(\|\Phi_m \Delta_m\|^2) = \sum_{l \in \mathbb{N}} E(\langle \Phi_m \Delta_m, f_l \rangle^2) = \sum_{l \in \mathbb{N}} E\big(E(\langle \Delta_m, \Phi_m^* f_l \rangle^2 | \mathcal{F}_{t_m})\big)$$

Taking an orthonormal basis e_k, $k \in \mathbb{N}$, of U we obtain that
$\Phi_m^* f_l = \sum_{k \in \mathbb{N}} \langle f_l, \Phi_m e_k \rangle e_k$. Since $\langle f_l, \Phi_m e_k \rangle$ is $\mathcal{F}_{t_m}$-measurable this implies that $\Phi_m^* f_l$ is $\mathcal{F}_{t_m}$-measurable by Proposition A.3. Using the fact that $\sigma(\Delta_m)$ is independent of $\mathcal{F}_{t_m}$ we obtain by Lemma 1.13 that

$$E(\langle \Delta_m, \Phi_m^* f_l \rangle^2 | \mathcal{F}_{t_m}) = F(\Phi_m^* f_l)$$

where $F(u) = E(\langle \Delta_m, u \rangle^2) = (t_{m+1} - t_m)\langle Qu, u \rangle$. Thus the symmetry of $Q^{\frac{1}{2}}$ finally provides that

$$\begin{aligned} E((\|\Phi_m \Delta_m\|^2) &= \sum_{l \in \mathbb{N}} E\big(E(\langle \Delta_m, \Phi_m^* f_l \rangle^2 | \mathcal{F}_{t_m})\big) \\ &= (t_{m+1} - t_m) \sum_{l \in \mathbb{N}} E(\langle Q \Phi_m^* f_l, \Phi_m^* f_l \rangle) \end{aligned}$$

$$\begin{aligned} &= (t_{m+1} - t_m) \sum_{l \in \mathbb{N}} E(\|Q^{\frac{1}{2}} \Phi_m^* f_l\|^2) \\ &= (t_{m+1} - t_m) E(\|(\Phi_m \circ Q^{\frac{1}{2}})^*\|^2_{L_2(H,U)}) \\ &= (t_{m+1} - t_m) E(\|\Phi_m \circ Q^{\frac{1}{2}}\|^2_{L_2(U,H)}) \end{aligned}$$

Hence the first assertion is proved and it only remains to verify the following claim.
Claim 2:

$$E(\langle \Phi_m \Delta_m, \Phi_n \Delta_n \rangle) = 0, \quad 0 \leq m < n \leq k-1$$

But this can be proved in a similar way as Claim 1:

$$\begin{aligned} E(\langle \Phi_m \Delta_m, \Phi_n \Delta_n \rangle) &= E(E(\langle \Phi_n^* \Phi_m \Delta_m, \Delta_n \rangle | \mathcal{F}_{t_n})) \\ &= E(\tilde{F}(\Phi_n^* \Phi_m \Delta_m)) \end{aligned}$$

where $\tilde{F}(u) = E(\langle u, \Delta_n \rangle_U) = 0$ for all $u \in U$ (see Proposition 1.13). Hence the assertion follows. □

In this way the right norm on $\mathcal{E}$ is actually found but strictly speaking $\| \ \|_T$ is only a semi norm on $\mathcal{E}$. Therefore we have to consider equivalence classes of elementary processes with respect to $\| \ \|_T$ to get a norm on $\mathcal{E}$. For simplicity we will not change the notation but we have to underline the following fact.

Remark 1.26. If two elementary processes Φ and $\tilde{\Phi}$ belong to one equivalence class with respect to $\| \ \|_T$ it does not follow that they are equal P_T-a.e. because their values only have to correspond on $Q^{\frac{1}{2}}(U)$ P_T-a.e..

Thus we finally have shown that

$$\mathrm{Int} : (\mathcal{E}, \| \ \|_T) \to (\mathcal{M}_T^2, \| \ \|_{\mathcal{M}_T^2})$$

is an isometric transformation. Since $\mathcal{E}$ is dense in the abstract completion $\bar{\mathcal{E}}$ of $\mathcal{E}$ with respect to $\| \ \|_T$ it is clear that there is a unique isometric extension of Int to $\bar{\mathcal{E}}$.

Step 3: To give an explicit representation of $\bar{\mathcal{E}}$ it is useful, at this moment, to introduce the subspace $U_0 := Q^{\frac{1}{2}}(U)$ with the inner product given by $\langle u_0, v_0 \rangle_0 := \langle Q^{-\frac{1}{2}} u_0, Q^{-\frac{1}{2}} v_0 \rangle_U$, $u_0, v_0 \in U_0$, where $Q^{-\frac{1}{2}}$ is the pseudo inverse of $Q^{\frac{1}{2}}$ in the case that Q is not one to one. Then we get by Proposition C.3 (i) that $(U_0, \langle \ , \ \rangle_0)$ is again a separable Hilbert space.
The separable Hilbert space $L_2(U_0, H)$ is called L_2^0. By Proposition C.3 (ii)

we know that $Q^{\frac{1}{2}}g_k$, $k \in \mathbb{N}$, is an orthonormal basis of $(U_0, \langle\ ,\ \rangle_0)$ if g_k, $k \in \mathbb{N}$, is an orthonormal basis of $(\operatorname{Ker} Q^{\frac{1}{2}})^{\perp}$. This basis can be supplemented to a basis of U by elements of $\operatorname{Ker} Q^{\frac{1}{2}}$. Thus we obtain that

$$\|L\|_{L_2^0} = \|L \circ Q^{\frac{1}{2}}\|_{L_2} \quad \text{for each } L \in L_2^0.$$

Since $Q^{\frac{1}{2}} \in L_2(U)$ it is clear that $L(U,H) \subset L_2^0$ and that the $\|\ \|_T$-norm of $\Phi \in \mathcal{E}$ can be written in the following way

$$\|\Phi\|_T = (E(\int_0^T \|\Phi(s)\|_{L_2^0}^2 \, ds))^{\frac{1}{2}}$$

Besides we need the following σ-field

$$\begin{aligned}\mathcal{P}_T &:= \sigma(\{]s,t] \times F_s \mid 0 \le s < t \le T,\ F_s \in \mathcal{F}_s\} \cup \{\{0\} \times F_0 \mid F_0 \in \mathcal{F}_0\}) \\ &= \sigma(Y : \Omega_T \to \mathbb{R} \mid Y \text{ is continuous on the left and adapted to} \\ &\qquad\qquad \mathcal{F}_t,\ t \in [0,T]).\end{aligned}$$

Let $\tilde{H}$ be an arbitrary separable Hilbert space. If $Y : \Omega_T \to \tilde{H}$ is $\mathcal{P}_T/\mathcal{B}(\tilde{H})$-measurable it is called ($\tilde{H}$-)predictable.
If, for example, the process Y itself is continuous on the left and adapted to $\mathcal{F}_t$, $t \in [0,T]$, then it is predictable.
So we are now able to characterize $\bar{\mathcal{E}}$.
Claim: There is an explicit representation of $\bar{\mathcal{E}}$ and it is given by

$$\mathcal{N}_W^2(0,T;H) := \{\Phi : [0,T] \times \Omega \to L_2^0 \mid \Phi \text{ is predictable and } \|\Phi\|_T < \infty\}$$

For simplicity we also write $\mathcal{N}_W^2(0,T)$ or $\mathcal{N}_W^2$ instead of $\mathcal{N}_W^2(0,T;H)$.

To prove this claim we first notice the following facts:

1. Since $L(U,H) \subset L_2^0$ and since any $\Phi \in \mathcal{E}$ is L_2^0-predictable by construction we have that $\mathcal{E} \subset \mathcal{N}_W^2$.

2. Because of the completeness of L_2^0 we get by Appendix A that $\mathcal{N}_W^2 = L^2(\Omega_T, \mathcal{P}_T, P_T; L_2^0)$ is also complete.

Therefore $\mathcal{N}_W^2$ is at least a candidate for a representation of $\bar{\mathcal{E}}$. Thus finally there only remains to show that $\mathcal{E}$ is even a dense subset of $\mathcal{N}_W^2$ but this is formulated in Proposition 1.28. It can be proved by the help of the following lemma.

Lemma 1.27. *There is an orthonormal basis of L_2^0 consisting of elements of $L(U,H)$. This implies especially that $L(U,H)$ is a dense subset of L_2^0.*

Proof. Since Q is symmetric, nonnegative and tr $Q < \infty$ we know by Lemma 1.4 that there exists an orthonormal basis e_k, $k \in \mathbb{N}$, of U such that $Qe_k = \lambda_k e_k$, $\lambda_k \geq 0$, $k \in \mathbb{N}$. In this case $Q^{\frac{1}{2}} e_k = \sqrt{\lambda_k} e_k$, $k \in \mathbb{N}$ with $\lambda_k > 0$, is an orthonormal basis of U_0 (see Proposition C.3 (ii)).
If f_k, $k \in \mathbb{N}$, is an orthonormal basis of H we set

$$f_j \otimes e_k := f_j \langle e_k, \cdot \rangle_U \in L(U, H) \subset L_2^0$$

for all $j, k \in \mathbb{N}$. Then we get of course that

$$\begin{aligned}
&\langle \frac{1}{\sqrt{\lambda_k}} f_j \otimes e_k, \frac{1}{\sqrt{\lambda_m}} f_l \otimes e_m \rangle_{L_2^0} \\
&= \sum_{n \in \mathbb{N}} \frac{1}{\sqrt{\lambda_k}\sqrt{\lambda_m}} \langle f_j, f_l \rangle \langle e_k, Q^{\frac{1}{2}} e_n \rangle_U \langle e_m, Q^{\frac{1}{2}} e_n \rangle_U \\
&= \delta_{j,l} \delta_{k,m}
\end{aligned}$$

for all $j, k, l, m \in \mathbb{N}$, with $\lambda_k, \lambda_m > 0$. Hence we have found an orthonormal system. Moreover we get for $L \in L_2^0$ that

$$\langle f_j \otimes e_k, L \rangle_{L_2^0} = \langle f_j, L Q^{\frac{1}{2}} e_k \rangle$$

for all $j, k \in \mathbb{N}$. This implies that $L = 0 \in L_2^0$ if $\langle f_j \otimes e_k, L \rangle_{L_2^0} = 0$ for all $j, k \in \mathbb{N}$ with $\lambda_k > 0$ and in this way we get that

$$\overline{\operatorname{span}(\frac{1}{\sqrt{\lambda_k}} f_j \otimes e_k \mid j, k \in \mathbb{N} \text{ with } \lambda_k > 0)} = L_2^0$$

by [Al 92, 7.4 Lemma, p.213]. □

Proposition 1.28. *If Φ is a L_2^0-predictable process such that $\|\Phi\|_T < \infty$ then there exists a sequence Φ_n, $n \in \mathbb{N}$, of $L(U, H)$-valued elementary processes such that*

$$\|\Phi - \Phi_n\|_T \longrightarrow 0 \quad \textit{as } n \to \infty$$

Proof. Step 1: If $\Phi \in \mathcal{N}_W^2$ there exists a sequence of simple random variables $\Phi_n = \sum_{k=1}^{M_n} L_k^n 1_{A_k^n}$, $A_k^n \in \mathcal{P}_T$ and $L_k^n \in L_2^0$, $n \in \mathbb{N}$, such that

$$\|\Phi - \Phi_n\|_T \longrightarrow 0 \quad \text{as } n \to \infty.$$

As L_2^0 is a Hilbert space this is a simple consequence of Lemma A.4 and Lebesgue's dominated convergence theorem.

Thus the assertion is reduced to the case that $\Phi = L1_A$ where $L \in L_2^0$ and $A \in \mathcal{P}_T$.

Step 2: Let $A \in \mathcal{P}_T$ and $L \in L_2^0$. Then there exists a sequence L_n, $n \in \mathbb{N}$, in $L(U, H)$ such that

$$\|L1_A - L_n 1_A\|_T \longrightarrow 0 \quad \text{as } n \to \infty$$

This result is obvious by Lemma 1.27 and thus now we only have to consider the case that $\Phi = L1_A$, $L \in L(U, H)$ and $A \in \mathcal{P}_T$.

Step 3: If $\Phi = L1_A$, $L \in L(U, H)$, $A \in \mathcal{P}_T$, then there is a sequence Φ_n, $n \in \mathbb{N}$, of elementary $L(U, H)$-valued processes in the sense of Definition 1.21 such that

$$\|L1_A - \Phi_n\|_T \longrightarrow 0 \quad \text{as } n \longrightarrow \infty$$

To get this result it is sufficient to prove that for any $\varepsilon > 0$ there is a finite sum $\Lambda = \bigcup_{n=1}^{N} A_n$ of pairwise disjoint predictable rectangles

$$A_n \in \{]s, t] \times F_s \mid 0 \le s < t \le T, \ F_s \in \mathcal{F}_s\} \cup \{\{0\} \times F_0 \mid F_0 \in \mathcal{F}_0\} =: \mathcal{A}$$

such that

$$P_T((A \setminus \Lambda) \cup (\Lambda \setminus A)) < \varepsilon$$

since then we get that $\sum_{n=1}^{N} L1_{A_n}$ is an elementary process and

$$\|L1_A - \sum_{n=1}^{N} L1_{A_n}\|_T = E(\int_0^T \|L(1_A - \sum_{n=1}^{N} 1_{A_n})\|_{L_2^0} \, ds) \le \varepsilon \|L\|_{L_2^0}$$

Hence we define

$$\mathcal{K} := \{\bigcup_{i \in I} A_i \mid I \text{ is finite and } A_i \in \mathcal{A}, \ i \in I, \ \text{are pairwise disjoint}\}$$

Then $\mathcal{K}$ is a π-field which means that $\mathcal{K}$ is stable under finite intersections. Now let $\mathcal{G}$ be the family of all $A \in \mathcal{P}_T$ which can be approximated by elements of $\mathcal{K}$ in the above sense. Then $\mathcal{G}$ is a Dynkin system and therefore $\mathcal{P}_T = \sigma(\mathcal{K}) = \mathcal{D}(\mathcal{K}) \subset \mathcal{G}$ as $\mathcal{K} \subset \mathcal{G}$. □

Step 4: Finally the so called localization procedure provides the possibility to extend the definition of the stochastic integral even to the linear space

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

For simplicity we also write $\mathcal{N}_W(0,T)$ or $\mathcal{N}_W$ instead of $\mathcal{N}_W(0,T;H)$ and $\mathcal{N}_W$ is called the class of stochastically integrable processes on $[0,T]$.

The extension is done in the following way:
For $\Phi \in \mathcal{N}_W$ we define

$$\tau_n := \inf\{t \in [0,T] \mid \int_0^t \|\Phi(s)\|_{L_2^0}^2\, ds > n\}$$

where we make the convention that the $\inf \emptyset := T$. Then we get by the continuity on the right of the filtration $\mathcal{F}_t$, $t \in [0,T]$, that

$$\begin{aligned}\{\tau_n \le t\} &= \bigcap_{m\in\mathbb{N}} \{\tau_n < t + \frac{1}{m}\} \\ &= \underbrace{\bigcap_{m\in\mathbb{N}} \bigcup_{q\in[0,t+\frac{1}{m}[\cap\mathbb{Q}} \underbrace{\{\int_0^q \|\Phi(s)\|_{L_2^0}^2\, ds > n\}}_{\in\mathcal{F}_q \text{ by the real Fubini theorem}}}_{\in\mathcal{F}_{t+\frac{1}{m}} \text{ and decreasing in } m} \in \mathcal{F}_t\end{aligned}$$

Therefore τ_n, $n \in \mathbb{N}$, is an increasing sequence of stopping times with respect to $\mathcal{F}_t$, $t \in [0,T]$, such that

$$E(\int_0^T \|1_{]0,\tau_n]}(s)\Phi(s)\|_{L_2^0}^2\, ds) < \infty$$

In addition, the processes $1_{]0,\tau_n]}\Phi$, $n \in \mathbb{N}$, are still L_2^0-predictable since

$$\begin{aligned}]0,\tau_n] &:= \{(s,\omega) \in \Omega_T \mid 0 < s \le \tau_n(\omega)\} \\ &= \Big(\{(s,\omega) \in \Omega_T \mid \tau_n(\omega) < s \le T\} \cup \{0\} \times \Omega\Big)^c \\ &= \Big(\bigcup_{q\in\mathbb{Q}} \underbrace{(]q,T] \times \underbrace{\{\tau_n \le q\}}_{\in\mathcal{F}_q})}_{\in\mathcal{P}_T} \cup \{0\} \times \Omega\Big)^c \in \mathcal{P}_T\end{aligned}$$

Thus we get that the stochastic integrals

$$\int_0^t 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s), \quad t \in [0,T],$$

are well defined for all $n \in \mathbb{N}$. For arbitrary $t \in [0,T]$ we set

$$\int_0^t \Phi(s)\, dW(s) := \int_0^t 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s)$$

where n is an arbitrary natural number such that $\tau_n \geq t$. (As the sequence τ_n, $n \in \mathbb{N}$, even reaches T P-a.s. it is also possible to define the integral $\int_0^T \Phi(s)\, dW(s)$ in this way.)
To show that this definition is consistent we have to prove that for arbitrary natural numbers $m < n$ and $t \in [0,T]$

$$\int_0^t 1_{]0,\tau_m]}(s)\Phi(s)\, dW(s) = \int_0^t 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s) \quad P\text{-a.s.}$$

on $\{\tau_m \geq t\} \subset \{\tau_n \geq t\}$. This result is provided by the following lemma.

Lemma 1.29. *Assume that $\Phi \in \mathcal{N}_W^2$ and that τ is a $\mathcal{F}_t$-stopping time such that $P(\tau \leq T) = 1$. Then there exists one P-null set $N \in \mathcal{F}$ independent of $t \in [0,T]$ such that*

$$\int_0^t 1_{]0,\tau]}(s)\Phi(s)\, dW(s) = \int_0^{\tau \wedge t} \Phi(s)\, dW(s) \quad \text{on } N^c \text{ for all } t \in [0,T].$$

Proof. Since both integrals which appear in the equation are P-a.s. continuous we only have to prove that they correspond P-a.s. at any fixed time $t \in [0,T]$.
Step 1: We first consider the case that $\Phi \in \mathcal{E}$ and that τ is a simple stopping time which means that it takes only a finite number of values.
Let $0 = t_0 < t_1 < \ldots < t_k \leq T$, $k \in \mathbb{N}$, and

$$\Phi = \sum_{m=0}^{k-1} \Phi_m 1_{]t_m,t_{m+1}]}$$

where $\Phi_m : \Omega \to L(U,H)$ is $\mathcal{F}_{t_m}$-measurable and takes only a finite number of values, $0 \leq m \leq k-1$.
If τ is a simple stopping time there exists a $n \in \mathbb{N}$ such that
$\tau(\Omega) = \{a_0, \ldots, a_n\}$ and

$$\tau = \sum_{j=0}^{n} a_j 1_{A_j}$$

where $0 \le a_j < a_{j+1} \le T$ and $A_j = \{\tau = a_j\} \in \mathcal{F}_{a_j}$. In this way we get that $1_{]\tau,T]}\Phi$ is an elementary process since

$$\begin{aligned}
1_{]\tau,T]}(s)\Phi(s) &= \sum_{m=0}^{k-1} \Phi_m 1_{]t_m,t_{m+1}]\cap]\tau,T]}(s) \\
&= \sum_{m=0}^{k-1}\sum_{j=0}^{n} 1_{A_j}\Phi_m 1_{]t_m,t_{m+1}]\cap]a_j,T]}(s) \\
&= \sum_{m=0}^{k-1}\sum_{j=0}^{n} \underbrace{1_{A_j}\Phi_m}_{\mathcal{F}_{t_m \vee a_j}\text{-measurable}} 1_{]t_m\vee a_j,t_{m+1}\vee a_j]}(s)
\end{aligned}$$

and concerning the integral, we are interested in, we obtain that

$$\begin{aligned}
&\int_0^t 1_{]0,\tau]}(s)\Phi(s)\, dW(s) \\
&= \int_0^t \Phi(s)\, dW(s) - \int_0^t 1_{]\tau,T]}(s)\Phi(s)\, dW(s) \\
&= \sum_{m=0}^{k-1} \Phi_m(W(t_{m+1}\wedge t) - W(t_m \wedge t)) \\
&\quad - \sum_{m=0}^{k-1}\sum_{j=0}^{n} 1_{A_j}\Phi_m(W((t_{m+1}\vee a_j)\wedge t) - W((t_m \vee a_j)\wedge t)) \\
&= \sum_{m=0}^{k-1} \Phi_m(W(t_{m+1}\wedge t) - W(t_m \wedge t)) \\
&\quad - \sum_{m=0}^{k-1}\sum_{j=0}^{n} 1_{A_j}\Phi_m(W((t_{m+1}\vee \tau)\wedge t) - W((t_m \vee \tau)\wedge t)) \\
&= \sum_{m=0}^{k-1} \Phi_m(W(t_{m+1}\wedge t) - W(t_m \wedge t)) \\
&\quad - \sum_{m=0}^{k-1} \Phi_m(W((t_{m+1}\vee \tau)\wedge t) - W((t_m \vee \tau)\wedge t)) \\
&= \sum_{m=0}^{k-1} \Phi_m\Big(\underbrace{W(t_{m+1}\wedge t) - W(t_m\wedge t) - W((t_{m+1}\vee\tau)\wedge t) + W((t_m\vee\tau)\wedge t)}_{= W(t_{m+1}\wedge\tau\wedge t) - W(t_m\wedge\tau\wedge t)}\Big) \\
&= \int_0^{t\wedge\tau} \Phi(s)\, dW(s)
\end{aligned}$$

Step 2: Now we consider the case that Φ is still an elementary process while τ is an arbitrary stopping time with $P(\tau \leq T) = 1$.

Then there exists a sequence $\tau_n = \sum_{k=0}^{2^n-1} T(k+1)2^{-n} 1_{]Tk2^{-n}, T(k+1)2^{-n}]} \circ \tau$, $n \in \mathbb{N}$, of simple stopping times such that $\tau_n \downarrow \tau$ as $n \to \infty$ and because of the continuity of the stochastic integral we get that

$$\int_0^{\tau_n \wedge t} \Phi(s)\, dW(s) \underset{n\to\infty}{\longrightarrow} \int_0^{\tau \wedge t} \Phi(s)\, dW(s) \quad P\text{-a.s.}$$

Besides we obtain (even for non elementary processes Φ) that

$$\|1_{]0,\tau_n]}\Phi - 1_{]0,\tau]}\Phi\|_T^2 = E\Big(\int_0^T 1_{]\tau,\tau_n]}(s)\|\Phi(s)\|_{L_2^0}^2\, ds\Big) \underset{n\to\infty}{\longrightarrow} 0$$

which by the definition of the integral implies especially that

$$E\Big(\Big\|\int_0^t 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s) - \int_0^t 1_{]0,\tau]}(s)\Phi(s)\, dW(s)\Big\|^2\Big) \underset{n\to\infty}{\longrightarrow} 0$$

for all $t \in [0, T]$. As by Step 1

$$\int_0^t 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s) = \int_0^{\tau_n \wedge t} \Phi(s)\, dW(s), \quad n \in \mathbb{N},\ t \in [0,T],$$

the assertion follows.

Step 3: Finally we generalize the statement to arbitrary $\Phi \in \mathcal{N}_W^2(0,T)$:
If $\Phi \in \mathcal{N}_W^2(0,T)$ then there exists a sequence of elementary processes Φ_n, $n \in \mathbb{N}$, such that

$$\|\Phi_n - \Phi\|_T \underset{n\to\infty}{\longrightarrow} 0$$

By the definition of the stochastic integral that means that

$$\int_0^{\cdot} \Phi_n(s)\, dW(s) \underset{n\to\infty}{\longrightarrow} \int_0^{\cdot} \Phi(s)\, dW(s) \quad \text{in } \mathcal{M}_T^2.$$

Hence it follows from Lemma 1.19 that there is one subsequence n_k, $k \in \mathbb{N}$, and one P-null set $N \in \mathcal{F}$ independent of $t \in [0,T]$ such that

$$\int_0^t \Phi_{n_k}(s)\, dW(s) \underset{k\to\infty}{\longrightarrow} \int_0^t \Phi(s)\, dW(s) \quad \text{on } N^c$$

for all $t \in [0, T]$ and therefore we get for all $t \in [0, T]$ that

$$\int_0^{\tau \wedge t} \Phi_{n_k}(s)\, dW(s) \underset{k\to\infty}{\longrightarrow} \int_0^{\tau \wedge t} \Phi(s)\, dW(s) \quad P\text{-a.s.}.$$

In addition, it is clear that

$$\|1_{]0,\tau]}\Phi_n - 1_{]0,\tau]}\Phi\|_T \underset{n\to\infty}{\longrightarrow} 0$$

which implies that

$$\int_0^t 1_{]0,\tau]}(s)\Phi_n(s)\, dW(s) \underset{n\to\infty}{\longrightarrow} \int_0^t 1_{]0,\tau]}(s)\Phi(s)\, dW(s)$$

As by Step 2

$$\int_0^t 1_{]0,\tau]}(s)\Phi_{n_k}(s)\, dW(s) = \int_0^{\tau\wedge t} \Phi_{n_k}(s)\, dW(s) \quad P\text{-a.s.}$$

for all $k \in \mathbb{N}$ the assertion follows. □

Therefore it is clear that for $m < n$ on $\{\tau_m \geq t\} \subset \{\tau_n \geq t\}$

$$\begin{aligned}
\int_0^t 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s) &= \int_0^{\tau_m \wedge t} 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s) \\
&= \int_0^t 1_{]0,\tau_m]}(s) 1_{]0,\tau_n]}(s)\Phi(s)\, dW(s) \\
&= \int_0^t 1_{]0,\tau_m]}(s)\Phi(s)\, dW(s) \quad P\text{-a.s.}
\end{aligned}$$

and the definition is consistent.

Remark 1.30. In fact it is easy to see that the definition of the stochastic integral does not depend on the choice of τ_n, $n \in \mathbb{N}$. If σ_n, $n \in \mathbb{N}$, is another sequence of stopping times such that $\sigma_n \uparrow T$ as $n \to \infty$ and $1_{]0,\sigma_n]}\Phi \in \mathcal{N}_W^2$ for all $n \in \mathbb{N}$ we also get that

$$\int_0^t \Phi(s)\, dW(s) = \lim_{n\to\infty} \int_0^t 1_{]0,\sigma_n]}(s)\Phi(s)\, dW(s) \quad P\text{-a.s. for all } t \in [0, T]$$

Proof. Let $t \in [0, T]$. Then we get that on the set $\{\tau_m \geq t\}$

$$\int_0^t \Phi(s)\, dW(s) = \int_0^t 1_{]0,\tau_m]}(s)\Phi(s)\, dW(s)$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \int_0^{t\wedge\sigma_n} 1_{]0,\tau_m]}(s)\Phi(s)\, dW(s) \\
&= \lim_{n\to\infty} \int_0^{t\wedge\tau_m} 1_{]0,\sigma_n]}(s)\Phi(s)\, dW(s) \\
&= \lim_{n\to\infty} \int_0^{t} 1_{]0,\sigma_n]}(s)\Phi(s)\, dW(s) \quad P\text{-a.s.}
\end{aligned}
$$

□

1.4 Properties of the stochastic integral

Let T be a positive real number and $W(t)$, $t \in [0,T]$, a Q-Wiener process as described at the beginning of the previous section.

Lemma 1.31. *Let Φ be a L_2^0-valued stochastically integrable process, $(\tilde{H}, \| \ \|_{\tilde{H}})$ a further separable Hilbert space and $L \in L(H, \tilde{H})$.*
Then the process $L(\Phi(t))$, $t \in [0,T]$, is an element of $\mathcal{N}_W(0,T;\tilde{H})$ and

$$
L\Big(\int_0^T \Phi(t)\, dW(t)\Big) = \int_0^T L\big(\Phi(t)\big)\, dW(t) \quad P\text{-a.s.}
$$

Proof. Since Φ is a stochastically integrable process and $\|L(\Phi(t))\|_{L_2(U_0,\tilde{H})} \leq \|L\|_{L(H,\tilde{H})} \|\Phi(t)\|_{L_2^0}$ it is obvious that $L(\Phi(t))$, $t \in [0,T]$, is $L_2(U_0,\tilde{H})$-predictable and

$$
P\Big(\int_0^T \|L(\Phi(t))\|_{L_2(U_0,\tilde{H})}\, dt < \infty\Big) = 1
$$

Step 1: As first step we consider the case that Φ is an elementary process, i.e.

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

where $0 = t_0 < t_1 < ... < t_k = T$, $\Phi_m : \Omega \to L(U,H)$ $\mathcal{F}_{t_m}$-measurable with $|\Phi_m(\Omega)| < \infty$ for $0 \leq m \leq k$. Then

$$
\begin{aligned}
L\Big(\int_0^T \Phi(t)\, dW(t)\Big) &= L\Big(\sum_{m=0}^{k-1} \Phi_m(W(t_{m+1}) - W(t_m))\Big) \\
&= \sum_{m=0}^{k-1} L\big(\Phi_m(W(t_{m+1}) - W(t_m))\big)
\end{aligned}
$$

$$= \int_0^T L\big(\Phi(t)\big)\, dW(t)$$

Step 2: Now let $\Phi \in \mathcal{N}_W^2(0,T)$. Then there exists a sequence Φ_n, $n \in \mathbb{N}$, of elementary processes such that

$$\|\Phi_n - \Phi\|_T = \Big(E\big(\int_0^T \|\Phi_n(t) - \Phi(t)\|_{L_2^0}^2\, dt\big)\Big)^{\frac{1}{2}} \underset{n\to\infty}{\longrightarrow} 0$$

Then $L(\Phi_n)$, $n \in \mathbb{N}$, is a sequence of elementary processes with values in $L(U, \tilde{H})$ and

$$\|L(\Phi_n) - L(\Phi)\|_T \le \|L\|_{L(H,\tilde{H})} \|\Phi_n - \Phi\|_T \underset{n\to\infty}{\longrightarrow} 0$$

By the definition of the stochastic integral, Step 1 and the continuity of L we get that there is a subsequence n_k, $k \in \mathbb{N}$, such that

$$\begin{aligned}
\int_0^T L\big(\Phi(t)\big)\, dW(t) &= \lim_{k\to\infty} \int_0^T L\big(\Phi_{n_k}(t)\big)\, dW(t) \\
&= \lim_{k\to\infty} L\big(\int_0^T \Phi_{n_k}(t)\, dW(t)\big) \\
&= L\big(\lim_{k\to\infty} \int_0^T \Phi_{n_k}(t)\, dW(t)\big) \\
&= L\big(\int_0^T \Phi(t)\, dW(t)\big) \quad P\text{-a.s.}
\end{aligned}$$

Step 3: Finally let $\Phi \in \mathcal{N}_W(0,T)$.
Let τ_n, $n \in \mathbb{N}$, be a sequence of stopping times such that $\tau_n \uparrow T$ as $n \to \infty$ and $1_{]0,\tau_n]}\Phi \in \mathcal{N}_W^2(0,T)$. Then $1_{]0,\tau_n]}L(\Phi) \in \mathcal{N}_W^2(0,T)$ for all $n \in \mathbb{N}$ and we obtain by Remark 1.30 and Step 2:

$$\begin{aligned}
\int_0^T L\big(\Phi(t)\big)\, dW(t) &= \lim_{n\to\infty} \int_0^T 1_{]0,\tau_n]}(t) L\big(\Phi(t)\big)\, dW(t) \\
&= \lim_{n\to\infty} L\big(\int_0^T 1_{]0,\tau_n]}(t)\Phi(t)\, dW(t)\big) \\
&= L\big(\lim_{n\to\infty} \int_0^T 1_{]0,\tau_n]}(t)\Phi(t)\, dW(t)\big) \\
&= L\big(\int_0^T \Phi(t)\, dW(t)\big) \quad P\text{-a.s.}
\end{aligned}$$

□

Lemma 1.32. *Let $\Phi \in \mathcal{N}_W(0,T)$ and ζ_n, $n \in \mathbb{N}$, a sequence in $C([0,T],H)$ which converges uniformly to ζ. Then there exists a subsequence ζ_{n_k}, $k \in \mathbb{N}$, such that*

$$\int_0^T \langle \Phi(t), \zeta_{n_k}(t)\rangle \, dW(t) \underset{k\to\infty}{\longrightarrow} \int_0^T \langle \Phi(t), \zeta(t)\rangle \, dW(t) \quad P\text{-}a.s.$$

Proof. Step 1: Let $\Phi \in \mathcal{N}_W^2(0,T)$.
Then we get that

$$\|\langle \Phi, \zeta_n\rangle - \langle \Phi, \zeta\rangle\|_T \le \sup_{t\in[0,T]} \|\zeta_n(t) - \zeta(t)\| \, \|\Phi\|_T$$

and therefore we get by the isometry that

$$\int_0^T \langle \Phi(t), \zeta_n(t)\rangle \, dW(t) \underset{n\to\infty}{\longrightarrow} \int_0^T \langle \Phi(t), \zeta(t)\rangle \, dW(t)$$

in $L^2(\Omega, \mathcal{F}, P; H)$ which implies that there is a subsequence n_k, $k \in \mathbb{N}$, such that

$$\int_0^T \langle \Phi(t), \zeta_{n_k}(t)\rangle \, dW(t) \underset{k\to\infty}{\longrightarrow} \int_0^T \langle \Phi(t), \zeta(t)\rangle \, dW(t) \quad P\text{-a.s.}$$

Step 2: Let $\Phi \in \mathcal{N}_W(0,T)$.
As in Step 4 of the definition of the stochastic integral we define the stopping times

$$\tau_m := \inf\{t \in [0,T] \mid \int_0^t \|\Phi(s)\|_{L_2}^2 \, ds > m\}$$

where $\inf \emptyset := T$.
Then the processes $\langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta_n(t)\rangle$ and $\langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta(t)\rangle$, $t \in [0,T]$, are in $\mathcal{N}_W^2(0,T;\mathbb{R})$ for all $n, m \in \mathbb{N}$. By Step 1 and the diagonalization procedure we get the existence of one subsequence n_k, $k \in \mathbb{N}$, such that

$$\int_0^T \langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta_{n_k}(t)\rangle \, dW(t) \underset{k\to\infty}{\longrightarrow} \int_0^T \langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta(t)\rangle \, dW(t) \quad P\text{-a.s.}$$

for all $m \in \mathbb{N}$. Hence, by the definition of the stochastic integral, we obtain that

$$
\begin{aligned}
&\int_0^T \langle \Phi(t), \zeta(t)\rangle \, dW(t) \\
&= \sum_{m=1}^{\infty} 1_{\{\tau_{m-1}<T\le\tau_m\}} \int_0^T \langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta(t)\rangle \, dW(t) \\
&= \sum_{m=1}^{\infty} 1_{\{\tau_{m-1}<T\le\tau_m\}} \lim_{k\to\infty} \int_0^T \langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta_{n_k}(t)\rangle \, dW(t) \\
&= \lim_{k\to\infty} \sum_{m=1}^{\infty} 1_{\{\tau_{m-1}<T\le\tau_m\}} \int_0^T \langle 1_{]0,\tau_m]}(t)\Phi(t), \zeta_{n_k}(t)\rangle \, dW(t) \\
&= \lim_{k\to\infty} \int_0^T \langle \Phi(t), \zeta_{n_k}(t)\rangle \, dW(t) \quad P\text{-a.s.}
\end{aligned}
$$

□

1.4.1 The Itô formula

We assume that

- $\Phi \in \mathcal{N}_W(0,T;H)$
- $\varphi : \Omega_T \to H$ is a predictable and P-a.s. Bochner integrable process on $[0,T]$
- $X(0) : \Omega \to H$ is $\mathcal{F}_0$-measurable
- $F : [0,T] \times H \to \mathbb{R}$ is Fréchet differentiable with derivatives

$$
\begin{aligned}
\frac{\partial F}{\partial t} &:= D_1 F : [0,T]\times H \to \mathbb{R} \\
DF &:= D_2 F : [0,T]\times H \to L(H,\mathbb{R}) \cong H \\
D^2 F &:= D_2^2 F : [0,T]\times H \to L(H),
\end{aligned}
$$

which are uniformly continuous on bounded subsets of $[0,T]\times H$.

Under these assumptions the process

$$
X(t) = X(0) + \int_0^t \varphi(s)\, ds + \int_0^t \Phi(s)\, dW(s), \quad t \in [0,T],
$$

is well defined and we get the following result.

Theorem 1.33 (Itô formula). *There exists a P-null set $N \in \mathcal{F}$, independent of $t \in [0,T]$, such that the following formula is fulfilled on N^c for all $t \in [0,T]$:*

$$\begin{aligned} F(t, X(t)) = F(0, X(0)) &+ \int_0^t \langle DF(s, X(s)), \Phi(s)\rangle \, dW(s) \\ &+ \int_0^t \frac{\partial F}{\partial t}(s, X(s)) + \langle DF(s, X(s)), \varphi(s)\rangle \\ &\quad + \frac{1}{2} \operatorname{tr} [D^2F(s, X(s))(\Phi(s)Q^{\frac{1}{2}})(\Phi(s)Q^{\frac{1}{2}})^*] \, ds \end{aligned}$$

Proof. [DaPrZa 92, Theorem 4.17, p.105] □

1.4.2 The Burkholder-Davis-Gundy inequality

Theorem 1.34 (Burkholder-Davis-Gundy inequality). *Let $p \geq 2$ and $\Phi \in \mathcal{N}_W(0,T;H)$. Then*

$$\begin{aligned} &\sup_{t\in[0,T]} E(\| \int_0^t \Phi(s) \, dW(s) \|^p) \\ &\leq \left(\frac{p}{2}(p-1)\right)^{\frac{p}{2}} \Big(\int_0^T \big(E\big(\|\Phi(s)\|^p_{L_2^0} \big)\big)^{\frac{2}{p}} \, ds \Big)^{\frac{p}{2}} \end{aligned}$$

Remark 1.35. If $\Phi \in \mathcal{N}_W^2(0,T)$ we get that $\int_0^t \Phi(s) \, dW(s)$, $t \in [0,T]$, is a martingale and therefore

$$\sup_{t\in[0,T]} E(\| \int_0^t \Phi(s) \, dW(s) \|^p) = E(\| \int_0^T \Phi(s) \, dW(s) \|^p)$$

Proof of Theorem 1.34. We can assume that $E(\int_0^T \|\Phi(t)\|^2_{L_2^0} \, dt) < \infty$ because otherwise the assertion follows immediately since

$$E(\int_0^T \|\Phi(t)\|^2_{L_2^0} \, dt) \leq \int_0^T \big(E\big(\|\Phi(t)\|^p_{L_2^0} \big)\big)^{\frac{2}{p}} \, dt$$

Then, in the case that $p = 2$, we already know by the definition of the stochastic integral that the Burkholder-Davis-Gundy inequality is true. It is even an equality.

If $p > 2$ we want to apply the Itô formula to

- $\Phi : \Omega_T \to L_2^0$
- $\varphi : \Omega_T \to \mathbb{R}$, $(t,\omega) \mapsto 0$
- $X(0) : \Omega \to \mathbb{R}$, $\omega \mapsto 0$
- $F : [0,T] \times H \to \mathbb{R}$, $(t,x) \mapsto \|x\|^p$

The derivatives of F exist and by calculation we get

$$\frac{\partial F}{\partial t}(t,x) = 0$$
$$DF(t,x) = p\|x\|^{p-2}x$$
$$D^2F(t,x) = \begin{cases} 2p\left(\frac{p}{2}-1\right)\|x\|^{p-4}x \otimes x + p\|x\|^{p-2}I_H & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

where $x \otimes y := x\langle y, \cdot\rangle$.
Then $F, \frac{\partial F}{\partial t}, DF$ and D^2F are uniformly continuous on bounded subsets of $[0,T] \times H$ and if we define

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

we get by the Itô formula that for all $t \in [0,T]$

$$\|X(t)\|^p = \int_0^t p\|X(s)\|^{p-2}\langle X(s), \Phi(s)\rangle\, dW(s) + \int_0^t \frac{1}{2}\,\mathrm{tr}\,[D^2F(s,X(s))(\Phi(s)Q^{\frac{1}{2}})(\Phi(s)Q^{\frac{1}{2}})^*]\, ds \tag{1.2}$$

outside a P-null set independent of $t \in [0,T]$. If we set

$$Y(t) := p\|X(t)\|^{p-2}\langle X(t), \Phi(t)\rangle, t \in [0,T],$$

then we have that

$$Y \in \mathcal{N}_W(0,T;\mathbb{R})$$

Now we define $\tau_n := \inf\{t \in [0,T] | \|X(t)\| > n\}, n \in \mathbb{N}$, where $\inf \emptyset := T$. Then

1. τ_n, $n \in \mathbb{N}$, are stopping times with respect to $\mathcal{F}_t$, $t \in [0,T]$

2. $\tau_n \uparrow T$ P-a.s. as $n \to \infty$ since the process $X(t)$, $t \in [0,T]$, is P-a.s. continuous and therefore we have that $t \mapsto X(t)$ is P-a.s. bounded on $[0,T]$.

3. $1_{]0,\tau_n]}Y \in \mathcal{N}_W^2(0,T;\mathbb{R})$ since:
$1_{]0,\tau_n]}Y$ is predictable and

$$\begin{aligned}
&E\Big(\int_0^T \|1_{]0,\tau_n]}(t)Y(t)\|^2_{L_2(U_0,\mathbb{R})}\,dt\Big)\\
&\le E\Big(\int_0^T 1_{]0,\tau_n]}(t)p^2\|X(t)\|^{2p-2}\|\Phi(t)\|^2_{L_2^0}\,dt\Big)\\
&\le p^2n^{2p-2}E\Big(\int_0^T \|\Phi(t)\|^2_{L_2^0}\,dt\Big) < \infty
\end{aligned}$$

If we stop the process $\|X\|^p$ by the stopping times τ_n, $n \in \mathbb{N}$, we get by (1.2) that

$$\begin{aligned}
\|X(T\wedge\tau_n)\|^p = &\int_0^{T\wedge\tau_n} Y(t)\,dW(t)\\
&+\int_0^{T\wedge\tau_n}\frac{1}{2}\,\mathrm{tr}\,[D^2F(t,X(t))(\Phi(t)Q^{\frac12})(\Phi(t)Q^{\frac12})^*]\,dt \quad P\text{-a.s.}
\end{aligned}$$

Taking the expectation on both sides and using Proposition 1.29 we obtain

$$\begin{aligned}
E\big(\|X(T\wedge\tau_n)\|^p\big) = &E\Big(\int_0^T 1_{]0,\tau_n]}(t)Y(t)\,dW(t)\Big)\\
&+E\Big(\int_0^{T\wedge\tau_n}\frac{1}{2}\,\mathrm{tr}\,[D^2F(t,X(t))(\Phi(t)Q^{\frac12})(\Phi(t)Q^{\frac12})^*]\,dt\Big)\\
&=E\Big(\int_0^{T\wedge\tau_n}\frac{1}{2}\,\mathrm{tr}\,[D^2F(t,X(t))(\Phi(t)Q^{\frac12})(\Phi(t)Q^{\frac12})^*]\,dt\Big)
\end{aligned}$$

since $1_{]0,\tau_n]}Y \in \mathcal{N}_W^2(0,T;\mathbb{R})$ and therefore $\int_0^t 1_{]0,\tau_n]}(s)Y(s)\,dW(s)$, $t \in [0,T]$, is a $(\mathcal{F}_t)_{t\in[0,T]}$-martingale which starts in zero. Substituting D^2F by its representation we obtain that

$$\begin{aligned}
&\mathrm{tr}\,[D^2F(t,X(t))(\Phi(t)Q^{\frac12})(\Phi(t)Q^{\frac12})^*]\\
=&\Big(2p\big(\frac{p}{2}-1\big)\|X(t)\|^{p-4}\,\mathrm{tr}\,[(X(t)\otimes X(t))(\Phi(t)Q^{\frac12})(\Phi(t)Q^{\frac12})^*]\\
&+p\|X(t)\|^{p-2}\,\mathrm{tr}\,[(\Phi(t)Q^{\frac12})(\Phi(t)Q^{\frac12})^*]\Big)1_{\{X(t)\neq 0\}}
\end{aligned}$$

Since $\Phi(t)Q^{\frac{1}{2}} \in L_2(U,H)$ and $X(t) \otimes X(t) \in L(H)$ we have that $(X(t) \otimes X(t))\Phi(t)Q^{\frac{1}{2}} \in L_2(U,H)$ and, by Remark B.6, that $\|\Phi(t)Q^{\frac{1}{2}}\|_{L_2} = \|(\Phi(t)Q^{\frac{1}{2}})^*\|_{L_2}$. By Proposition B.8 we therefore obtain that

$$\begin{aligned}
&\text{tr}\,[(X(t) \otimes X(t))(\Phi(t)Q^{\frac{1}{2}})(\Phi(t)Q^{\frac{1}{2}})^*] \\
&\le \|(X(t) \otimes X(t))\Phi(t)Q^{\frac{1}{2}}\|_{L_2}\|\Phi(t)Q^{\frac{1}{2}}\|_{L_2} \\
&\le \|X(t) \otimes X(t)\|_{L(H)}\|\Phi(t)\|^2_{L_2^0} \le \|X(t)\|^2\|\Phi(t)\|^2_{L_2^0}
\end{aligned}$$

and

$$\text{tr}\,[(\Phi(t)Q^{\frac{1}{2}})(\Phi(t)Q^{\frac{1}{2}})^*] = \|\Phi(t)\|^2_{L_2^0}$$

So finally we get that

$$\begin{aligned}
&\text{tr}\,[D^2F(t,X(t))(\Phi(t)Q^{\frac{1}{2}})(\Phi(t)Q^{\frac{1}{2}})^*] \\
&\le 2p\big(\frac{p}{2}-1\big)\|X(t)\|^{p-4}\|X(t)\|^2\|\Phi(t)\|^2_{L_2^0} + p\|X(t)\|^{p-2}\|\Phi(t)\|^2_{L_2^0} \\
&= p(p-1)\|X(t)\|^{p-2}\|\Phi(t)\|^2_{L_2^0}
\end{aligned}$$

Therefore we have for all $n \in \mathbb{N}$ the following estimation for $E(\|X(T \wedge \tau_n)\|^p)$:

$$\begin{aligned}
&E(\|X(T \wedge \tau_n)\|^p) \\
&\le \frac{p}{2}(p-1)E\Big(\int_0^{T\wedge\tau_n} \|X(t)\|^{p-2}\|\Phi(t)\|^2_{L_2^0}\,dt\Big) \\
&= \frac{p}{2}(p-1)E\Big(\int_0^{T\wedge\tau_n} \|X(t\wedge\tau_n)\|^{p-2}\|\Phi(t)\|^2_{L_2^0}\,dt\Big) \\
&\le \frac{p}{2}(p-1)E\Big(\int_0^{T} \|X(t\wedge\tau_n)\|^{p-2}\|\Phi(t)\|^2_{L_2^0}\,dt\Big) \\
&= \frac{p}{2}(p-1)\int_0^T E\big(\|X(t\wedge\tau_n)\|^{p-2}\|\Phi(t)\|^2_{L_2^0}\big)\,dt \\
&\quad \text{(real Fubini theorem)} \\
&\le \frac{p}{2}(p-1)\int_0^T \Big(E\big(\|X(t\wedge\tau_n)\|^p\big)\Big)^{\frac{p-2}{p}}\Big(E\big(\|\Phi(t)\|^p_{L_2^0}\big)\Big)^{\frac{2}{p}}\,dt \\
&\quad \text{(Hölder inequality for } \frac{p}{2} \text{ and } \frac{p}{p-2}\text{)} \\
&\le \frac{p}{2}(p-1)\int_0^T \Big(E\big(\|X(T\wedge\tau_n)\|^p\big)\Big)^{\frac{p-2}{p}}\Big(E\big(\|\Phi(t)\|^p_{L_2^0}\big)\Big)^{\frac{2}{p}}\,dt \\
&\quad (\|X(t\wedge\tau_n)\|^p,\ t\in[0,T], \text{ is a submartingale})
\end{aligned}$$

Dividing both sides of the above inequality by $E\big(\|X(T\wedge\tau_n)\|^p\big)^{\frac{p-2}{p}}<\infty$ we get that

$$\Big(E\big(\|X(T\wedge\tau_n)\|^p\big)\Big)^{\frac{2}{p}}\le\frac{p}{2}(p-1)\int_0^T\Big(E\big(\|\Phi(t)\|^p_{L_2^0}\big)\Big)^{\frac{2}{p}}\,dt$$

for all $n\in\mathbb{N}$.
Since $\tau_n\uparrow T$ and $t\mapsto X(t)=\displaystyle\int_0^t\Phi(s)\,dW(s)$ is P-a.s. continuous we finally obtain by Fatou's lemma that

$$\begin{aligned}\sup_{t\in[0,T]}\Big(E\big(\|\int_0^t\Phi(s)\,dW(s)\|^p\big)\Big)^{\frac{2}{p}}&=\Big(E\big(\|\int_0^T\Phi(t)\,dW(t)\|^p\big)\Big)^{\frac{2}{p}}\\&=\Big(E\big(\|X(T)\|^p\big)\Big)^{\frac{2}{p}}\\&=\Big(E\big(\liminf_{n\to\infty}\|X(T\wedge\tau_n)\|^p\big)\Big)^{\frac{2}{p}}\\&\le\liminf_{n\to\infty}\Big(E\big(\|X(T\wedge\tau_n)\|^p\big)\Big)^{\frac{2}{p}}\\&\le\frac{p}{2}(p-1)\int_0^T\Big(E\big(\|\Phi(t)\|^p_{L_2^0}\big)\Big)^{\frac{2}{p}}\,dt\end{aligned}$$

□

Remark 1.36. The main part of the proof of the Burkholder-Davis-Gundy inequality can be found in [DaPrZa 92, Lemma 7.7, p.194], but we think that in this proof it is not taken into consideration that it could happen that $E\big(\|X(T)\|^p\big)=\infty$. Therefore it could be impossible to divide by this term in the last step without introducing stopping times.

1.4.3 Stochastic Fubini theorem

We assume that

1. $(E,\mathcal{E},\mu)$ is a measure space where μ is finite,
2. $\Phi:\Omega_T\times E\to L_2^0$, $(t,\omega,x)\mapsto\Phi(t,\omega,x)$ is $\mathcal{P}_T\otimes\mathcal{E}/\mathcal{B}(L_2^0)$-measurable, thus in particular $\Phi(\cdot,\cdot,x)$ is a predictable L_2^0-valued process for all $x\in E$.

Theorem 1.37 (Stochastic Fubini theorem). *Assume 1., 2. and that*

$$\int_E\|\Phi(\cdot,\cdot,x)\|_T\,\mu(dx)=\int_E\Big(E\big(\int_0^T\|\Phi(t,\cdot,x)\|^2_{L_2^0}\,dt\big)\Big)^{\frac{1}{2}}\,\mu(dx)<\infty.$$

Then

$$\int_E \Big[\int_0^T \Phi(t,x)\, dW(t) \Big]\, \mu(dx) = \int_0^T \Big[\int_E \Phi(t,x)\, \mu(dx) \Big]\, dW(t) \quad P\text{-}a.s.$$

Proof. [DaPrZa 92, Theorem 4.18, p.109] □

1.5 The stochastic integral for cylindrical Wiener processes

Until now we have considered the case that $W(t)$, $t \in [0,T]$, was a standard Q-Wiener process where $Q \in L(U)$ was nonnegative, symmetric and with finite trace. We could integrate processes in

$$\mathcal{N}_W := \{\Phi : \Omega_T \to L_2(Q^{\frac{1}{2}}(U), H) \mid \Phi \text{ is predictable and } P(\int_0^T \|\Phi(s)\|^2_{L_2^0}\, ds < \infty) = 1\}$$

In fact it is possible to extend the definition of the stochastic integral to the case that Q is not necessarily of finite trace. To this end we first have to introduce the concept of cylindrical Wiener processes.

1.5.1 Cylindrical Wiener processes

Let $Q \in L(U)$ be nonnegative and symmetric. Remember that in the case that Q is of finite trace the Q-Wiener process has the following representation

$$W(t) = \sum_{k \in \mathbb{N}} \beta_k(t) e_k, \quad t \in [0,T],$$

where e_k, $k \in \mathbb{N}$, is an orthonormal basis of $Q^{\frac{1}{2}}(U) = U_0$ and β_k, $k \in \mathbb{N}$, is a family of independent real valued Brownian motions. The series converges in $L^2(\Omega, \mathcal{F}, P; U)$. In the case that Q is no longer of finite trace one loses this convergence. Nevertheless it is possible to define the Wiener process under the following assumption:
There is a further Hilbert space $(U_1, \langle\ ,\ \rangle_1)$ such that there exists a Hilbert-Schmidt embedding

$$J : (U_0, \langle\ ,\ \rangle_0) \to (U_1, \langle\ ,\ \rangle_1).$$

Then the process given by the following Proposition is called *cylindrical Q-Wiener process* in U.

Proposition 1.38. *If e_k, $k \in \mathbb{N}$, is an orthonormal basis of $U_0 = Q^{\frac{1}{2}}(U)$ and β_k, $k \in \mathbb{N}$, is a family of independent real valued Brownian motions then there exists $Q_1 \in L(U_1)$ nonnegative, symmetric and with finite trace such that the series*

$$W(t) = \sum_{k\in\mathbb{N}} \beta_k(t) J(e_k), \quad t \in [0,T],$$

converges in $\mathcal{M}_T^2(U_1)$ and defines a Q_1-Wiener process on U_1. Moreover we have that $\operatorname{Im} Q_1^{\frac{1}{2}} = J(U_0)$ *and for all $u_0 \in U_0$ we get that*

$$\|u_0\|_0 = \|Q_1^{-\frac{1}{2}} J(u_0)\|_1$$

Proof. Step 1: We prove that $W(t)$, $t \in [0,T]$, describes a Q_1-Wiener process in U_1, where $Q_1 = JJ^* \in L(U_1)$:
If we set $\xi_j(t) := \beta_j(t)J(e_j)$, $j \in \mathbb{N}$, we obtain that $\xi_j(t)$, $t \in [0,T]$, is a continuous martingale with respect to $\mathcal{G}_t := \sigma(\bigcup_{j\in\mathbb{N}} \sigma(\beta_j(s)|s \le t))$, $t \in [0,T]$, since

$$E(\beta_j(t) \mid \mathcal{G}_s) = E(\beta_j(t) \mid \sigma(\beta_j(u)|u \le s)) = \beta_j(s) \quad \text{for all } 0 \le s < t \le T$$

as $\sigma\big(\sigma(\beta_j(u)|u \le s) \cup \sigma(\beta_j(t))\big)$ is independent of $\sigma\big(\bigcup_{\substack{k\in\mathbb{N}\\ k\neq j}} \sigma(\beta_k(u)|u \le s)\big)$.

Then it is clear that

$$W_n(t) := \sum_{j=1}^{n} \beta_j(t) J(e_j), \quad t \in [0,T],$$

is also a continuous martingale with respect to $\mathcal{G}_t$, $t \in [0,T]$. In addition, we obtain that

$$E(\|\sum_{j=n}^{m} \beta_j(t) J(e_j)\|_1^2) = t \sum_{j=n}^{m} \|J(e_j)\|_1^2, \quad m \ge n \ge 1,$$

where $\sum_{j\in\mathbb{N}} \|J(e_j)\|_1^2 = \|J\|_{L_2(U_0,U_1)} < \infty$. Therefore we get the convergence of $W_n(t)$, $t \in [0,T]$, in $\mathcal{M}_T^2(U_1)$. This implies especially that the limit $W(t)$, $t \in [0,T]$, is P-a.s. continuous.
Now we want to show that $P \circ (W(t) - W(s))^{-1} = N(0, (t-s)JJ^*)$. Analogously to the second part of the proof of Proposition 1.5 we get that $\langle W(t) - W(s), u_1 \rangle_1$ is normal distributed for all $0 \le s < t \le T$ and $u_1 \in U_1$.

It is easy to see that the mean is equal to zero and concerning the covariance of $\langle W(t)-W(s),u_1\rangle_1$ and $\langle W(t)-W(s),v_1\rangle_1$, $u_1,v_1\in U_1$, we obtain that

$$\begin{aligned}
&E(\langle W(t)-W(s),u_1\rangle_1\langle W(t)-W(s),v_1\rangle_1)\\
=&\sum_{k\in\mathbb{N}}(t-s)\langle Je_k,u_1\rangle_1\langle Je_k,v_1\rangle_1\\
=&(t-s)\sum_{k\in\mathbb{N}}\langle e_k,J^*u_1\rangle_0\langle e_k,J^*v_1\rangle_0\\
=&(t-s)\langle J^*u_1,J^*v_1\rangle_0=(t-s)\langle JJ^*u_1,v_1\rangle_1
\end{aligned}$$

In this way Q_1 is completely determined and thus we obtain that $Q_1=JJ^*$. Thus there only remains to show that the increments of $W(t)$, $t\in[0,T]$, are independent but this can be done in the same way as in the proof of Proposition 1.8.

Step 2: We prove that Im $Q_1^{\frac{1}{2}}=J(U_0)$ and that $\|u_0\|_0=\|Q_1^{-\frac{1}{2}}Ju_0\|_1$ for all $u_0\in U_0$:
Since $Q_1=JJ^*$ we obtain for all $u_1\in U_1$ that

$$\|Q_1^{\frac{1}{2}}u_1\|_1^2=\langle JJ^*u_1,u_1\rangle_1=\|J^*u_1\|_0^2$$

Because of Corollary C.6 this result already implies that Im $Q_1^{\frac{1}{2}}=J(U_0)$ and that $\|Q_1^{-\frac{1}{2}}u_1\|_1=\|J^{-1}u_1\|_0$ for all $u_1\in J(U_0)$. If we replace u_1 by $J(u_0)$, $u_0\in U_0$, we finally get the last assertion because $J:U_0\to U_1$ is one to one. □

1.5.2 The definition of the stochastic integral for cylindrical Wiener processes

We fix $Q\in L(U)$ nonnegative, symmetric but not necessarily of finite trace. After the preparations of the previous section we are now able to define the stochastic integral with respect to a cylindrical Q-Wiener process $W(t)$, $t\in[0,T]$:
Basically we integrate with respect to the standard U_1-valued Q_1-Wiener process given by Proposition 1.38. In this sense we first get that a process $\Phi(t)$, $t\in[0,T]$, is integrable with respect to $W(t)$, $t\in[0,T]$, if it takes values in $L_2(Q_1^{\frac{1}{2}}(U_1),H)$, is predictable and if

$$P\Big(\int_0^T\|\Phi(s)\|^2_{L_2(Q_1^{\frac{1}{2}}(U_1),H)}\,ds<\infty\Big)=1.$$

But in addition, we have by Proposition 1.38 that $Q_1^{\frac{1}{2}}(U_1) = J(U_0)$ and that

$$\langle Ju_0, Jv_0\rangle_{Q_1^{\frac{1}{2}}(U_1)} = \langle Q_1^{-\frac{1}{2}} Ju_0, Q_1^{-\frac{1}{2}} Jv_0\rangle_1 = \langle u_0, v_0\rangle_0 = \langle Q^{-\frac{1}{2}} u_0, Q^{-\frac{1}{2}} v_0\rangle_U$$

for all $u_0, v_0 \in U_0$. That means especially that Je_k, $k \in \mathbb{N}$, is an orthonormal basis of $Q_1^{\frac{1}{2}}(U_1)$ if e_k, $k \in \mathbb{N}$, is an orthonormal basis of U_0 and in this way we get that

$$\Phi \in L_2^0 = L_2(Q^{\frac{1}{2}}(U), H) \Longleftrightarrow \Phi \circ J^{-1} \in L_2(Q_1^{\frac{1}{2}}(U_1), H)$$

since

$$\begin{aligned} \|\Phi\|_{L_2^0}^2 &= \sum_{k \in \mathbb{N}} \langle \Phi e_k, \Phi e_k\rangle \\ &= \sum_{k \in \mathbb{N}} \langle \Phi \circ J^{-1}(Je_k), \Phi \circ J^{-1}(Je_k)\rangle = \|\Phi \circ J^{-1}\|^2_{L_2(Q_1^{\frac{1}{2}}(U_1), H)} \end{aligned}$$

If we set now

$$\int_0^t \Phi(s)\, dW(s) := \int_0^t \Phi(s) \circ (J^{-1})_{|Q_1^{\frac{1}{2}}(U_1)}\, dW(s), \quad t \in [0, T],$$

the class of all integrable processes is given by

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

as in the case that $W(t)$, $t \in [0, T]$, is a standard Wiener process in U. Especially we notice that the space $\mathcal{N}_W(0, T)$ does not depend on the embedding $J : U_0 \to U_1$.

Remark 1.39. (i) If $Q \in L(U)$ is nonnegative, symmetric and with finite trace the standard Q-Wiener process can also be considered as a cylindrical Q-Wiener process by setting $J = I : U_0 \to U$ where I is the identity function. In this case both definitions of the stochastic integral correspond.
(ii) We are especially interested in the case where $Q = I : U \to U$. Then it is clear that Q is not of finite trace but if we introduce weights $\lambda_k > 0$, $k \in \mathbb{N}$, such that $\sum_{k \in \mathbb{N}} \lambda_k^2 < \infty$ and define

$$\begin{aligned} J : U_0 &= U \to U \\ u &= \sum_{k \in \mathbb{N}} \alpha_k e_k \mapsto \sum_{k \in \mathbb{N}} \lambda_k \alpha_k e_k \end{aligned}$$

where e_k, $k \in \mathbb{N}$, is an orthonormal basis of U, then J is in fact a Hilbert-Schmidt embedding. Thus we are in the setting of cylindrical Wiener processes and it is possible to define the stochastic integral for all processes

$$\{\Phi : \Omega_T \to L_2(U,H) \mid \Phi \text{ predictable and } P(\int_0^T \|\Phi(s)\|_{L_2}^2 \, ds < \infty) = 1\}$$

Chapter 2

Strong, Mild and Weak Solutions

As in the previous chapter let $(U, \| \|_U)$ and $(H, \| \|)$ be separable Hilbert spaces. We take $Q = I_U$ and fix a cylindrical Q-Wiener process $W(t)$, $t \geq 0$, in U on a probability space $(\Omega, \mathcal{F}, P)$ with a normal filtration $\mathcal{F}_t$, $t \geq 0$. Moreover we fix $T > 0$ and consider the following type of stochastic differential equations in H

$$(2.1) \qquad \begin{cases} dX(t) & = [AX(t) + F(X(t))]\, dt + B(X(t))\, dW(t), \quad t \in [0,T] \\ X(0) & = \xi \end{cases}$$

where

- $A : D(A) \to H$ is the infinitesimal generator of a C_0-semigroup $S(t)$, $t \geq 0$, of linear operators on H,
- $F : H \to H$ is $\mathcal{B}(H)/\mathcal{B}(H)$-measurable,
- $B : H \to L(U, H)$,
- ξ is a H-valued, $\mathcal{F}_0$-measurable random variable.

Definition 2.1 (Strong solution). A $D(A)$-valued predictable process $X(t)$, $t \in [0,T]$, (i. e. $(s, \omega) \mapsto X(s, \omega)$ is $\mathcal{P}_T/\mathcal{B}(H)$-measurable) is called a *strong solution* of problem (2.1) if

$$(2.2) \qquad X(t) = \xi + \int_0^t AX(s) + F(X(s))\, ds + \int_0^t B(X(s))\, dW(s) \quad P\text{-a.s.}$$

for each $t \in [0,T]$. In particular, the integrals on the right hand side have to be well defined, i.e. that $AX(t)$, $F(X(t))$, $t \in [0,T]$, are P-a.s. Bochner integrable and that $B(X(t))$, $t \in [0,T]$, is stochastically integrable.

Definition 2.2 (Mild solution). A H-valued predictable process $X(t)$, $t \in [0,T]$, is called a *mild solution* of problem (2.1) if

$$\begin{aligned} X(t) = S(t)\xi &+ \int_0^t S(t-s)F(X(s))\,ds \\ &+ \int_0^t S(t-s)B(X(s))\,dW(s) \quad P\text{-a.s.} \end{aligned} \tag{2.3}$$

for each $t \in [0,T]$. In particular, as in Definition 2.1, the appearing integrals have to be well defined.

Remark 2.3. While the strong solution has to take values in $D(A)$ the mild solution $X(t)$, $t \in [0,T]$, has only to be H-valued.
Besides, for the definition of a mild solution, it suffices that the process $1_{[0,t[}(s)S(t-s)B(X(s))$, $s \in [0,T]$, is $L_2(U,H)$-valued for each $t \in [0,T]$. For the definition of the strong solution we have to claim that the process $B(X(s))$, $s \in [0,T]$, itself is $L_2(U,H)$-valued.

Remark 2.4. Let $\Phi \in \mathcal{N}_W$. Then the process of the stochastic integrals $\int_0^t \Phi(s)\ dW(s)$, $t \in [0,T]$, is well defined outside one P-null set N independent of $t \in [0,T]$ so that, in this context, it is possible to consider the trajectories

$$\int_0^t \Phi(s)\ dW(s)(\omega),\ t \in [0,T],\ \text{for } \omega \in N^c.$$

If we consider different processes $\Phi_t(s)$, $s \in [0,T]$, $t \in [0,T]$, the stochastic integrals

$$\int_0^u \Phi_t(s)\ dW(s),\ u \in [0,T],$$

have well defined trajectories outside P-null sets $N(t)$, $t \in [0,T]$, but there is no pathwise definition of

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

outside one P-null set N independent of $t \in [0,T]$. Since

$$\int_0^t S(t-s)B(X(s))\ dW(s),\ t \in [0,T],$$

is a process of this type one has problems to analyze pathwise properties of mild solutions without introducing the following notion.

Definition 2.5. Let $X(t)$, $t \in [0,T]$, be a H-valued stochastic process on $(\Omega, \mathcal{F}, P)$. Another H-valued process $\tilde{X}(t)$, $t \in [0,T]$, is called *version* of X if

$$P(X(t) = \tilde{X}(t)) = 1 \qquad \text{for all } t \in [0,T].$$

Definition 2.6 (Weak solution). A H-valued predictable process $X(t)$, $t \in [0,T]$, is called a *weak solution* of problem (2.1) if

$$\begin{aligned} \langle X(t), \zeta \rangle = \langle \xi, \zeta \rangle &+ \int_0^t \langle X(s), A^*\zeta \rangle + \langle F(X(s)), \zeta \rangle \, ds \\ &+ \int_0^t \langle B(X(s)), \zeta \rangle \, dW(s) \quad P\text{-a.s.} \end{aligned} \tag{2.4}$$

for each $t \in [0,T]$ and $\zeta \in D(A^*)$.
In particular, as in Definition 2.1, the appearing integrals have to be well-defined.

Remark 2.7. Without any additional assumptions, the process $\langle B(X(t)), x \rangle$, $t \in [0,T]$, takes values in $L_2(U, \mathbb{R})$ for each $x \in H$ even if the original process $B(X(t))$, $t \in [0,T]$, is only $L(U,H)$-valued.

Proof. Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of the Hilbert space U and $L \in L(U,H)$. Then we get by the Parseval identity and the fact that $\|L\|_{L(U,H)} = \|L^*\|_{L(H,U)}$ that

$$\begin{aligned} \|\langle L, h \rangle\|^2_{L_2(U,\mathbb{R})} &= \sum_{k \in \mathbb{N}} \langle Le_k, h \rangle_H^2 = \sum_{k \in \mathbb{N}} \langle e_k, L^*h \rangle_U^2 = \|L^*h\|_U^2 \le \|L^*\|^2_{L(H,U)} \|h\|^2 \\ &= \|L\|^2_{L(U,H)} \|h\|^2 < \infty \end{aligned}$$

□

Lemma 2.8. *Let $X(t)$, $t \in [0,T]$, be a stochastic process with values in H.*

(i) If $X(t)$, $t \in [0,T]$, is $\mathcal{P}_T/\mathcal{B}(H)$-measurable and $D(A)$-valued. Then $AX(t)$, $t \in [0,T]$, is also $\mathcal{P}_T/\mathcal{B}(H)$-measurable.

(ii) If $B(X(t))$, $t \in [0,T]$, takes values in $L_2(U,H)$ and if $\langle B(X(t)), \zeta \rangle$, $t \in [0,T]$, is $\mathcal{P}_T/\mathcal{B}(L_2(U,\mathbb{R}))$-measurable for all $\zeta \in D(A^*)$ then $B(X(t))$, $t \in [0,T]$, itself is $L_2(U,H)$-predictable.

Proof. (i): Let $\zeta \in D(A^*)$. Then the mapping

$$(t,\omega) \mapsto \langle AX(t,\omega), \zeta\rangle = \langle X(t,\omega), A^*\zeta\rangle \text{ is } \mathcal{P}_T\text{-measurable.}$$

Since $D(A^*)$ is dense in H (see [Pa 83, Corollary 2.5, p.5; Lemma 10.5, p.40]) we get the assertion.
(ii): Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U and f_k, $k \in \mathbb{N}$, be an orthonormal basis of H. Then $f_k \otimes e_j = f_k\langle e_j, \cdot\rangle$ is an orthonormal basis of $L_2(U,H)$ (see Proposition B.7).
Since $D(A^*)$ is dense in H (see [Pa 83, Corollary 2.5, p.5; Lemma 10.5, p.40]) we get that $\langle B(X(t)), x\rangle$, $t \in [0,T]$, is $\mathcal{P}_T/\mathcal{B}(L_2(U,\mathbb{R}))$-measurable for each $x \in H$. Therefore we can conclude that

$$\langle B(X(t)), f_k \otimes e_j\rangle_{L_2} = \langle B(X(t))e_j, f_k\rangle, \quad t \in [0,T],$$

is $\mathcal{P}_T/\mathcal{B}(\mathbb{R})$-measurable for all $j,k \in \mathbb{N}$ and the assertion follows. □

Proposition 2.9. *(i) Every strong solution of problem (2.1) is also a weak solution.*

(ii) Let $X(t)$, $t \in [0,T]$, be a weak solution of problem (2.1) with values in $D(A)$ such that $B(X(t))$ takes values in $L_2(U,H)$ for all $t \in [0,T]$. Besides we assume that

$$P\Big(\int_0^T \|AX(t)\|\, dt < \infty\Big) = 1$$
$$P\Big(\int_0^T \|F(X(t))\|\, dt < \infty\Big) = 1$$
$$P\Big(\int_0^T \|B(X(t))\|_{L_2}^2\, dt < \infty\Big) = 1.$$

Then the process is also a strong solution.

Proof. (i): Follows immediately from Lemma A.6 and Lemma 1.31.
(ii): Let $X(t)$, $t \in [0,T]$, be a weak solution of problem (2.1) with the required properties.
Then we get by Lemma 2.8 and the assumptions that we can apply Lemma A.6 and Lemma 1.31 to get for all $t \in [0,T]$ and for all $\zeta \in D(A^*)$ that

$$\langle X(t),\zeta\rangle = \langle \xi,\zeta\rangle + \int_0^t \langle X(s), A^*\zeta\rangle + \langle F(X(s)),\zeta\rangle\, ds + \int_0^t \langle B(X(s)),\zeta\rangle\, dW(s)$$

$$= \langle \xi + \int_0^t AX(s) + F(X(s))\, ds + \int_0^t B(X(s))\, dW(s), \zeta\rangle \quad P\text{-a.s.}$$

Since $D(A^*)$ is a dense subset of H (see [Pa 83, Corollary 2.5, p.5; Lemma 10.5, p.40]) there exists for each $t \in [0,T]$ and each $x \in H$ a P-null set $N(t,x) \in \mathcal{F}$ such that the equation

$$\langle X(t), x\rangle = \langle \xi + \int_0^t AX(s) + F(X(s))\, ds + \int_0^t B(X(s))\, dW(s), x\rangle$$

holds outside $N(t,x)$.
As H is separable we can choose the P-null set $N(t,x) = N(t)$ independent of x for each $t \in [0,T]$ and therefore the assertion follows. □

Proposition 2.10. *(i) Let $X(t)$, $t \in [0,T]$, be a weak solution of problem (2.1) such that $B(X(t))$ takes values in $L_2(U,H)$ for all $t \in [0,T]$. Besides we assume that*

$$P(\int_0^T \|X(t)\|\, dt < \infty) = 1$$
$$P(\int_0^T \|F(X(t))\|\, dt < \infty) = 1$$
$$P(\int_0^T \|B(X(t))\|_{L_2}^2\, dt < \infty) = 1.$$

Then the process is also a mild solution.

(ii) Let $X(t)$, $t \in [0,T]$, be a mild solution of problem (2.1) such that the mappings

$$(t,\omega) \mapsto \int_0^t S(t-s)F(X(s,\omega))\, ds$$
$$(t,\omega) \mapsto \int_0^t S(t-s)B(X(s))\, dW(s)(\omega)$$

have predictable versions. In addition, we require that

$$P(\int_0^T \|F(X(t))\|\, dt < \infty) = 1$$
$$\int_0^T E(\int_0^t \|\langle S(t-s)B(X(s)), A^*\zeta\rangle\|_{L_2(U,\mathbb{R})}^2\, ds)\, dt < \infty$$

for all $\zeta \in D(A^*)$.
Then the process is also a weak solution.

For the proof we need some preparations. At first we introduce the following notations.
Let $(E, \| \,\|_E)$ be an arbitrary Banach space and $k \in \mathbb{N} \cup \{0\}$. If we define

$$C^k([0,T],E) := \{f : [0,T] \to E \mid f \text{ is } k\text{-times differentiable with continuous derivatives } f^{(j)},\ 0 \le j \le k\},$$

$$\|f\|_{C^k([0,T],E)} := \sup_{t\in[0,T]} \sum_{j=0}^{k} \|f^{(j)}(t)\|_E, \quad f \in C^k([0,T],E)$$

then $(C^k([0,T],E), \| \,\|_{C^k([0,T],E)})$ is a Banach space. If $E = \mathbb{R}$ we set $C^k([0,T],\mathbb{R}) = C^k([0,T])$.

Lemma 2.11. *The space* $\operatorname{span}(f\cdot x \mid f \in C^1([0,T]),\ x \in E)$ *is dense in* $(C^1([0,T],E), \| \,\|_{C^1([0,T],E)})$.

Proof. Claim 1: It is enough to show that $\operatorname{span}(f\cdot x \mid f \in C([0,T]), x \in E)$ is dense in $(C([0,T],E), \| \,\|_{C([0,T],E)})$:
We assume that $\operatorname{span}(f \cdot x \mid f \in C([0,T]),\ x \in E)$ is dense in $(C([0,T],E), \| \,\|_{C([0,T],E)})$ and take an arbitrary $g \in C^1([0,T],E)$.
Then $g' \in C([0,T],E)$. Therefore by assumption there exists a sequence $\tilde{g}_n := f_n x_n$, $n \in \mathbb{N}$, in $\operatorname{span}(f \cdot x \mid f \in C([0,T]),\ x \in E)$ such that $\|g' - \tilde{g}_n\|_{C([0,T],E)} \longrightarrow 0$ as $n \to \infty$.
If we define g_n, $n \in \mathbb{N}$, by $g_n(t) := x_n \displaystyle\int_0^t f_n(s)\,ds + g(0)$, $t \in [0,T]$, then g_n is an element of $\operatorname{span}(f \cdot x \mid f \in C^1([0,T]),\ x \in E)$ with $g_n' = \tilde{g}_n$ and by the fundamental Theorem for Bochner integrals A.7 and the Bochner inequality A.5 we get that

$$\begin{aligned}
&\|g - g_n\|_{C^1([0,T],E)} \\
&= \sup_{t\in[0,T]} (\|g(t) - g_n(t)\|_E + \|g'(t) - \tilde{g}_n(t)\|_E) \\
&= \sup_{t\in[0,T]} (\|\int_0^t g'(s) - \tilde{g}_n(s)\,ds\|_E + \|g'(t) - \tilde{g}_n(t)\|_E) \\
&\le T \sup_{t\in[0,T]} \|g'(t) - \tilde{g}_n(t)\|_E + \sup_{t\in[0,T]} \|g'(t) - \tilde{g}_n(t)\|_E \underset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

Claim 2: The space $\operatorname{span}(f \cdot x \mid f \in C([0,T]),\ x \in E)$ is dense in $(C([0,T],E), \| \,\|_{C([0,T],E)})$:
Let f be an element of $C([0,T],E)$. Then f is uniformly continuous, i.e. that for all $\varepsilon > 0$ there exists a $n = n(\varepsilon) \in \mathbb{N}$ such that $\|f(t) - f(s)\|_E \le \frac{\varepsilon}{2}$

for all $s, t \in [0,T]$ with $|s-t| \leq \frac{T}{n}$. If we set

$$f_n := f(0)1_{\{0\}} + \sum_{k=1}^{n} f(\frac{kT}{n}) 1_{]\frac{(k-1)T}{n}, \frac{kT}{n}]}$$

then

$$\sup_{t\in[0,T]} \|f_n(t) - f(t)\|_E \leq \frac{\varepsilon}{2}.$$

For $0 \leq m \leq n$ we define $f_{n,m} : [0,T] \to [0,1]$ by

$$f_{n,m}(t) := \begin{cases} 0 & \text{if } t \in [\frac{(m-1)T}{n}, \frac{(m+1)T}{n}]^c \cap [0,T] \\ \frac{n}{T}(t - \frac{(m-1)T}{n}) & \text{if } t \in [\frac{(m-1)T}{n}, \frac{mT}{n}[\cap [0,T] \\ -\frac{n}{T}(t - \frac{(m+1)T}{n}) & \text{if } t \in [\frac{mT}{n}, \frac{(m+1)T}{n}] \cap [0,T] \end{cases}$$

Then $f_{n,m} \in C([0,T], E)$ for all $0 \leq m \leq n$. Therefore

$$\tilde{f}_n := \sum_{m=0}^{n} f_{n,m} f(\frac{mT}{n}) \in \text{span}\{f \cdot x | f \in C([0,T]),\ x \in E\}$$

and the following holds:

1. $\|\tilde{f}_n(0) - f_n(0)\|_E = \|f_{n,0}(0) f(0) - f(0)\|_E = 0$
2. For $t \in]0,T]$ there exists $m \in \{1, ..., n\}$, such that $t \in]\frac{(m-1)T}{n}, \frac{mT}{n}]$.

   Then we can conclude that

$$\begin{aligned} &\|\tilde{f}_n(t) - f_n(t)\|_E \\ &= \|f_{n,m-1}(t) f(\frac{(m-1)T}{n}) + f_{n,m}(t) f(\frac{mT}{n}) - f(\frac{mT}{n})\|_E \\ &= \|f_{n,m-1}(t) f(\frac{(m-1)T}{n}) - f_{n,m-1}(t) f(\frac{mT}{n})\|_E \\ &\quad + \|f_{n,m-1}(t) f(\frac{mT}{n}) + f_{n,m}(t) f(\frac{mT}{n}) - f(\frac{mT}{n})\|_E \\ &= |f_{n,m-1}(t)| \|f(\frac{(m-1)T}{n}) - f(\frac{mT}{n})\|_E \\ &\quad + \|\underbrace{\Big(-\frac{n}{T}(t - \frac{mT}{n}) + \frac{n}{T}(t - \frac{(m-1)T}{n})\Big)}_{=1} f(\frac{mT}{n}) - f(\frac{mT}{n})\|_E \\ &\leq \|f(\frac{(m-1)T}{n}) - f(\frac{mT}{n})\|_E \leq \frac{\varepsilon}{2} \end{aligned}$$

So finally

$$\sup_{t\in[0,T]} \|\tilde{f}_n(t) - f(t)\|_E \leq \sup_{t\in[0,T]} \|\tilde{f}_n(t) - f_n(t)\|_E + \sup_{t\in[0,T]} \|f_n(t) - f(t)\|_E \leq \varepsilon$$

and the assertion is proved. □

The following Proposition can be proved in a similar way as [DaPrZa 92, Lemma 5.5, p.122] where $F = 0$ and $B(x) = B \in L_2(U,H)$ for all $x \in H$. The appearing space $D(A^*)$ is equipped with the graphnorm given by $\|x\|_{D(A^*)} := \|x\| + \|A^*x\|$ for all $x \in D(A^*)$.

Proposition 2.12. *Let $X(t)$, $t \in [0,T]$, be a weak solution of problem (2.1) such that $B(X(t))$ takes values in $L_2(U,H)$ for all $t \in [0,T]$. Besides we assume that*

$$P(\int_0^T \|X(t)\|\, dt < \infty) = 1$$
$$P(\int_0^T \|F(X(t))\|\, dt < \infty) = 1$$
$$P(\int_0^T \|B(X(t))\|_{L_2}^2\, dt < \infty) = 1.$$

Then for arbitrary $\zeta \in C^1([0,T], D(A^*))$ the following equation holds P-a.s.

$$\begin{aligned}\langle X(T), \zeta(T)\rangle &= \langle \xi, \zeta(0)\rangle \\ &\quad + \int_0^T \langle B(X(t)), \zeta(t)\rangle\, dW(t) \\ &\quad + \int_0^T \langle X(t), A^*\zeta(t) + \zeta'(t)\rangle + \langle F(X(t)), \zeta(t)\rangle\, dt\end{aligned} \tag{2.5}$$

Proof. Step 1: We first prove the assertion for functions ζ of the form $\zeta(t) = \eta f(t)$, $t \in [0,T]$, where $f \in C^1([0,T])$ and $\eta \in D(A^*)$.
To this end we set for $t \in [0,T]$

$$\begin{aligned}Y(t) := \langle X(t), \eta\rangle = \langle \xi, \eta\rangle &+ \int_0^t \langle B(X(s)), \eta\rangle\, dW(s) \\ &+ \int_0^t \langle X(s), A^*\eta\rangle + \langle F(X(s)), \eta\rangle\, ds\end{aligned}$$

Then $Y(0)$ is a $\mathbb{R}$-valued $\mathcal{F}_0$-measurable random variable and $\Phi(t) := \langle B(X(t)), \eta\rangle$, $t \in [0,T]$, is a $L_2(U,\mathbb{R})$-valued stochastically integrable process on $[0,T]$. Besides $\varphi(t) := \langle X(t), A^*\eta\rangle + \langle F(X(t)), \eta\rangle$, $t \in [0,T]$, is a $\mathbb{R}$-valued predictable process, as $X(t)$, $t \in [0,T]$, is predictable and $F : H \to H$ is $\mathcal{B}(H)/\mathcal{B}(H)$-measurable by assumption. In addition $\varphi(t)$, $t \in [0,T]$, is P-a.s. Bochner integrable.

We consider the mapping $\quad G{:}[0,T] \times \mathbb{R} \to \mathbb{R}$

$$(t,x) \mapsto f(t)x.$$

Then the partial derivatives $\frac{\partial G}{\partial t}(t,x) = f'(t)x$, $DG(t,x) = f(t)$, $D^2G(t,x) = 0$ exist for each $(t,x) \in [0,T] \times \mathbb{R}$ and they are uniformly continuous on bounded subsets of $[0,T] \times \mathbb{R}$. Therefore we can apply the Itô formula 1.33 and we get that

$$\begin{aligned}
&\langle X(T), \zeta(T)\rangle = G(T, Y(T)) \\
&= G(0, Y(0)) + \int_0^T DG(t, Y(t))\Phi(t)\, dW(t) \\
&\quad + \int_0^T \frac{\partial G}{\partial t}(t, Y(t)) + DG(t, Y(t))\varphi(t) + \frac{1}{2}\operatorname{tr}[D^2G(t, Y(t))\Phi(t)\Phi^*(t)]\, dt \\
&= \langle \xi, f(0)\eta\rangle + \int_0^T f(t)\langle B(X(t)), \eta\rangle\, dW(t) \\
&\quad + \int_0^T f'(t)\langle X(t), \eta\rangle + f(t)(\langle X(t), A^*\eta\rangle + \langle F(X(t)), \eta\rangle)\, dt \\
&= \langle \xi, \zeta(0)\rangle + \int_0^T \langle B(X(t)), \zeta(t)\rangle\, dW(t) \\
&\quad + \int_0^T \langle X(t), \zeta'(t)\rangle + \langle X(t), A^*\zeta(t)\rangle + \langle F(X(t)), \zeta(t)\rangle\, dt \quad P\text{-a.s.}
\end{aligned}$$

Step 2. Because of the linearity of the inner product, the integrals, the mapping $\zeta \mapsto \zeta'$ and A^*, all $\zeta \in \operatorname{span}(f \cdot \eta \mid f \in C^1([0,T]), \eta \in D(A^*))$ fulfill equation (2.5).

Step 3. Now let $\zeta \in C^1([0,T], D(A^*))$.

Since $D(A^*)$ is a Banach space (see [Pa 83, Corollary 2.5, p.5, Lemma 10.6, p.41]) with respect to the graphnorm we can apply Lemma 2.11 to get the existence of a sequence ζ_n, $n \in \mathbb{N}$, in $\operatorname{span}(f \cdot \eta \mid f \in C^1([0,T]), \eta \in D(A^*))$ such that $\|\zeta - \zeta_n\|_{C^1([0,T],D(A^*))} \longrightarrow 0$ as $n \to \infty$. Then:

1. By Lemma 1.32 we can find a subsequence n_k, $k \in \mathbb{N}$, such that

$$\int_0^T \langle B(X(t)), \zeta_{n_k}(t)\rangle\, dW(t) \underset{k\to\infty}{\longrightarrow} \int_0^T \langle B(X(t)), \zeta(t)\rangle\, dW(t) \quad P\text{-a.s.}$$

2. $$\int_0^T \langle X(t), \zeta_n'(t)\rangle + \langle X(t), A^*\zeta_n(t)\rangle + \langle F(X(t)), \zeta_n(t)\rangle\, dt$$
$$\underset{n\to\infty}{\longrightarrow} \int_0^T \langle X(t), \zeta'(t)\rangle + \langle X(t), A^*\zeta(t)\rangle + \langle F(X(t)), \zeta(t)\rangle\, dt \quad P\text{-a.s.}$$

because of the following estimation:
Let $\omega \in \Omega$ with $\int_0^T \|X(t,\omega)\|\, dt < \infty$ and $\int_0^T \|F(X(t,\omega))\|\, dt < \infty$, then

$$\begin{aligned}
&\Big|\int_0^T \langle X(t,\omega), \zeta_n'(t) - \zeta'(t)\rangle + \langle X(t,\omega), A^*(\zeta_n(t) - \zeta(t))\rangle \\
&\qquad + \langle F(X(t,\omega)), \zeta_n(t) - \zeta(t)\rangle\, dt\Big| \\
&\le \int_0^T |\langle X(t,\omega), \zeta_n'(t) - \zeta'(t)\rangle + \langle X(t,\omega), A^*(\zeta_n(t) - \zeta(t))\rangle \\
&\qquad + \langle F(X(t,\omega)), \zeta_n(t) - \zeta(t)\rangle|\, dt \\
&\le \int_0^T \|X(t,\omega)\|\big(\|\zeta_n'(t) - \zeta'(t)\| + \|A^*(\zeta_n(t) - \zeta(t))\|\big) \\
&\qquad + \|F(X(t,\omega))\|\|\zeta_n(t) - \zeta(t)\|\, dt \\
&\le \|\zeta_n - \zeta\|_{C^1([0,T],D(A^*))} \int_0^T \|X(t,\omega)\| + \|F(X(t,\omega))\|\, dt \underset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

Therefore by Step 2 we finally get the existence of a subsequence n_k, $k \in \mathbb{N}$, such that

$$\begin{aligned}
&\langle X(T), \zeta(T)\rangle = \lim_{k\to\infty} \langle X(T), \zeta_{n_k}(T)\rangle \\
&= \lim_{k\to\infty} \Big(\langle \xi, \zeta_{n_k}(0)\rangle + \int_0^T \langle B(X(t)), \zeta_{n_k}(t)\rangle\, dW(t) \\
&\qquad + \int_0^T \langle X(t), A^*\zeta_{n_k}(t)\rangle + \langle X(t), \zeta_{n_k}'(t)\rangle + \langle F(X(t)), \zeta_{n_k}(t)\rangle\, dt\Big) \\
&= \langle \xi, \zeta(0)\rangle + \int_0^T \langle B(X(t)), \zeta(t)\rangle\, dW(t) \\
&\quad + \int_0^T \langle X(t), A^*\zeta(t)\rangle + \langle X(t), \zeta'(t)\rangle + \langle F(X(t)), \zeta(t)\rangle\, dt
\end{aligned}$$

□

Now we want to prove Proposition 2.10. The original idea for the proof of (i) was to apply Proposition 2.12 to the functions given by $\zeta(s) = S^*(t-s)\eta$, $s \in [0,t]$, $\eta \in D(A^*)$ (see [DaPrZa 92, Proposition 6.3, p.153]). Since these functions theirselves are not in $C^1([0,t], D(A^*))$ but only in

$C([0,t], D(A^*)) \cap C^1([0,t], H)$ we have to introduce the resolvent of A^*.
The main part of the proof of (ii) can be found in [DaPrZa 92, Proposition 6.4 (i), p.154].

Proof of 2.10. Proof of (i): We fix $t \in [0,T]$. Since $\mathcal{P}_T \cap ([0,t] \times \Omega) = \mathcal{P}_t$ we have that $X(s)$, $s \in [0,t]$, is a weak solution of problem (2.1) on $[0,t]$.
As $S^*(t)$, $t \in [0,T]$, is a C_0-semigroup with generator A^* (see [Pa 83, Corollary 10.6, p.41]) there exist constants $\omega \geq 0$ and $M \geq 1$ such that $\|S^*(t)\|_{L(H)} \leq Me^{\omega t}$ for $0 \leq t \leq T$. By [Pa 83, Theorem 5.3, p.20] we get that the resolvent set $\rho(A^*)$ of A^* contains the ray $]\,\omega, \infty[$.
That means that we can define the resolvent $R^*_\alpha : H \to D(A^*)$, $\alpha > \omega$, by $R^*_\alpha x = [\alpha I - A^*]^{-1} x = \int_0^\infty e^{-\alpha t} S^*(t) x \, dt$ with $\|R^*_\alpha\|_{L(H)} \leq \frac{M}{\alpha - \omega}$ (see [Pa 83, Theorem 5.3, p.20]).
Then the functions $\zeta : [0,t] \to D(A^*)$ given by $\zeta(s) = S^*(t-s) n R^*_n \eta$, $n > \omega$ and $\eta \in D(A^*)$, are in $C^1([0,t], D(A^*))$ since $A^* S^*(t) n R^*_n \eta = S^*(t) A^* n R^*_n \eta$ $= S^*(t) n R^*_n A^* \eta$ (see [Pa 83, Theorem 2.4 (c), p.4] and the proof of [MaRö 92, Proposition 1.10, p.10]). Hence we can apply Proposition 2.12 and obtain that the following equation holds P-a.s.:

$$
\begin{aligned}
&\langle X(t), nR^*_n\eta\rangle = \langle X(t), S^*(t-t) nR^*_n \eta\rangle \\
&= \langle \xi, S^*(t-0) nR^*_n\eta\rangle + \int_0^t \langle B(X(s)), S^*(t-s) nR^*_n \eta\rangle \, dW(s) \\
&\quad + \int_0^t \langle X(s), A^* S^*(t-s) nR^*_n\eta\rangle + \langle X(s), -A^* S^*(t-s) nR^*_n \eta\rangle \, ds \\
&\quad + \int_0^t \langle F(X(s)), S^*(t-s) nR^*_n \eta\rangle \, ds \\
&= \langle S(t)\xi + \int_0^t S(t-s) B(X(s)) \, dW(s) + \int_0^t S(t-s) F(X(s)) \, ds, nR^*_n\eta\rangle
\end{aligned}
$$

Since for all $n > \omega$ and $\eta \in D(A^*)$

$$
\|nR^*_n\eta - \eta\| = \|R^*_n(n\eta - (nI - A^*)\eta)\| = \|R^*_n A^* \eta\| \leq \frac{M}{n-\omega} \|A^*\eta\| \underset{n\to\infty}{\longrightarrow} 0
$$

we get P-a.s. that

$$
\begin{aligned}
&\langle X(t), \eta\rangle \\
&= \langle S(t)\xi + \int_0^t S(t-s) B(X(s)) \, dW(s) + \int_0^t S(t-s) F(X(s)) \, ds, \eta\rangle
\end{aligned}
$$

for each $\eta \in D(A^*)$. As $D(A^*)$ is dense in H (see e. g. [Pa 83, Corollary 2.5, p.5, Lemma 10.5, p.40]) and H is separable the assertion follows.

Proof of (ii): By the notations

$$\int_0^t S(t-s)F(X(s))\,ds \quad \text{and} \quad \int_0^t S(t-s)B(X(s))\,dW(s), \quad t\in[0,T],$$

we understand predictable versions of the respective processes.
By the Cauchy-Schwarz inequality and the Bochner inequality we get for all $\zeta\in D(A^*)$ that

$$\begin{aligned}
&\int_0^T |\langle \int_0^t S(t-s)F(X(s))\,ds, A^*\zeta\rangle|\,dt \\
&\leq \|A^*\zeta\| \int_0^T \int_0^t \|S(t-s)F(X(s))\|\,ds\,dt \\
&\leq \|A^*\zeta\| M_T \int_0^T \int_0^t \|F(X(s))\|\,ds\,dt \\
&\leq \|A^*\zeta\| M_T T \int_0^T \|F(X(s))\|\,ds < \infty \quad P\text{-a.s.}.
\end{aligned}$$

Using Lemma 1.31, the Hölder inequality and the isometry for stochastic integrals we have in addition that

$$\begin{aligned}
&E(\int_0^T |\langle \int_0^t S(t-s)B(X(s))\,dW(s), A^*\zeta\rangle|\,dt) \\
&= E(\int_0^T |\int_0^t \langle S(t-s)B(X(s)), A^*\zeta\rangle\,dW(s)|\,dt) \\
&\leq T^{\frac{1}{2}}\Big(E(\int_0^T |\int_0^t \langle S(t-s)B(X(s)), A^*\zeta\rangle\,dW(s)|^2\,dt)\Big)^{\frac{1}{2}} \\
&= T^{\frac{1}{2}}\Big(\int_0^T E(|\int_0^t \langle S(t-s)B(X(s)), A^*\zeta\rangle\,dW(s)|^2)\,dt\Big)^{\frac{1}{2}} \\
&= T^{\frac{1}{2}}\Big(\int_0^T E(\int_0^t \|\langle S(t-s)B(X(s)), A^*\zeta\rangle\|^2_{L_2(U,\mathbb{R})}\,ds)\,dt\Big)^{\frac{1}{2}} < \infty
\end{aligned}$$

for all $\zeta\in D(A^*)$. Therefore the processes

$$\langle \int_0^t S(t-s)F(X(s))\,ds, A^*\zeta\rangle \quad \text{and} \quad \langle \int_0^t S(t-s)B(X(s))\,dW(s), A^*\zeta\rangle,$$

$t\in[0,T]$, are P-a.s. Bochner integrable and we obtain by the real Fubini Theorem that

$$E\Big(\Big|\int_0^t \langle X(s), A^*\zeta\rangle\,ds - \int_0^t \langle S(s)\xi, A^*\zeta\rangle\,ds$$

$$
\begin{aligned}
&- \int_0^t \langle \int_0^s S(s-u)F(X(u))\,du, A^*\zeta\rangle\,ds \\
&- \int_0^t \langle \int_0^s S(s-u)B(X(u))\,dW(u), A^*\zeta\rangle\,ds\Big|\Big) \\
\leq\ & E\Big(\int_0^t \big|\big\langle X(s) - S(s)\xi - \int_0^s S(s-u)F(X(u))\,du \\
&- \int_0^s S(s-u)B(X(u))\,dW(u), A^*\zeta\big\rangle\big|\,ds\Big) \\
=\ & \int_0^t E\Big(\big|\big\langle X(s) - S(s)\xi - \int_0^s S(s-u)F(X(u))\,du \\
&- \int_0^s S(s-u)B(X(u))\,dW(u), A^*\zeta\big\rangle\big|\Big)\,ds = 0
\end{aligned}
$$

since $X(t)$, $t \in [0,T]$, is a mild solution. Thus we get for all $\zeta \in D(A^*)$ and $t \in [0,T]$ that the following equation holds P-a.s.:

$$
\begin{aligned}
&\int_0^t \langle X(s), A^*\zeta\rangle\,ds \\
=\ & \int_0^t \langle S(s)\xi, A^*\zeta\rangle\,ds \\
&+ \int_0^t \langle \int_0^s S(s-u)F(X(u))\,du, A^*\zeta\rangle\,ds \\
&+ \int_0^t \langle \int_0^s S(s-u)B(X(u))\,dW(u), A^*\zeta\rangle\,ds
\end{aligned}
$$

By [Pa 83, Corollary 10.6, p.41] $S^*(t)$, $t \in [0,T]$, is a C_0-semigroup with infinitesimal generator A^*. Then by [Pa 83, Theorem 2.4 c, p.4] we get that $S^*(t)\zeta \in D(A^*)$ for all $t \in [0,T]$ and $\frac{d}{dt}S^*(t)\zeta = A^*S^*(t)\zeta = S^*(t)A^*\zeta$ for all $\zeta \in D(A^*)$.
That is why we can conclude by the fundamental Theorem for Bochner integrals A.7 that the following equations hold P-a.s.:

1.
$$
\begin{aligned}
&\int_0^t \langle S(s)\xi, A^*\zeta\rangle\,ds \\
=\ & \int_0^t \langle \xi, S^*(s)A^*\zeta\rangle\,ds \\
=\ & \int_0^t \langle \xi, \frac{d}{ds}S^*(s)\zeta\rangle\,ds \\
=\ & \langle \xi, S^*(t)\zeta - \zeta\rangle = \langle S(t)\xi, \zeta\rangle - \langle \xi, \zeta\rangle
\end{aligned}
$$

2.
$$\int_0^t \langle \int_0^s S(s-u)F(X(u))\,du, A^*\zeta\rangle\,ds$$
$$= \int_0^t \int_0^t 1_{[0,s]}(u)\langle F(X(u)), S^*(s-u)A^*\zeta\rangle\,du\,ds$$
(see Lemma A.6)
$$= \int_0^t \int_u^t \langle F(X(u)), \frac{d}{ds}S^*(s-u)\zeta\rangle\,ds\,du$$
(real Fubini Theorem)
$$= \int_0^t \langle F(X(u)), S^*(t-u)\zeta\rangle - \langle F(X(u)), \zeta\rangle\,du$$
$$= \langle \int_0^t S(t-s)F(X(s))\,ds, \zeta\rangle - \int_0^t \langle F(X(s)), \zeta\rangle\,ds$$
(see Lemma A.6)

3.
$$\int_0^t \langle \int_0^s S(s-u)B(X(u))\,dW(u), A^*\zeta\rangle\,ds$$
$$= \int_0^t \int_0^t 1_{[0,s]}(u)\langle B(X(u)), S^*(s-u)A^*\zeta\rangle\,dW(u)\,ds$$
(see Lemma 1.31)
$$= \int_0^t \int_u^t \langle B(X(u)), S^*(s-u)A^*\zeta\rangle\,ds\,dW(u)$$
(stochastic Fubini Theorem 1.37)
$$= \int_0^t \int_u^t \langle B(X(u)), \frac{d}{ds}S^*(s-u)\zeta\rangle\,ds\,dW(u)$$
$$= \int_0^t \langle B(X(u)), S^*(t-u)\zeta - \zeta\rangle\,dW(u)$$
$$= \langle \int_0^t S(t-s)B(X(s))\,dW(s), \zeta\rangle - \int_0^t \langle B(X(s)), \zeta\rangle\,dW(s)$$
(see Lemma 1.31)

In this way we have proved that the mild solution $X(t)$, $t \in [0,T]$, fulfills the following equation P-a.s..

$$\int_0^t \langle X(s), A^*\zeta\rangle\,ds$$
$$= \langle S(t)\xi, \zeta\rangle + \langle \int_0^t S(t-s)F(X(s))\,ds, \zeta\rangle + \langle \int_0^t S(t-s)B(X(s))\,dW(s), \zeta\rangle$$

$$
\begin{aligned}
&- \langle \xi, \zeta\rangle - \int_0^t \langle F(X(s)), \zeta\rangle \, ds - \int_0^t \langle B(X(s)), \zeta\rangle \, dW(s) \\
&= \langle X(t), \zeta\rangle - \langle \xi, \zeta\rangle - \int_0^t \langle F(X(s)), \zeta\rangle \, ds - \int_0^t \langle B(X(s)), \zeta\rangle \, dW(s)
\end{aligned}
$$

and therefore we finally get that for all $t \in [0, T]$ and $\zeta \in D(A^*)$

$$
\begin{aligned}
\langle X(t), \zeta\rangle = \langle \xi, \zeta\rangle &+ \int_0^t \langle X(s), A^*\zeta\rangle + \langle F(X(s)), \zeta\rangle \, ds \\
&+ \int_0^t \langle B(X(s)), \zeta\rangle \, dW(s) \quad P\text{-a.s.}
\end{aligned}
$$

□

Chapter 3

Existence, Uniqueness and Continuity of the Mild Solution

3.1 Existence, uniqueness and continuity with respect to the initial data

We consider the stochastic equation (2.1) of the previous chapter

$$\begin{cases} dX(t) & = [AX(t) + F(X(t))]\, dt + B(X(t))\, dW(t) \\ X(0) & = \xi \end{cases}$$

To get the existence of a mild solution on $[0, T]$ we make the following usual assumptions (see [DaPrZa 96, Hypothesis 5.1, p.65]).

Hypothesis H.0

- $A : D(A) \to H$ is the infinitesimal generator of a C_0-semigroup $S(t)$, $t \geq 0$.
- $F : H \to H$ is Lipschitz continuous, i.e. that there exists a constant $C > 0$ such that

  $$\|F(x) - F(y)\| \leq C\|x - y\| \quad \text{for all } x, y \in H.$$

- $B : H \to L(U, H)$ is strongly continuous, i.e. that the mapping

  $$x \mapsto B(x)u$$

  is continuous from H to H for each $u \in U$.

- For all $t \in]0,T]$ and $x \in H$ we have that

$$S(t)B(x) \in L_2(U,H).$$

- There is a square integrable mapping $K : [0,T] \to [0,\infty[$ such that

$$\|S(t)(B(x) - B(y))\|_{L_2} \leq K(t)\|x - y\|$$
and
$$\|S(t)B(x)\|_{L_2} \leq K(t)(1 + \|x\|)$$

for all $t \in]0,T]$ and $x, y \in H$.

Remark 3.1. (i) If we call $M_T := \sup_{t\in[0,T]} \|S(t)\|_{L(H)}$ then $M_T < \infty$.
(ii) To get the last assumption it is even enough to verify that there exists an $\varepsilon \in]0,T]$ such that the inequalities hold for $0 < t \leq \varepsilon$ and

$$\int_0^\varepsilon K^2(s)\, ds < \infty.$$

(iii) The Lipschitz constant of F in Hypothesis H.0 can be chosen in such a way that we also have

$$\|F(x)\| \leq C(1 + \|x\|) \quad \text{for all } x \in H.$$

Proof. (i): For example by [Pa 83, Theorem 2.2, p.4] there exist constants $\omega \geq 0$ and $M \geq 1$ such that

$$\|S(t)\|_{L(H)} \leq Me^{\omega t} \quad \text{for all } t \geq 0$$

(ii): Let $\tilde{K} : [0,\varepsilon] \to [0,\infty[$ be square integrable with

$$\|S(t)(B(x) - B(y))\|_{L_2} \leq \tilde{K}(t)\|x - y\|$$

and

$$\|S(t)B(x)\|_{L_2} \leq \tilde{K}(t)(1 + \|x\|)$$

for all $t \in]0,T]$ and $x, y \in H$. Then we choose $N \in \mathbb{N}$ such that $\frac{T}{N} \leq \varepsilon$ and set

$$K(t) := M_T\tilde{K}(\frac{t}{N}) \quad \text{for } t \in [0,T]$$

where $M_T = \sup_{t\in[0,T]} \|S(t)\|_{L(H)}$. Then it is clear that $K : [0,T] \to [0,\infty[$ is square integrable and for all $x, y \in H$, $t \in]0,T]$ we get by the semigroup property that

$$\begin{aligned}\|S(t)(B(x)-B(y))\|_{L_2} &= \|S(\frac{Nt-t}{N})S(\frac{t}{N})(B(x)-B(y))\|_{L_2}\\ &\le M_T\|S(\frac{t}{N})(B(x)-B(y))\|_{L_2}\\ &\le M_T\tilde{K}(\frac{t}{N})\|x-y\| = K(t)\|x-y\|\end{aligned}$$

and

$$\begin{aligned}\|S(t)B(x)\|_{L_2} &\le M_T\|S(\frac{t}{N})B(x)\|_{L_2}\\ &\le M_T\tilde{K}(\frac{t}{N})(1+\|x\|) = K(t)(1+\|x\|)\end{aligned}$$

(iii): For all $x \in H$ we have that

$$\begin{aligned}\|F(x)\| &\le \|F(x)-F(0)\| + \|F(0)\|\\ &\le C\|x\| + \|F(0)\|\\ &\le (C \vee \|F(0)\|)\,(1+\|x\|)\end{aligned}$$

and of course we still have that

$$\|F(x)-F(y)\| \le (C \vee \|F(0)\|)\,\|x-y\| \quad \text{for all } x,y \in H$$

$\square$

Now we introduce the spaces where we want to find the mild solution of the above problem:
For each $T > 0$ and $p \ge 2$ we define $\mathcal{H}^p(T,H)$ to be the space of all H-valued predictable processes Y with

$$\|Y\|_{\mathcal{H}^p} := \sup_{t\in[0,T]} \big(E(\|Y(t)\|^p)\big)^{\frac{1}{p}} < \infty$$

Then $(\mathcal{H}^p(T,H), \|\ \|_{\mathcal{H}^p})$ is a Banach space.
For technical reasons we also consider the norms $\|\ \|_{p,\lambda,T}$, $\lambda \ge 0$, on $\mathcal{H}^p(T,H)$ given by

$$\|Y\|_{p,\lambda,T} := \sup_{t\in[0,T]} e^{-\lambda t}\big(E(\|Y(t)\|^p)\big)^{\frac{1}{p}}$$

Then $\|\ \|_{\mathcal{H}^p} = \|\ \|_{p,0,T}$ and for all $\lambda > 0$, $Y \in \mathcal{H}^p(T,H)$ we get that

$$\|Y\|_{p,\lambda,T} \le \|Y\|_{\mathcal{H}^p} \le e^{\lambda T}\|Y\|_{p,\lambda,T}$$

which means that all norms $\| \; \|_{p,\lambda,T}$, $\lambda \geq 0$, are equivalent. For simplicity we introduce the following notations

$$\mathcal{H}^p(T,H) := (\mathcal{H}^p(T,H), \| \; \|_{\mathcal{H}^p})$$

and

$$\mathcal{H}^{p,\lambda}(T,H) := (\mathcal{H}^p(T,H), \| \; \|_{p,\lambda,T})$$

Theorem 3.2. *Under Hypothesis H.0 there exists a unique mild solution $X(\xi) \in \mathcal{H}^p(T,H)$ of problem (2.1) with initial condition*
$\xi \in L^p(\Omega, \mathcal{F}_0, P; H) =: L^p_0$.
In addition we even obtain that the mapping

$$\begin{aligned} X : L^p_0 &\to \mathcal{H}^p(T,H) \\ \xi &\mapsto X(\xi) \end{aligned}$$

is Lipschitz continuous with Lipschitz constant $L_{T,p}$.

Remark 3.3. The above result can be found in [DaPrZa 96, Theorem 5.3.1, p.66]. In order to make the Lipschitz property explicit we added part (ii) to the abstract implicit function Theorem D.1 which is the basis for the proof of Theorem 3.2.
Checking that we are in the setting of the abstract implicit function Theorem D.1 there appears especially the problem to verify that there is a predictable version of

$$\int_0^t S(t-s)B(Y(s))\, dW(s), \quad t \in [0,T],$$

for all $Y \in \mathcal{H}^p(T,H)$. In [DaPrZa 96, Proposition 6.2, p.153] this is solved in the case that $B(Y) \in \mathcal{N}_W$. As we do not demand that B itself takes values in $L_2(U,H)$ we have to modify the proof.

Remark 3.4. It follows from the Lipschitz continuity of X that there exists a constant $C_{T,p}$ independent of $\xi \in L^p_0$ such that

$$\|X(\xi)\|_{\mathcal{H}^p} \leq C_{T,p}(1 + \|\xi\|_{L^p})$$

Before giving the proof of the theorem we need the following lemmas.

Lemma 3.5. *If $Y : \Omega_T \to H$ is $\mathcal{P}_T$-measurable then the mapping*

$$\begin{aligned} \tilde{Y} : \quad & \Omega_T \to H \\ & (s,\omega) \mapsto 1_{[0,t[}(s) S(t-s) Y(s,\omega) \end{aligned}$$

is also $\mathcal{P}_T$-measurable for each fixed $t \in [0,T]$.

Proof. Step 1: We prove the assertion for simple processes Y given by

$$Y = \sum_{k=1}^{n} x_k 1_{A_k}$$

where $n \in \mathbb{N}$, $x_k \in H$, $1 \le k \le n$, and $A_k \in \mathcal{P}_T$, $1 \le k \le n$, is a disjoint covering of Ω_T. Then we get that

$$\begin{aligned} \tilde{Y}: \quad & \Omega_T \to H \\ (s,\omega) \mapsto & \, 1_{[0,t[}(s) S(t-s) Y(s,\omega) = 1_{[0,t[}(s) \sum_{k=1}^{n} S(t-s) x_k 1_{A_k}(s,\omega) \end{aligned}$$

is $\mathcal{P}_T$-measurable since for each $B \in \mathcal{B}(H)$

$$\tilde{Y}^{-1}(B) = \bigcup_{k=1}^{n} \underbrace{(\underbrace{\{s \in [0,T] | 1_{[0,t[}(s) S(t-s) x_k \in B\}}_{\in \mathcal{B}([0,T])} \times \Omega) \cap A_k}_{\in \mathcal{P}_T}$$

because of the strong continuity of the semigroup.
Step 2: We prove the assertion for an arbitrary predictable process Y.
If $Y : \Omega_T \to H$ is $\mathcal{P}_T$-measurable there exists a sequence Y_n, $n \in \mathbb{N}$, of simple predictable processes such that $Y_n(s,\omega) \longrightarrow Y(s,\omega)$ as $n \to \infty$ for all $(s,\omega) \in \Omega_T$ (see Lemma A.4). Since $S(t) \in L(H)$ for all $t \in [0,T]$ we obtain that

$$\tilde{Y}(s,\omega) := 1_{[0,t[}(s) S(t-s) Y(s,\omega) = \lim_{n \to \infty} \underbrace{1_{[0,t[}(s) S(t-s) Y_n(s,\omega)}_{=: \tilde{Y}_n(s,\omega)}$$

By Step 1 $\tilde{Y}_n$, $n \in \mathbb{N}$, are predictable and therefore Proposition A.3 implies that $\tilde{Y}$ is also predictable. □

Lemma 3.6. *If Y is a predictable H-valued process and Hypothesis H.0 is fulfilled then the mapping*

$$(s,\omega) \mapsto 1_{[0,t[}(s) S(t-s) B(Y(s,\omega))$$

is $\mathcal{P}_T/\mathcal{B}(L_2)$-measurable.

Proof. Let f_k, $k \in \mathbb{N}$, be an orthonormal basis of H and e_k, $k \in \mathbb{N}$, an orthonormal basis of U. Then $f_k \otimes e_j = f_k \langle e_j, \cdot \rangle_U$, $k, j \in \mathbb{N}$, is an orthonormal basis of $L_2(U,H)$ (see Proposition B.7). Because of the strong continuity of B we obtain that

$$(s,\omega) \mapsto B(Y(s,\omega))e_j$$

is predictable for all $j \in \mathbb{N}$. Hence the previous lemma provides that

$$\begin{aligned}(s,\omega) \mapsto &\langle f_k \otimes e_j, 1_{[0,t[}(s)S(t-s)B(Y(s,\omega))\rangle_{L_2} \\ &= \langle f_k, 1_{[0,t[}(s)S(t-s)B(Y(s,\omega))e_j\rangle\end{aligned}$$

is predictable for all $j,k \in \mathbb{N}$. This is enough to conclude that

$$(s,\omega) \mapsto 1_{[0,t[}(s)S(t-s)B(Y(s,\omega))$$

is predictable. □

Lemma 3.7. *If a mapping $g : \Omega_T \to \mathbb{R}$ is $\mathcal{P}_T/\mathcal{B}(\mathbb{R})$-measurable then the mapping*

$$\begin{aligned}\tilde{Y} : \Omega_T &\to \mathbb{R} \\ (s,\omega) &\mapsto 1_{[0,t[}(s)g(s,\omega)\end{aligned}$$

is $\mathcal{B}([0,T]) \otimes \mathcal{F}_t/\mathcal{B}(\mathbb{R})$-measurable for each $t \in [0,T]$.

Proof. We have to show that $([0,t[\times\Omega) \cap \mathcal{P}_T \subset \mathcal{B}([0,T]) \otimes \mathcal{F}_t$.
Let $t \in [0,T]$. If we set

$$\mathcal{A} := \{A \in \mathcal{P}_T \mid A \cap ([0,t[\times\Omega) \in \mathcal{B}([0,T]) \otimes \mathcal{F}_t\}$$

it is clear that $\mathcal{A}$ is σ-field which contains the predictable rectangles $]s,u] \times F_s$, $F_s \in \mathcal{F}_s$, $0 \le s \le u \le T$ and $\{0\} \times F_0$, $F_0 \in \mathcal{F}_0$. Therefore $\mathcal{A} = \mathcal{P}_T$. □

Lemma 3.8. *If a process Φ is adapted to $\mathcal{F}_t$, $t \in [0,T]$, and stochastically continuous with values in a Banach space E then there exists a predictable version of Φ.*

Proof. [DaPrZa 92, Proposition 3.6 (ii), p.76] □

Lemma 3.9. *Let Φ be a predictable H-valued process which is P-a.s. Bochner integrable. Then the process given by*

$$\int_0^t S(t-s)\Phi(s)\, ds, \quad t \in [0,T],$$

is P-a.s. continuous and adapted to $\mathcal{F}_t$, $t \in [0,T]$. This especially implies that it is predictable.

Proof. By Lemma 3.5 the inner process $1_{[0,t[}(s)S(t-s)\Phi(s)$, $s \in [0,T]$, is predictable and in addition $\|1_{[0,t[}(s)S(t-s)\Phi(s)\| \leq M_T\|\Phi(s)\|$, $s \in [0,T]$. Hence the integrals $\int_0^t S(t-s)\Phi(s)\,ds$, $t \in [0,T]$, are well defined.
First we want to prove the continuity. To this end let $0 \leq s \leq t \leq T$. Then we get that

$$\begin{aligned}
&\|\int_0^s S(s-u)\Phi(u)\,du - \int_0^t S(t-u)\Phi(u)\,du\| \\
&\leq \|\int_0^s [S(s-u)-S(t-u)]\Phi(u)\,du\| + \|\int_s^t S(t-u)\Phi(u)\,du\| \\
&\leq \int_0^s \|[S(s-u)-S(t-u)]\Phi(u)\|\,du + \int_s^t \|S(t-u)\Phi(u)\|\,du
\end{aligned}$$

where the first summand converges to zero as $s \uparrow t$ or $t \downarrow s$ because of Lebesgue's dominated convergence theorem:

$$\|1_{[0,s[}(u)[S(s-u)-S(t-u)]\Phi(u)\| \to 0 \quad \text{as } s \uparrow t \text{ or } t \downarrow s$$

for all $u \in [0,T]$ because of the strong continuity of the semigroup $S(u)$, $u \in [0,T]$. Moreover

$$\begin{aligned}
&\|1_{[0,s[}(u)[S(s-u)-S(t-u)]\Phi(u)\| \\
&\leq 1_{[0,s[}(u)(\|S(s-u)\|_{L(H)} + \|S(t-u)\|_{L(H)})\|\Phi(u)\| \\
&\leq 2M_T\|\Phi(u)\|
\end{aligned}$$

where $\|\Phi\| \in L^1([0,T],dx) := L^1([0,T],\mathcal{B}([0,T]),dx;\mathbb{R})$ P-a.s..
Concerning the second summand we get the same result since

$$\begin{aligned}
&\int_s^t \|S(t-u)\Phi(u)\|\,du \\
&\leq \int_s^t M_T\|\Phi(u)\|\,du \longrightarrow 0 \quad \text{as } s \uparrow t \text{ or } t \downarrow s
\end{aligned}$$

P-a.s. by Lebesgues's dominated convergence theorem.
In order to prove that the process of the integrals is adapted we fix $t \in [0,T]$. By Lemma 3.7 the mapping

$$(s,\omega) \mapsto 1_{[0,t[}(s)S(t-s)\Phi(s,\omega)$$

is $\mathcal{B}([0,T]) \otimes \mathcal{F}_t$-measurable. Hence, by Proposition A.6, we get for each $x \in H$ that the mapping

$$\begin{aligned}\omega \mapsto & \langle \int_0^t S(t-s)\Phi(s,\omega)\ ds, x\rangle \\ = & \int_0^t \langle S(t-s)\Phi(s,\omega), x\rangle\ ds \\ = & \int_0^T \langle 1_{[0,t[}(s) S(t-s)\Phi(s,\omega), x\rangle\ ds\end{aligned}$$

is $\mathcal{F}_t$-measurable by the real Fubini theorem and therefore the integral itself is $\mathcal{F}_t$-measurable. □

Lemma 3.10. *Let $(x_{n,m})_{m\in\mathbb{N}}$, $n \in \mathbb{N}$, be sequences of real numbers such that for each $n \in \mathbb{N}$ there exists $x_n \in \mathbb{R}$ with*

$$x_{n,m} \longrightarrow x_n \quad \textit{as } m \to \infty.$$

If there exists a further sequence y_n, $n \in \mathbb{N}$, such that $|x_{n,m}| \leq y_n$ for all $m \in \mathbb{N}$ and $\sum_{n\in\mathbb{N}} y_n < \infty$ then

$$\sum_{n\in\mathbb{N}} x_{n,m} \longrightarrow \sum_{n\in\mathbb{N}} x_n \quad \textit{as } m \to \infty.$$

Proof. The assertion is a simple consequence of Lebesgue's dominated convergence theorem with respect to the measure $\mu := \sum_{n\in\mathbb{N}} \delta_n$ where δ_n is the Dirac measure in n. □

Proof of Theorem 3.2:
Idea: Let $p \geq 2$. For $t \in [0,T]$, $\xi \in L_0^p$ and $Y \in \mathcal{H}^p(T,H)$ we define

$$\mathcal{F}(\xi, Y)(t) := S(t)\xi + \int_0^t S(t-s)F(Y(s))\ ds + \int_0^t S(t-s)B(Y(s))\ dW(s)$$

and prove that

$$\mathcal{F} : L_0^p \times \mathcal{H}^p(T,H) \to \mathcal{H}^p(T,H)$$

Since $X(\xi) \in \mathcal{H}^p(T,H)$ is a mild solution of problem (2.1) if and only if $\mathcal{F}(\xi, X(\xi)) = X(\xi)$ we have to search for an implicit function $X : L_0^p \to \mathcal{H}^p(T,H)$ such that the previous equation holds for arbitrary $\xi \in L_0^p$. For this end we show that there exists a $\lambda = \lambda(p) \geq$ such that

$$\mathcal{F} : L_0^p \times \mathcal{H}^{p,\lambda,T}(T,H) \to \mathcal{H}^{p,\lambda,T}(T,H)$$

is a contraction with respect to the second variable, i.e. that there exists $\alpha(p) < 1$ such that for all $\xi \in L_0^p$ and Y, $\tilde{Y} \in \mathcal{H}^p(T,H)$

$$\|\mathcal{F}(\xi, Y) - \mathcal{F}(\xi, \tilde{Y})\|_{p,\lambda,T} \leq \alpha(p)\|Y - \tilde{Y}\|_{p,\lambda,T}.$$

Setting $G := \mathcal{F}$, $\Lambda := L_0^p$ and $E := \mathcal{H}^p(T,H)$ we are hence in the situation described at the beginning of Appendix D. Therefore it is clear that the implicit function $X = \varphi$ exists and that it is unique.
To get the Lipschitz continuity of the mapping $X : L_0^p \to \mathcal{H}^p(T,H)$ we verify that the condition of Theorem D.1 (ii) is fulfilled. Because of the equivalence of the norms $\| \; \|_{p,\lambda,T}$ and $\| \; \|_{\mathcal{H}^p}$ that means that we check the existence of a constant $L_{T,p} > 0$ such that

$$\|\mathcal{F}(\xi, Y) - \mathcal{F}(\tilde{\xi}, Y)\|_{\mathcal{H}^p} \leq L_{T,p}\|\xi - \tilde{\xi}\|_{L^p}$$

for all ξ, $\tilde{\xi} \in L_0^p$ and $Y \in \mathcal{H}^p(T,H)$.

Step 1: We prove that the mapping $\mathcal{F}$ is well defined.
Let $\xi \in L_0^p$ and $Y \in \mathcal{H}^p(T,H)$.
1. The Bochner integral $\int_0^t S(t-s)F(Y(s))\,ds$, $t \in [0,T]$, is well defined by Lemma 3.9:
(i) Because of the continuity of $F : H \to H$ it is clear that $F(Y(t))$, $t \in [0,T]$, is predictable.
(ii) In addition the process $F(Y(t))$, $t \in [0,T]$, is P-a.s. Bochner integrable since

$$E\Big(\int_0^t \|F(Y(s))\|\,ds\Big) \leq \int_0^t E(C(1 + \|Y(s)\|))\,ds \leq CT(1 + \|Y\|_{\mathcal{H}^p}) < \infty$$

2. The stochastic integrals $\int_0^t S(t-s)B(Y(s))\,dW(s)$, $t \in [0,T]$, are well defined since the processes $1_{[0,t[}(s)S(t-s)B(Y(s))$, $s \in [0,T]$, are in $\mathcal{N}_W^2(0,T)$ for all $t \in [0,T]$:
(i) The mapping

$$(s,\omega) \mapsto 1_{[0,t[}(s)S(t-s)B(Y(s,\omega))$$

is $\mathcal{P}_T/\mathcal{B}(L_2(U,H))$-measurable by Lemma 3.6.
(ii) With respect to the norm we obtain that

$$\begin{aligned}
\|1_{[0,t[}S(t-\cdot)B(Y)\|_T^2 &= E\Big(\int_0^t \|S(t-s)B(Y(s))\|_{L_2}^2\,ds\Big) \\
&= \int_0^t E(\|S(t-s)B(Y(s))\|_{L_2}^2)\,ds \\
&\leq \int_0^t K^2(t-s)E\big((1 + \|Y(s)\|)^2\big)\,ds
\end{aligned}$$

$$\leq 2\int_0^t K^2(t-s)E(1+\|Y(s)\|^2)\,ds$$
$$\leq 2\,(1+\sup_{s\in[0,T]} E(\|Y(s)\|^2)\int_0^t K^2(t-s)\,ds$$
$$\leq 2\,(1+\|Y\|^2_{\mathcal{H}^p})\int_0^t K^2(s)\,ds<\infty$$

Step 2: We prove that $\mathcal{F}(\xi,Y)\in\mathcal{H}^p(T,H)$ for all $\xi\in L_0^p$ and $Y\in\mathcal{H}^p(T,H)$. Let $\xi\in L_0^p$ and $Y\in\mathcal{H}^p(T,H)$.

1. The first summand $S(t)\xi$, $t\in[0,T]$, is an element of $\mathcal{H}^p(T,H)$:

(i) The mapping

$$(s,\omega)\mapsto S(t)\xi(\omega)$$

is predictable since for fixed $\omega\in\Omega$

$$t\mapsto S(t)\xi(\omega)$$

is a continuous mapping from $[0,T]$ to H and for fixed $t\in[0,T]$

$$\omega\mapsto S(t)\xi(\omega)$$

is not only $\mathcal{F}_t$- but even $\mathcal{F}_0$-measurable.

(ii) With respect to the norm we obtain that

$$\|S(\cdot)\xi\|_{\mathcal{H}^p}=\sup_{t\in[0,T]}(E(\|S(t)\xi\|^p))^{\frac{1}{p}}\leq M_T\|\xi\|_{L^p}<\infty$$

2. There is a version of the second summand $\int_0^t S(t-s)F(Y(s))\,ds$, $t\in[0,T]$, which is an element of $\mathcal{H}^p(T,H)$:

(i) First we notice that the mapping

$$(s,\omega)\mapsto\int_0^t S(t-s)F(Y(s,\omega))\,ds$$

has a predictable version because the assumptions of Lemma 3.9 are fulfilled (see Step 1, 1.).

(ii) Concerning the norm we prove that

$$\|\int_0^{\cdot} S(\cdot - s)F(Y(s))\, ds\|_{\mathcal{H}^p} \le CTM_T(1 + \|Y\|_{\mathcal{H}^p}).$$

To verify the assertion we take t from $[0,T]$ and show that the L^p-norm of $\|\int_0^t S(t-s)F(Y(s))\, ds\|$ can be estimated independently of $t \in [0,T]$:

$$\|\int_0^t S(t-s)F(Y(s))\, ds\|^p \le C^p T^{p-1} M_T^p \int_0^t (1 + \|Y(s)\|)^p\, ds \quad P\text{-a.s.}.$$

Taking the expectation we get that

$$\begin{aligned}
&\big(E(\|\int_0^t S(t-s)F(Y(s))\, ds\|^p)\big)^{\frac{1}{p}}\\
&\le CT^{\frac{p-1}{p}} M_T \big(E(\int_0^t (1 + \|Y(s)\|)^p\, ds\big)^{\frac{1}{p}}\\
&\le CT^{\frac{p-1}{p}} M_T \big[\big(E(\int_0^T 1\, ds)\big)^{\frac{1}{p}} + \big(\int_0^T E(\|Y(s)\|^p)\, ds\big)^{\frac{1}{p}}\big]\\
&\le CTM_T(1 + \|Y\|_{\mathcal{H}^p}) < \infty
\end{aligned}$$

In this way the inequality is proved.

3. There is a version of $\int_0^t S(t-s)B(Y(s))\, dW(s)$, $t \in [0,T]$, which is in $\mathcal{H}^p(T,H)$:

(i) First we show that there is a predictable version of the process. To do so we proceed in several steps.

Claim 1: If $\alpha > 1$ the process $\int_0^{\frac{t}{\alpha}} S(t-s)B(Y(s))\, dW(s)$, $t \in [0,T]$, has a predictable version.

To prove this we first use the semigroup property and get that

$$\begin{aligned}
&\int_0^{\frac{t}{\alpha}} S(t-s)B(Y(s))\, dW(s)\\
&= \int_0^{\frac{t}{\alpha}} S(t-\alpha s)S((\alpha-1)s)B(Y(s))\, dW(s), \quad t \in [0,T],
\end{aligned}$$

where we set

$$1_{]0,T]}(s)S((\alpha-1)s)B(Y(s)) =: \Phi^\alpha(s)$$

Then it is clear that $\Phi^\alpha(t)$, $t \in [0,T]$, is an element of $\mathcal{N}_W^2(0,T)$:
The fact that there is a predictable version of

$$(s,\omega) \mapsto 1_{]0,T]}(s) S((\alpha-1)s) B(Y(s,\omega))$$

can be proved in the same way as Lemma 3.6 and of course

$$\begin{aligned}
&E\Big(\int_0^T \|S((\alpha-1)s)B(Y(s))\|_{L_2}^2 \, ds\Big) \\
&\leq \int_0^T E\big(K^2((\alpha-1)s)(1+\|Y(s)\|)^2\big) \, ds \\
&\leq 2\int_0^T K^2((\alpha-1)s)(1+\sup_{s\in[0,T]} E(\|Y(s)\|^2)) \, ds \\
&= 2(1+\|Y\|_{\mathcal{H}^2(T,H)}^2) \int_0^{(\alpha-1)T} \frac{1}{\alpha-1} K^2(s) \, ds < \infty
\end{aligned}$$

Therefore we have to prove now that the process

$$\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s) \, dW(s), \quad t \in [0,T],$$

has a predictable version for each $\alpha > 1$ and $\tilde{\Phi} \in \mathcal{N}_W^2(0,T)$.

(a) We first consider the case that $\tilde{\Phi}(t)$, $t \in [0,T]$, is a simple process of the form

$$\tilde{\Phi} = \sum_{k=1}^m L_k 1_{A_k}$$

where $m \in \mathbb{N}$, $L_k \in L_2(U,H)$ and $A_k \in \mathcal{P}_T$, $1 \leq k \leq m$. To get the required measurability we check the conditions of Lemma 3.8.
At first it is clear that

$$\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s) \, dW(s)$$

is $\mathcal{F}_{\frac{t}{\alpha}}$- and therefore also $\mathcal{F}_t$-measurable for each $t \in [0,T]$ since the process

$$1_{[0,\frac{t}{\alpha}[}(s) S(t-\alpha s)\tilde{\Phi}(s), \quad s \in [0,T],$$

lies in $\mathcal{N}_W^2(0,T)$ (see the proof of Lemma 3.6) and therefore the process

$$\int_0^u 1_{[0,\frac{t}{\alpha}[}(s) S(t-\alpha s)\tilde{\Phi}(s) \, dW(s), \quad u \in [0,T],$$

is a H-valued martingale with respect to $\mathcal{F}_u$, $u \in [0,T]$. Besides we show now that

$$t \mapsto \int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s)$$

is continuous in the mean square and therefore stochastically continuous. To this end we take arbitrary $0 \le t < u \le T$ and get that

$$\begin{aligned}
&\Big(E(\|\int_0^{\frac{u}{\alpha}} S(u-\alpha s)\tilde{\Phi}(s)\, dW(s) - \int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s)\|^2)\Big)^{\frac{1}{2}} \\
&\le \Big(E(\|\int_0^{\frac{t}{\alpha}} [S(u-\alpha s) - S(t-\alpha s)]\tilde{\Phi}(s)\, dW(s)\|^2)\Big)^{\frac{1}{2}} \\
&\quad + \Big(E(\|\int_0^{\frac{u}{\alpha}} 1_{[\frac{t}{\alpha},\frac{u}{\alpha}[}(s) S(u-\alpha s)\tilde{\Phi}(s)\, dW(s)\|^2)\Big)^{\frac{1}{2}} \\
&= \Big(E(\int_0^{\frac{t}{\alpha}} \|[S(u-\alpha s) - S(t-\alpha s)]\tilde{\Phi}(s)\|_{L_2}^2\, ds)\Big)^{\frac{1}{2}} \\
&\quad + \Big(E(\int_{\frac{t}{\alpha}}^{\frac{u}{\alpha}} \|S(u-\alpha s)\tilde{\Phi}(s)\|_{L_2}^2\, ds)\Big)^{\frac{1}{2}} \\
&\le \sum_{k=1}^m \Big(E(\int_0^{\frac{t}{\alpha}} 1_{A_k}(s,\cdot)\|[S(u-\alpha s) - S(t-\alpha s)]L_k\|_{L_2}^2\, ds)\Big)^{\frac{1}{2}} \\
&\quad + \sum_{k=1}^m \Big(E(\int_{\frac{t}{\alpha}}^{\frac{u}{\alpha}} 1_{A_k}(s,\cdot)\|S(u-\alpha s)L_k\|_{L_2}^2\, ds)\Big)^{\frac{1}{2}}
\end{aligned}$$

$$\le \sum_{k=1}^m \Big(\int_0^{\frac{t}{\alpha}} \|[S(u-\alpha s) - S(t-\alpha s)]L_k\|_{L_2}^2\, ds\Big)^{\frac{1}{2}} \tag{3.1}$$

$$+ \sum_{k=1}^m \Big(\int_{\frac{t}{\alpha}}^{\frac{u}{\alpha}} \|S(u-\alpha s)L_k\|_{L_2}^2\, ds\Big)^{\frac{1}{2}} \tag{3.2}$$

where the first summand (3.1) converges to zero as $t \uparrow u$ or $u \downarrow t$ for the following reason:
Let e_n, $n \in \mathbb{N}$, be an orthonormal basis of U. Then we get for each $s \in [0,T]$ and $1 \le k \le m$ that

$$\begin{aligned}
&1_{[0,\frac{t}{\alpha}[}(s)\|[S(u-\alpha s) - S(t-\alpha s)]L_k\|_{L_2}^2 \\
&= \sum_{n\in\mathbb{N}} 1_{[0,\frac{t}{\alpha}[}(s)\|[S(u-\alpha s) - S(t-\alpha s)]L_k e_n\|^2
\end{aligned}$$

where

$$1_{[0,\frac{t}{\alpha}[}(s)\|[S(u-\alpha s) - S(t-\alpha s)]L_k e_n\|^2 \longrightarrow 0 \quad \text{as } t \uparrow u \text{ or } u \downarrow t$$

and at the same time

$$1_{[0,\frac{t}{\alpha}[}(s)\|[S(u-\alpha s)-S(t-\alpha s)]L_k e_n\|^2 \le 4M_T^2\|L_k e_n\|^2$$

for all $n \in \mathbb{N}$, $1 \le k \le m$. By Lemma 3.10 this result implies the pointwise convergence

$$1_{[0,\frac{t}{\alpha}[}(s)\|[S(u-\alpha s)-S(t-\alpha s)]L_k\|_{L_2}^2 \longrightarrow 0 \quad \text{as } t \uparrow u \text{ or } u \downarrow t.$$

Since there is the following upper bound

$$1_{[0,\frac{t}{\alpha}[}(s)\|[S(u-\alpha s)-S(t-\alpha s)]L_k\|_{L_2}^2 \le 4M_T^2\|L_k\|_{L_2}^2 \in L^1([0,T],dx)$$

for all $s \in [0,T]$, $0 \le t < u \le T$, we get the required convergence of the integrals $\int_0^{\frac{t}{\alpha}} \|[S(u-\alpha s)-S(t-\alpha s)]L_k\|_{L_2}^2\, ds$, $1 \le k \le m$, by Lebesgue's dominated convergence theorem.

The second summand (3.2) of the above equation also converges to zero since for each $1 \le k \le m$

$$\int_{\frac{t}{\alpha}}^{\frac{u}{\alpha}} \|S(u-\alpha s)L_k\|_{L_2}^2\, ds \le \frac{u-t}{\alpha} M_T^2 \|L_k\|_{L_2}^2 \longrightarrow 0 \quad \text{as } u \downarrow t \text{ or } t \uparrow u$$

Hence Lemma 3.8 implies that there is a predictable version of

$$\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s), \quad t \in [0,T],$$

if $\tilde{\Phi}$ is elementary.

(b) Now we generalize this result to arbitrary $\tilde{\Phi} \in \mathcal{N}_W^2$:

If $\tilde{\Phi}$ is an arbitrary process in $\mathcal{N}_W^2(0,T)$ there exists a sequence $\tilde{\Phi}_n$, $n \in \mathbb{N}$ of simple processes of the form we considered in (a) such that

$$E\left(\int_0^T \|\tilde{\Phi}(s)-\tilde{\Phi}_n(s)\|_{L_2}^2\, ds\right) \underset{n\to\infty}{\longrightarrow} 0$$

(see Lemma A.4). Hence let $\Psi_n(t)$, $t \in [0,T]$, $n \in \mathbb{N}$, be a predictable version of $\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}_n(s)\, dW(s)$, $t \in [0,T]$, $n \in \mathbb{N}$, which exists by step (a). To get the measurability of $\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\Phi(s)\, dW(s)$ we prove that there is a subsequence n_k, $k \in \mathbb{N}$, such that

$$\Psi_{n_k}(t) \underset{k\to\infty}{\longrightarrow} \int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s) \quad P\text{-a.s. for all } t \in [0,T].$$

To this end we take $c > 0$, $t \in [0,T]$ and obtain that

$$\begin{aligned}
&P(\|\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s) - \Psi_n(t)\| > c) \\
&\le \frac{1}{c^2} E(\|\int_0^{\frac{t}{\alpha}} S(t-\alpha s)[\tilde{\Phi}(s) - \tilde{\Phi}_n(s)]\, dW(s)\|^2) \\
&= \frac{1}{c^2} E(\int_0^{\frac{t}{\alpha}} \|S(t-\alpha s)[\tilde{\Phi}(s) - \tilde{\Phi}_n(s)]\|_{L_2}^2\, ds) \\
&\le \frac{M_T^2}{c^2} E(\int_0^{\frac{t}{\alpha}} \|\tilde{\Phi}(s) - \tilde{\Phi}_n(s)\|_{L_2}^2\, ds) \le \frac{M_T^2}{c^2} E(\int_0^T \|\tilde{\Phi}(s) - \tilde{\Phi}_n(s)\|_{L_2}^2\, ds)
\end{aligned}$$

As this upper bound is independent of $t \in [0,T]$ this implies that

$$\begin{aligned}
&\sup_{t\in[0,T]} P(\|\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s) - \Psi_n(t)\| > c) \\
&\le \frac{M_T^2}{c^2} E(\int_0^T \|\tilde{\Phi}(s) - \tilde{\Phi}_n(s)\|_{L_2}^2\, ds) \underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Therefore we get that there is one sequence n_k, $k \in \mathbb{N}$, such that

$$P(\|\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s) - \Psi_{n_k}\| > 2^{-k}) \le 2^{-k}$$

for all $t \in [0,T]$, $k \in \mathbb{N}$. By the lemma of Borel-Cantelli it follows that

$$\Psi_{n_k}(t) \longrightarrow \int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s) \quad P\text{-a.s. as } k \to \infty$$

for all $t \in [0,T]$. If we set now

$$A := \{(t,\omega) \in \Omega_T \mid (\Psi_{n_k}(t,\omega))_{k\in\mathbb{N}} \text{ is convergent in } H\}$$

then $A \in \mathcal{P}_T$ and the process defined by

$$\Psi(t,\omega) := \begin{cases} \lim_{k\to\infty} \Psi_{n_k}(t,\omega) & \text{if } (t,\omega) \in A \\ 0 & \text{otherwise} \end{cases}$$

is a predictable version of $\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\tilde{\Phi}(s)\, dW(s)$, $t \in [0,T]$.
Taking $\tilde{\Phi} = \Phi^\alpha$ we hence obtain that

$$\int_0^{\frac{t}{\alpha}} S(t-\alpha s)\Phi^\alpha(s)\, dW(s) = \int_0^{\frac{t}{\alpha}} S(t-s)B(Y(s))\, dW(s), \quad t \in [0,T],$$

has a predictable version. By the help of this result we can prove the assertion we are interested in.

Claim 2: The process $\int_0^t S(t-s)B(Y(s))\,dW(s),\ t\in[0,T]$, has a predictable version.
Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers such that $\alpha_n \downarrow 1$ as $n\to\infty$. By Claim 1 there is a predictable version $\Psi_{\alpha_n}(t)$, $t\in[0,T]$, of $\int_0^{\frac{t}{\alpha_n}} S(t-s)B(Y(s))\,dW(s)$, $t\in[0,T]$, $n\in\mathbb{N}$. If we define

$$B := \{(t,\omega)\in\Omega_T \mid (\Psi_{\alpha_n}(t,\omega))_{n\in\mathbb{N}} \text{ is convergent}\}$$

it is clear that $B\in\mathcal{P}_T$ and the process given by

$$\Psi(t,\omega) := \begin{cases} \lim_{n\to\infty}\Psi_{\alpha_n}(t,\omega) & \text{if } (t,\omega)\in B \\ 0 & \text{otherwise}\end{cases}$$

is predictable. Besides we get that for each $t\in[0,T]$

$$\Psi(t) = \int_0^t S(t-s)B(Y(s))\,dW(s) \quad P\text{-a.s.}$$

since

$$\Psi_{\alpha_n}(t) = \int_0^{\frac{t}{\alpha_n}} S(t-s)B(Y(s))\,dW(s) \underset{n\to\infty}{\longrightarrow} \int_0^t S(t-s)B(Y(s))\,dW(s)$$

P-a.s. because of the continuity of the stochastic integrals

$$\int_0^u 1_{[0,t[}(s)S(t-s)B(Y(s))\,dW(s), \quad u\in[0,T].$$

Therefore the predictable version is found.

(ii) Concerning the norm we get that

$$\begin{aligned}&\|\int_0^{\cdot} S(\cdot-s)B(Y(s))\,dW(s)\|_{\mathcal{H}^p}\\ &\le \left(\frac{p}{2}(p-1)\right)^{\frac12}\Big(\int_0^T K^2(s)\,ds\Big)^{\frac12}(1+\|Y\|_{\mathcal{H}^p})\end{aligned}$$

since we obtain by the help of the Burkholder-Davis-Gundy inequality that

$$\begin{aligned}
&\Big(E(\|\int_0^t S(t-s)B(Y(s))\,dW(s)\|^p)\Big)^{\frac{1}{p}}\\
&\le (\frac{p}{2}(p-1))^{\frac{1}{2}}\Big(\int_0^t \big(E(\|S(t-s)B(Y(s))\|_{L_2}^p)\big)^{\frac{2}{p}}\,ds\Big)^{\frac{1}{2}}\\
&\le (\frac{p}{2}(p-1))^{\frac{1}{2}}\Big(\int_0^t K^2(t-s)\big(E((1+\|Y(s)\|)^p)\big)^{\frac{2}{p}}\,ds\Big)^{\frac{1}{2}}\\
&\le (\frac{p}{2}(p-1))^{\frac{1}{2}}\Big(\int_0^t K^2(t-s)(1+\|Y(s)\|_{L^p})^2\,ds\Big)^{\frac{1}{2}}\\
&\le (\frac{p}{2}(p-1))^{\frac{1}{2}}\Big(\int_0^T K^2(s)\,ds\Big)^{\frac{1}{2}}(1+\|Y\|_{\mathcal{H}^p})
\end{aligned}$$

Therefore we have finally proved that

$$\mathcal{F}: L_0^p \times \mathcal{H}^p(T,H) \to \mathcal{H}^p(T,H).$$

Step 3: For each $p \ge 2$ there is a $\lambda(p) =: \lambda$ such that

$$\mathcal{F}(\xi,\cdot): \mathcal{H}^{p,\lambda}(T,H) \to \mathcal{H}^{p,\lambda}(T,H)$$

is a contraction for all $\xi \in L_0^p$ where the contraction constant $\alpha(\lambda) < 1$ does not depend on ξ:
Let Y, $\tilde{Y} \in \mathcal{H}^p(T,H)$, $\xi \in L_0^p$ and $t \in [0,T]$. Then we get that

$$\begin{aligned}
&\|[\mathcal{F}(\xi,Y) - \mathcal{F}(\xi,\tilde{Y})](t)\|_{L^p}\\
&\le \|\int_0^t S(t-s)[F(Y(s)) - F(\tilde{Y}(s))]\,ds\|_{L^p}\\
&\quad + \|\int_0^t S(t-s)[B(Y(s)) - B(\tilde{Y}(s))]\,dW(s)\|_{L^p}
\end{aligned}$$

where the first summand can be estimated in the following way:

$$\|\int_0^t S(t-s)[F(Y(s)) - F(\tilde{Y}(s))]\,ds\|^p \le M_T^p C^p T^{p-1}\int_0^t \|Y(s)-\tilde{Y}(s)\|^p\,ds$$

This implies that

$$\begin{aligned}
&\big(E(\|\int_0^t S(t-s)[F(Y(s)) - F(\tilde{Y}(s))]\,ds\|^p)\big)^{\frac{1}{p}}\\
&\le M_T C T^{\frac{p-1}{p}}(\int_0^t E(\|Y(s)-\tilde{Y}(s)\|^p)\,ds)^{\frac{1}{p}}\\
&= M_T C T^{\frac{p-1}{p}}(\int_0^t e^{\lambda p s}\underbrace{e^{-\lambda p s}E(\|Y(s)-\tilde{Y}(s)\|^p)}_{\le \|Y-\tilde{Y}\|_{p,\lambda,T}^p}\,ds)^{\frac{1}{p}}
\end{aligned}$$

$$\le M_T C T^{\frac{p-1}{p}} (\int_0^t e^{\lambda p s}\, ds)^{\frac{1}{p}} \|Y - \tilde{Y}\|_{p,\lambda,T}$$
$$\le M_T C T^{\frac{p-1}{p}} e^{\lambda t} (\frac{1}{\lambda p})^{\frac{1}{p}} \|Y - \tilde{Y}\|_{p,\lambda,T}$$

Dividing by $e^{\lambda t}$ provides the following result

$$\|\int_0^{\cdot} S(\cdot - s)[F(Y(s)) - F(\tilde{Y}(s))]\, ds\|_{p,\lambda,T} \le \underbrace{M_T C T^{\frac{p-1}{p}} (\frac{1}{\lambda p})^{\frac{1}{p}}}_{\longrightarrow 0 \quad \text{as } \lambda \to \infty} \|Y - \tilde{Y}\|_{p,\lambda,T}$$

By the Burkholder-Davis-Gundy inequality we get the following estimate for the second summand:

$$\Big(E(\|\int_0^t S(t-s)[B(Y(s)) - B(\tilde{Y}(s))]\, dW(s)\|^p)\Big)^{\frac{1}{p}}$$
$$\le (\frac{p}{2}(p-1))^{\frac{1}{2}} \Big(\int_0^t \big(E(\|S(t-s)[B(Y(s)) - B(\tilde{Y}(s))]\|_{L_2}^p)\big)^{\frac{2}{p}}\, ds\Big)^{\frac{1}{2}}$$
$$\le (\frac{p}{2}(p-1))^{\frac{1}{2}} \Big(\int_0^t K^2(t-s) \|Y(s) - \tilde{Y}(s)\|_{L^p}^2\, ds\Big)^{\frac{1}{2}}$$
$$= (\frac{p}{2}(p-1))^{\frac{1}{2}} \Big(\int_0^t K^2(t-s) e^{2\lambda s} \underbrace{e^{-2\lambda s} \|Y(s) - \tilde{Y}(s)\|_{L^p}^2}_{\le \|Y - \tilde{Y}\|_{p,\lambda,T}^2}\, ds\Big)^{\frac{1}{2}}$$
$$\le (\frac{p}{2}(p-1))^{\frac{1}{2}} \Big(\int_0^t K^2(t-s) e^{2\lambda s}\, ds\Big)^{\frac{1}{2}} \|Y - \tilde{Y}\|_{p,\lambda,T}$$
$$\le (\frac{p}{2}(p-1))^{\frac{1}{2}} e^{\lambda t} \Big(\int_0^T K^2(s) e^{-2\lambda s}\, ds\Big)^{\frac{1}{2}} \|Y - \tilde{Y}\|_{p,\lambda,T}$$

As for the first summand this implies that

$$\|\int_0^{\cdot} S(\cdot - s)[B(Y(s)) - B(\tilde{Y}(s))]\, dW(s)\|_{p,\lambda,T}$$
$$\le \underbrace{(\frac{p}{2}(p-1))^{\frac{1}{2}} \Big(\int_0^T K^2(s) e^{-2\lambda s}\, ds\Big)^{\frac{1}{2}}}_{\longrightarrow 0 \quad \text{as } \lambda \to \infty} \|Y - \tilde{Y}\|_{p,\lambda,T}$$

Therefore we have finally proved that there is a $\lambda = \lambda(p)$ such that there exists an $\alpha(\lambda) < 1$ with

$$\|\mathcal{F}(\xi, Y) - \mathcal{F}(\xi, \tilde{Y})\|_{p,\lambda,T} \le \alpha(\lambda) \|Y - \tilde{Y}\|_{p,\lambda,T}$$

for all $\xi \in L_0^p$, Y, $\tilde{Y} \in \mathcal{H}^{p,\lambda}(T,H)$. Thus the existence of a unique implicit function

$$\begin{aligned} X : L_0^p &\to \mathcal{H}^p(T,H) \\ \xi &\mapsto X(\xi) = \mathcal{F}(\xi, X(\xi)) \end{aligned}$$

is verified.
Step 4: We prove the Lipschitz continuity of $X : L_0^p \to \mathcal{H}^p(T,H)$.
By Theorem D.1 (ii) and the equivalence of the norms $\|\ \|_{\mathcal{H}^p}$ and $\|\ \|_{p,\lambda,T}$ we only have to check that the mappings

$$\mathcal{F}(\cdot, Y) : L_0^p \to \mathcal{H}^p(T,H)$$

are Lipschitz continuous for all $Y \in \mathcal{H}^p(T,H)$ where the Lipschitz constant does not depend on Y.
But this assertion holds as for all ξ, $\zeta \in L_0^p$ and $Y \in \mathcal{H}^p(T,H)$

$$\|\mathcal{F}(\xi, Y) - \mathcal{F}(\zeta, Y)\|_{\mathcal{H}^p} = \|S(\cdot)(\xi - \zeta)\|_{\mathcal{H}^p} \le M_T \|\xi - \zeta\|_{L^p}$$

3.2 Smoothing property of the semigroup: Pathwise continuity of the mild solution

Let $X(\xi)$ be the mild solution of problem (2.1) with initial condition $\xi \in L_0^p$. The aim of this section is to prove that the mapping $t \mapsto X(\xi)(t)$ has a continuous version. Because of Lemma 3.9 we already know that the process of the Bochner integrals

$$\int_0^t S(t-s) F(X(\xi)(s))\, ds, \quad t \in [0,T],$$

is P-a.s. continuous. Hence it only remains to show that the process

$$\int_0^t S(t-s) B(X(\xi)(s))\, dW(s), \quad t \in [0,T],$$

has a continuous version. To this end we use the method which is presented in [DaPrZa 96, Theorem 5.2.6, p.59; Proposition A.1.1, p.307]. In contrast to [DaPrZa 96] we do not demand that the semigroup is analytic and therefore we only get continuity instead of Hölder continuity as [DaPrZa 96].
First we have to introduce the general concept of the stochastic convolution.

Definition 3.11 (Stochastic convolution). If $\Phi(t)$, $t \in [0,T]$, is a $L(U,H)$-valued predictable process such that the stochastic integrals

$$W_A^\Phi(t) := \int_0^t S(t-s)\Phi(s)\, dW(s),\ t \in [0,T],$$

are well defined, then the process $W_A^\Phi(t)$, $t \in [0,T]$, is called stochastic convolution.

The following result (see [DaPrZa 96, Theorem 5.2.5, p.58]) is a corollary of the stochastic Fubini Theorem 1.37.

Theorem 3.12 (Factorization formula). *Let $\alpha \in]0,1[$ and Φ be a $L(U,H)$-valued predictable process. Assume that*

1. *$S(t-s)\Phi(s)$ is $L_2(U,H)$-valued for all $s \in [0,t[$, $t \in [0,T]$,*
2. *$\int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha} E\big(\|S(s-u)\Phi(u)\|_{L_2}^2\big)\, du\Big]^{\frac{1}{2}} ds < \infty$ for all $t \in [0,T]$.*

Then there is the following representation of the stochastic convolution.

$$\int_0^t S(t-s)\Phi(s)\, dW(s) = \frac{\sin \alpha\pi}{\pi} \int_0^t (t-s)^{\alpha-1} S(t-s) Y_\alpha^\Phi(s)\, ds \tag{3.3}$$

P-a.s. for all $t \in [0,T]$, where $Y_\alpha^\Phi(s)$, $s \in [0,T]$, is a $\mathcal{F}_T \otimes \mathcal{B}([0,T])$-measurable version of

$$\int_0^s (s-u)^{-\alpha} S(s-u)\Phi(u)\, dW(u),\ s \in [0,T]. \tag{3.4}$$

Proof. First we check that $\int_0^s (s-u)^{-\alpha} S(s-u)\Phi(u)\, dW(u)$, $s \in [0,T]$, is well defined and that there is a $\mathcal{F}_T \otimes \mathcal{B}([0,T])$-measurable version. But this is true since first we have that the mapping

$$\varphi : (u,\omega,s) \mapsto 1_{[0,s[}(u)(s-u)^{-\alpha} S(s-u)\Phi(u),\ u \in [0,T],$$

is $\mathcal{P}_T \otimes \mathcal{B}([0,T])/\mathcal{B}(L_2)$-measurable by assumption 1. (The proof can be done in a similar way as the proof of Lemma 3.5 and the proof of Lemma 3.6). Secondly, by assumption 2., we obtain that

$$\int_0^T \Big[E\Big(\int_0^T \|1_{[0,s[}(u)(s-u)^{-\alpha} S(s-u)\Phi(u)\|_{L_2}^2\, du\Big)\Big]^{\frac{1}{2}} ds$$

$$= \int_0^T \Big[\int_0^s (s-u)^{-2\alpha} E\big(\|S(s-u)\Phi(u)\|_{L_2}^2\big)\, du \Big]^{\frac{1}{2}} < \infty.$$

Therefore the mapping $\varphi : \Omega_T \times [0,T] \to L_2(U,H)$ fulfills the conditions of the stochastic Fubini Theorem 1.37 and thus the process of the integrals is well defined and has a product measurable version Y_α^Φ (see proof of [DaPrZa 92, Theorem 4.18, p.109]).
Besides the mapping φ_t given by

$$\begin{aligned} &\varphi_t : \Omega_T \times [0,T] \to L_2(U,H) \\ &\qquad (u,\omega,s) \mapsto 1_{[0,t[}(s)(t-s)^{\alpha-1} 1_{[0,s[}(u)(s-u)^{-\alpha} S(t-u)\Phi(u,\omega) \end{aligned}$$

also fulfills the conditions of the stochastic Fubini Theorem 1.37 for the following reasons:
For all $t \in [0,T]$ we have that the mapping

$$\varphi_t : \Omega_T \times [0,T] \to L_2(U,H)$$

is $\mathcal{P}_T \otimes \mathcal{B}([0,T])/\mathcal{B}(L_2)$-measurable. Moreover we get by assumption 2. that

$$\begin{aligned} &\int_0^T \Big(E\Big(\int_0^T 1_{[0,t[}(s) 1_{[0,s[}(u)(t-s)^{2(\alpha-1)}(s-u)^{-2\alpha} \\ &\qquad\qquad \|S(t-u)\Phi(u)\|_{L_2}^2\, du \Big)\Big)^{\frac{1}{2}}\, ds \\ &= \int_0^T 1_{[0,t[}(s)(t-s)^{\alpha-1} \Big(\int_0^T 1_{[0,s[}(u)(s-u)^{-2\alpha} \\ &\qquad\qquad E\big(\|S(t-u)\Phi(u)\|_{L_2}^2\big)\, du \Big)^{\frac{1}{2}}\, ds \\ &\quad \text{(by the real Fubini theorem)} \\ &\le M_T \int_0^T 1_{[0,t[}(s)(t-s)^{\alpha-1} \Big(\int_0^T 1_{[0,s[}(u)(s-u)^{-2\alpha} \\ &\qquad\qquad E\big(\|S(s-u)\Phi(u)\|_{L_2}^2\big)\, du \Big)^{\frac{1}{2}}\, ds < \infty \\ &\quad \text{(by the semigroup property)} \end{aligned}$$

Therefore there exists a product measurable version of

$$\int_0^s 1_{[0,t[}(s)(t-s)^{\alpha-1}(s-u)^{-\alpha} S(t-u)\Phi(u)\, dW(u), \quad s \in [0,T],$$

and for this version we get that

$$E\Big(\| \int_0^t (t-s)^{\alpha-1} S(t-s) Y_\alpha^\Phi(s)\, ds \\ - \int_0^T \int_0^s 1_{[0,t[}(s)(t-s)^{\alpha-1}(s-u)^{-\alpha} S(t-u)\Phi(u)\, dW(u)\, ds\|\Big) = 0.$$

Now can use the stochastic Fubini Theorem 1.37 to exchange the integration and thus we finally obtain that

$$\begin{aligned} &\int_0^t (t-s)^{\alpha-1} S(t-s) Y_\alpha^\Phi(s)\, ds \\ &= \int_0^t \int_0^s (t-s)^{\alpha-1}(s-u)^{-\alpha} S(t-u)\Phi(u)\, dW(u)\, ds \\ &= \int_0^t \int_u^t (t-s)^{\alpha-1}(s-u)^{-\alpha} S(t-u)\Phi(u)\, ds\, dW(u) \\ &= \int_0^t \Big(\int_u^t (t-s)^{\alpha-1}(s-u)^{-\alpha}\, ds\Big) S(t-u)\Phi(u)\, dW(u) \\ &= \frac{\pi}{\sin \alpha\pi} \int_0^t S(t-u)\Phi(u)\, dW(u) \quad P\text{-a.s.} \end{aligned}$$

since $\int_u^t (t-s)^{\alpha-1}(s-u)^{-\alpha}\, ds = \frac{\pi}{\sin \alpha\pi}$ □

Using this representation of the stochastic convolution we are now able to prove the demanded pathwise continuity. To this end we first consider the following case:
Let $\alpha \in]0,1[$ and $p > \frac{1}{\alpha}$.
For $\varphi \in L^p([0,T];H) := L^p([0,T],\mathcal{B}([0,T]),dx;H)$ we define

$$R_\alpha\varphi(t) := \int_0^t (t-s)^{\alpha-1} S(t-s)\varphi(s)\, ds,\ t \in [0,T].$$

Then $R_\alpha\varphi$ is well defined since

$$\int_0^t \|(t-s)^{\alpha-1} S(t-s)\varphi(s)\|\, ds \le \Big(\int_0^t s^{(\alpha-1)\frac{p}{p-1}}\, ds\Big)^{\frac{p-1}{p}} M_T \|\varphi\|_{L^p} < \infty$$

since $\alpha > \frac{1}{p}$ and therefore $(\alpha-1)\frac{p}{p-1} > -1$.

Proposition 3.13.

$$R_\alpha : L^p([0,T];H) \to C([0,T];H)$$

Remark 3.14. If one assumes that the semigroup $S(t)$, $t \in [0,T]$, is analytic one even gets that $R_\alpha\varphi$ is Hölder continuous for all $\varphi \in L^p([0,T];H)$ (see [DaPrZa 96, Proposition A.1.1, p.307]).

Proof of 3.13. Let $\varphi \in L^p([0,T];H)$, $t \in [0,T]$, and t_n, $n \in \mathbb{N}$, a sequence in $[0,T]$ such that $t_n \underset{n\to\infty}{\longrightarrow} t$. Then

$$
\begin{aligned}
&\|R_\alpha\varphi(t_n) - R_\alpha\varphi(t)\| \\
&= \|\int_0^{t_n} (t_n - s)^{\alpha-1} S(t_n - s)\varphi(s)\, ds - \int_0^t (t-s)^{\alpha-1} S(t-s)\varphi(s)\, ds\| \\
&\le \int_0^T \|1_{[0,t_n[}(s)(t_n - s)^{\alpha-1} S(t_n - s)\varphi(s) - 1_{[0,t[}(s)(t-s)^{\alpha-1} S(t-s)\varphi(s)\|\, ds
\end{aligned}
$$

Concerning the inner term we obtain that

$$
\|1_{[0,t_n[}(s)(t_n - s)^{\alpha-1} S(t_n - s)\varphi(s) - 1_{[0,t[}(s)(t-s)^{\alpha-1} S(t-s)\varphi(s)\| \underset{n\to\infty}{\longrightarrow} 0
$$

for dx-a.e. $s \in [0,T]$. Moreover the family

$$
(\|1_{[0,t_n[}(t_n - \cdot)^{\alpha-1} S(t_n - \cdot)\varphi(\cdot) - 1_{[0,t[}(t - \cdot)^{\alpha-1} S(t - \cdot)\varphi(\cdot)\|)_{n\in\mathbb{N}}
$$

is uniformly integrable :
For $t \in [0,T]$ we set

$$
F_t(s) := 1_{[0,t[}(s)(t-s)^{\alpha-1}\|S(t-s)\varphi(s)\|,\ s \in [0,T].
$$

Since $(\alpha - 1)\frac{p}{p-1} > -1$, there exists $\varepsilon > 0$ such that

$$
(\alpha - 1)(1+\varepsilon)\frac{p}{p-1-\varepsilon} > -1 \text{ and } \frac{p}{1+\varepsilon} > 1
$$

then

$$
\begin{aligned}
\int_0^T F_t^{1+\varepsilon}(s)\, ds &\le \Big(\int_0^t (t-s)^{(\alpha-1)(1+\varepsilon)\frac{p}{p-1-\varepsilon}}\, ds\Big)^{\frac{p-1-\varepsilon}{p}} M_T^{1+\varepsilon}\|\varphi\|_{L^p}^{1+\varepsilon} \\
&\le \Big(\int_0^T s^{(\alpha-1)(1+\varepsilon)\frac{p}{p-1-\varepsilon}}\, ds\Big)^{\frac{p-1}{p}} M_T^{1+\varepsilon}\|\varphi\|_{L^p}^{1+\varepsilon}
\end{aligned}
$$

Therefore $\sup_{t\in[0,T]} \int_0^T F_t^{1+\varepsilon}(s)\, ds < \infty$ and hence F_t, $t \in [0,T]$, is uniformly integrable. Since

$$
\begin{aligned}
&\|1_{[0,t_n[}(s)(t_n - s)^{\alpha-1} S(t_n - s)\varphi(s) - 1_{[0,t[}(s)(t-s)^{\alpha-1} S(t-s)\varphi(s)\| \\
&\le F_{t_n}(s) + F_t(s)
\end{aligned}
$$

for all $s \in [0,T]$ the assertion follows. □

In this way we have found an instrument to check if the process

$$\int_0^t S(t-s)B(X(\xi)(s))\,dW(s), \quad t\in[0,T],$$

has a P-a.s. continuous version. Altogether we can make the following statement concerning the mild solution of problem (2.1).

Proposition 3.15. *Assume that the mappings A, F and B satisfy Hypothesis H.0.*
If there exists $\alpha\in]0,\frac{1}{2}[$ such that

$$\int_0^T s^{-2\alpha}K^2(s)\,ds<\infty$$

then the mild solution $X(\xi)$ of problem (2.1) has a continuous version for all initial conditions $\xi\in L_0^p$, $p>\frac{1}{\alpha}$.

Proof. $S(\cdot)\xi$ is P-a.s. continuous because of the strong continuity of the semigroup.
In Step 2, 2. (i) of the proof of Theorem 3.2 we already have shown that the process of the Bochner integrals

$$\int_0^t S(t-s)F(X(\xi)(s))\,ds, \quad t\in[0,T],$$

has P-a.s. continuous trajectories.
Thus, in fact, it only remains to prove that the process

$$\int_0^t S(t-s)B(X(\xi)(s))\,dW(s), \quad t\in[0,T],$$

has a continuous version:
Since

$$\begin{aligned}
&\int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha}E\big(\|S(s-u)B(X(\xi)(u))\|_{L_2}^2\big)\,du\Big]^{\frac{1}{2}}\,ds\\
&\le \int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha}K^2(s-u)E\big((1+\|X(\xi)(u)\|)^2\big)\,du\Big]^{\frac{1}{2}}\,ds\\
&\le (1+\|X(\xi)\|_{\mathcal{H}^2})\int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha}K^2(s-u)\,du\Big]^{\frac{1}{2}}\,ds\\
&\le (1+\|X(\xi)\|_{\mathcal{H}^2})\Big(\int_0^T u^{-2\alpha}K^2(u)\,du\Big)^{\frac{1}{2}}\int_0^T s^{\alpha-1}\,ds<\infty
\end{aligned}$$

we have by Theorem 3.12 (factorization formula) that

$$\begin{aligned}
&\int_0^t S(t-s)B(X(\xi)(s))\,dW(s)\\
&= \frac{\sin\alpha\pi}{\pi}\int_0^t (t-s)^{\alpha-1}S(t-s)\int_0^s (s-u)^{-\alpha}S(s-u)B(X(\xi)(u))\,dW(u)\,ds\\
&= \frac{\sin\alpha\pi}{\pi}R_\alpha\big(\underbrace{\int_0^{\cdot}(\cdot-u)^{-\alpha}S(\cdot-u)B(X(\xi)(u))\,dW(u)}_{=:\,Y_\alpha}\big)(t)\quad P\text{-a.s.}
\end{aligned}$$

Since the mapping $\varphi : \Omega_T \times [0,T] \to L_2(U,H)$ given by

$$\varphi(u,\omega,s) := 1_{[0,s[}(u)(s-u)^{-\alpha}S(s-u)B(X(\xi)(u,\omega))$$

fulfills the conditions of the stochastic Fubini Theorem 1.37 the process Y_α can be understood as a $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable version of $\int_0^t (t-u)^{-\alpha}S(t-u)B(X(\xi)(u))\,dW(u)$, $t\in[0,T]$ (see proof of [DaPrZa 92, Theorem 4.18, p.109]).
To get the P-a.s. continuity of the stochastic integral we have to show that

$$Y_\alpha \in L^p([0,T];H)\quad P\text{-a.s. for } p > \frac{1}{\alpha}.$$

By the Burkholder-Davis-Gundy inequality we can estimate $E\big(\|Y_\alpha(t)\|^p\big)$ independently of $t\in[0,T]$ in the following way

$$\begin{aligned}
&E\big(\|Y_\alpha(t)\|^p\big)\\
&\le \big(\tfrac{p}{2}(p-1)\big)^{\frac{p}{2}}\Big(\int_0^t (t-s)^{-2\alpha}\big(E(\|S(t-s)B(X(\xi)(s))\|_{L_2}^p)\big)^{\frac{2}{p}}\,ds\Big)^{\frac{p}{2}}\\
&\le \big(\tfrac{p}{2}(p-1)\big)^{\frac{p}{2}}\Big(\int_0^t (t-s)^{-2\alpha}K^2(t-s)\big(E((1+\|X(\xi)(s)\|)^p)\big)^{\frac{2}{p}}\,ds\Big)^{\frac{p}{2}}\\
&\le \big(\tfrac{p}{2}(p-1)\big)^{\frac{p}{2}}\Big(\int_0^t (t-s)^{-2\alpha}K^2(t-s)(1+\|X(\xi)\|_{\mathcal{H}^p})^2\,ds\Big)^{\frac{p}{2}}\\
&\le \big(\tfrac{p}{2}(p-1)\big)^{\frac{p}{2}}(1+\|X(\xi)\|_{\mathcal{H}^p})^p\Big(\int_0^T s^{-2\alpha}K^2(s)\,ds\Big)^{\frac{p}{2}} < \infty
\end{aligned}$$

Finally we obtain by the real Fubini theorem that

$$E\Big(\int_0^T \|Y_\alpha(t)\|^p\,dt\Big) = \int_0^T E\big(\|Y_\alpha(t)\|^p\big)\,dt < \infty.$$

□

Chapter 4

Differentiable Dependence on Initial Data for the Mild Solution

In this chapter we analyze the first order differentiability of the mapping

$$X : L_0^p \to \mathcal{H}^p(T, H), \quad p \geq 2,$$

where $X(\xi)$ is still the mild solution of problem (2.1) with initial condition $\xi \in L_0^p$. (For details about the different concepts of differentiability see Appendix D.) To this end we make the following assumptions

Hypothesis H.1

- F and B are Fréchet differentiable where the derivatives

$$\begin{aligned} &DF : H \to L(H) \\ &DB : H \to L(H, L(U, H)) \text{ are continuous.} \end{aligned}$$

- For all $t \in]0, T]$ the mapping

$$S(t)DB : H \to L(H, L_2(U, H)) \text{ is continuous.}$$

Remark 4.1. (i) Comparing these assumptions with the usual ones (see [DaPrZa 96, Theorem 5.4.1 (i), p.69]) one will notice that on the one hand we do not demand the boundedness of DB while on the other hand the assumption concerning the continuity of $S(t)DB$ is additional. Even to prove the Gâteaux differentiability of X we need this stronger assumption.
(ii) We are especially interested in delimiting the Gâteaux from the Fréchet differentiability of X. Concerning the respective definitions see Appendix D.

Remark 4.2. (i) If F satisfies Hypotheses H.0 and H.1 we obtain that $\|DF(x)\|_{L(H)} \leq C$ for all $x \in H$.
(ii) If we just assume Hypothesis H.0 and the Fréchet differentiability of the mapping $B : H \to L(U, H)$ we already get for all $t \in]0, T]$ and $x \in H$ that $S(t)DB(x) \in L(H, L_2(U, H))$ with

$$\|S(t)DB(x)\|_{L(H,L_2)} \leq K(t)$$

Proof. (ii): By Hypothesis H.0 $\|S(t)B(z) - S(t)B(\tilde{z})\|_{L_2} \leq K(t)\|z - \tilde{z}\|$ for all $t \in]0, T]$ and $z, \tilde{z} \in H$.
Therefore we obtain for all $t \in]0, T]$ and $x, y \in H$ that

$$\|\frac{1}{h}(S(t)B(x + hy) - S(t)B(x))\|_{L_2} \leq K(t)\|y\|$$

If e_n, $n \in \mathbb{N}$, is an orthonormal basis of U then

$$\frac{1}{h}\big(S(t)B(x + hy) - S(t)B(x)\big)e_n \underset{h\to 0}{\longrightarrow} S(t)DB(x)y\, e_n \quad \text{for all } n \in \mathbb{N}.$$

By the lemma of Fatou it follows that

$$\begin{aligned}
\|S(t)DB(x)y\|_{L_2}^2 &= \sum_{n\in\mathbb{N}} \|S(t)DB(x)y\, e_n\|^2 \\
&= \sum_{n\in\mathbb{N}} \liminf_{h\to 0} \|\frac{1}{h}\big(S(t)B(x + hy) - S(t)B(x)\big)e_n\|^2 \\
&\leq \liminf_{h\to 0} \sum_{n\in\mathbb{N}} \|\frac{1}{h}\big(S(t)B(x + hy) - S(t)B(x)\big)e_n\|^2 \\
&= \liminf_{h\to 0} \|\frac{1}{h}\big(S(t)B(x + hy) - S(t)B(x)\big)\|_{L_2}^2 \\
&\leq K^2(t)\|y\|^2
\end{aligned}$$

and thus $\|S(t)DB(x)\|_{L(H,L_2)} \leq K(t)$ for all $t \in]0, T]$ and $x \in H$. □

Theorem 4.3. *Assume that the mappings A, F and B satisfy Hypothesis H.0 and Hypothesis H.1 and let $p \geq 2$. Then the following statements hold.*

(i) The mild solution

$$\begin{aligned} X : L_0^p &\to \mathcal{H}^p(T, H) \\ \xi &\mapsto X(\xi) \end{aligned}$$

is Gâteaux differentiable and the mapping

$$\begin{aligned} \partial X : L_0^p \times L_0^p &\to \mathcal{H}^p(T, H) \\ (\xi, \zeta) &\mapsto \partial X(\xi)\zeta \end{aligned}$$

is continuous.

(ii) The Gâteaux derivative of X fulfills the following equation for all $\xi, \zeta \in L_0^p$ and $t \in [0,T]$

$$\begin{aligned}\partial X(\xi)\zeta(t) =& S(t)\zeta + \int_0^t S(t-s)DF(X(\xi)(s))\partial X(\xi)\zeta(s)\, ds \\ &+ \int_0^t S(t-s)DB(X(\xi)(s))\partial X(\xi)\zeta(s)\, dW(s) \quad P\text{-a.s.}\end{aligned}$$

(iii) In addition the following estimate is true

$$\|\partial X(\xi)\zeta\|_{\mathcal{H}^p} \leq L_{T,p}\|\zeta\|_{L^p}$$

for all $\xi, \zeta \in L_0^p$ where $L_{T,p}$ is the Lipschitz constant of the mapping $X : L_0^p \to \mathcal{H}^p(T,H)$.

(iv) If $2 \leq p < q < \infty$ the mapping

$$\begin{aligned}X : L_0^q &\to \mathcal{H}^p(T,H) \\ \xi &\mapsto X(\xi)\end{aligned}$$

is continuously Fréchet differentiable.
In particular, the mapping

$$\begin{aligned}X : H &\to \mathcal{H}^p(T,H) \\ x &\mapsto X(x)\end{aligned}$$

is continuously Fréchet differentiable for all $p \geq 2$.

Remark 4.4. To prove the first assertion (i) we follow the way which is presented in [DaPrZa 96, Theorem 5.4.1 (i), p.69]. That means applying Theorem D.8 (i) (see [Za 98, Theorem 10.2, p.207]). But to get that the mild solution is (continuously) Fréchet differentiable we have to modify the abstract implicit function Theorem D.8 as in this case we have to consider not only one Banach space $E = \mathcal{H}^p(T,H)$ but also a further one $E_0 = \mathcal{H}^q(T,H) \subset \mathcal{H}^p(T,H)$, $q > p$. Zabczyk and Da Prato already have proposed this approach in the case of the second order derivatives (see Theorem D.13 (i)).

Before we can prove Theorem 4.3 we need the following lemmas.

Lemma 4.5. *Assume that the mapping B satisfies the conditions H.0 and H.1. Then for all $t \in]0,T]$ and $x, y \in H$*

$$\begin{aligned}&\left\|\frac{S(t)B(x+hy)-S(t)B(x)}{h} - S(t)DB(x)y\right\|_{L_2}\\ &\leq \frac{1}{h}\int_0^h \|S(t)DB(x+sy)y - S(t)DB(x)y\|_{L_2}\, ds\end{aligned}$$

and therefore we have in particular that

$$\frac{S(t)B(x+hy)-S(t)B(x)}{h} \underset{h\to 0}{\longrightarrow} S(t)DB(x)y \quad \textit{in } L_2(U,H)$$

Proof. For all $t \in]0,T]$ and $x, y \in H$ the mapping $s \mapsto S(t)B(x+sy)$ is continuously differentiable as a mapping from $\mathbb{R}$ to $L(U,H)$ with derivative

$$s \mapsto S(t)DB(x+sy)y \in L(U,H).$$

Therefore we get by the fundamental theorem for Bochner integrals, Theorem A.7, that

$$\begin{aligned}&\frac{S(t)B(x+hy)-S(t)B(x)}{h} - S(t)DB(x)y\\ &= \frac{1}{h}\int_0^h S(t)DB(x+sy)y - S(t)DB(x)y\, ds \quad \text{in } L(U,H).\end{aligned}$$

Besides we know by assumption that the mapping $s \mapsto S(t)DB(x+sy)y$ is continuous even as a mapping from $\mathbb{R}$ to $L_2(U,H)$. In particular it is $\mathcal{B}(\mathbb{R})/\mathcal{B}(L_2)$-measurable and it is possible to define

$$\int_0^h S(t)DB(x+sy)y - S(t)DB(x)y\, ds$$

also as $L_2(U,H)$-valued Bochner integral.
Since both integrals are identical the Bochner inequality with respect to the Hilbert-Schmidt norm $\|\cdot\|_{L_2}$ provides the desired result

$$\begin{aligned}&\left\|\frac{S(t)B(x+hy)-S(t)B(x)}{h} - S(t)DB(x)y\right\|_{L_2}\\ &\leq \frac{1}{h}\int_0^h \|S(t)DB(x+sy)y - S(t)DB(x)y\|_{L_2}\, ds\end{aligned}$$

The last assertion is a consequence of the continuity of $x \mapsto S(t)DB(x)y$ as a mapping to $L_2(U,H)$. $\square$

Lemma 4.6. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space where μ is finite and let (E, d) be a polish space.*
Moreover let $Y, Y_n, n \in \mathbb{N}$, be E-valued random variables on $(\Omega, \mathcal{F}, \mu)$ such that

$$Y_n \longrightarrow Y \quad \textit{in measure as } n \to \infty.$$

Let $(\tilde{E}, \tilde{d})$ be an arbitrary metric space and $f : (E, d) \to (\tilde{E}, \tilde{d})$ a continuous mapping. Then

$$f \circ Y_n \longrightarrow f \circ Y \quad \textit{in measure as } n \to \infty.$$

Proof. Without loss of generality we can assume that μ is a probability measure. Then we define $\nu := \mu \circ Y^{-1}$ and $\nu_n := \mu \circ Y_n^{-1}$, $n \in \mathbb{N}$, on $(E, \mathcal{B}(E))$.
Since $Y_n \longrightarrow Y$ in measure as $n \to \infty$ we have that $\nu_n \longrightarrow \nu$ weakly.
Then by [Bi 68, Theorem 6.2, p.37] $\{\nu_n \mid n \in \mathbb{N}\} \cup \{\nu\}$ is tight, i.e. for all $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset E$ such that

$$\nu(K_\varepsilon^c) \leq \varepsilon \text{ and } \nu_n(K_\varepsilon^c) \leq \varepsilon \text{ for all } n \in \mathbb{N}.$$

Let now $\varepsilon > 0$ and $c > 0$. Then

$$\begin{aligned} &\mu(\tilde{d}(f \circ Y_n, f \circ Y) \geq c) \\ &\leq \mu(\{\tilde{d}(f \circ Y_n, f \circ Y) \geq c\} \cap \{Y_n \in K_\varepsilon\} \cap \{Y \in K_\varepsilon\}) + 2\varepsilon \end{aligned}$$

Since $f : K_\varepsilon \to \tilde{E}$ is uniformly continuous there exists $\delta > 0$ such that $\tilde{d}(f(x), f(y)) < c$ for all $x, y \in K_\varepsilon$ with $d(x, y) < \delta$.
Therefore we can conclude

$$\begin{aligned} &\mu(\{\tilde{d}(f \circ Y_n, f \circ Y) \geq c\} \cap \{Y_n \in K_\varepsilon\} \cap \{Y \in K_\varepsilon\}) \\ &= \mu(\{\tilde{d}(f \circ Y_n, f \circ Y) \geq c\} \cap \{Y_n \in K_\varepsilon\} \cap \{Y \in K_\varepsilon\} \cap \{d(Y_n, Y) \geq \delta\}) \\ &\leq \mu(d(Y_n, Y) \geq \delta) < \varepsilon \end{aligned}$$

for n big enough since $Y_n \longrightarrow Y$ in measure.
Consequently $\mu(\tilde{d}(f \circ Y_n, f \circ Y) \geq c) \leq 3\varepsilon$ for n big enough. □

Proof of Theorem 4.3. Idea: In order to prove the Gâteaux resp. the Fréchet differentiability of the $\mathcal{H}^p(T, H)$-valued mapping X we want to apply Theorem D.8 to the spaces $\Lambda_0 := L_0^q$, $\Lambda := L_0^p$, $E_0 := \mathcal{H}^{q,\lambda(q)}(T, H)$ and $E := \mathcal{H}^{p,\lambda(p)}(T, H)$ and the mapping $G := \mathcal{F}$.

To be in the setting of Theorem D.8 we have to choose $\lambda(p), \lambda(q) > 0$ (from the proof of Theorem 3.2) such that

$$\begin{aligned} \mathcal{F} : L_0^p \times \mathcal{H}^{p,\lambda(p)}(T,H) &\to \mathcal{H}^{p,\lambda(p)}(T,H) \\ \text{and } \mathcal{F} : L_0^q \times \mathcal{H}^{q,\lambda(q)}(T,H) &\to \mathcal{H}^{q,\lambda(q)}(T,H) \end{aligned}$$

are contractions in the second variable. In this way we get differentiability properties of the mapping

$$X : L_0^r \to \mathcal{H}^{p,\lambda(p)}(T,H), \quad r \in \{p,q\}$$

and therefore because of the equivalence of the norms $\| \ \|_{p,\lambda,T}$, $\lambda \geq 0$, also of the mapping

$$X : L_0^r \to \mathcal{H}^p(T,H), \quad r \in \{p,q\}.$$

Since it is easier from a technical point of view we will check that the spaces $\Lambda_0 := L_0^q$, $\Lambda := L_0^p$, $E_0 := \mathcal{H}^q(T,H)$ and $E := \mathcal{H}^p(T,H)$ and the mapping $G := \mathcal{F}$ fulfill the conditions of
Theorem D.8.

Proof of (i): In Theorem 3.2 we already proved the continuity of the mapping

$$L_0^p \to \mathcal{H}^p(T,H), \quad \xi \mapsto \mathcal{F}(\xi, Y), \quad \text{for all } Y \in \mathcal{H}^p(T,H).$$

Hence it remains to verify the second condition of Theorem D.8 (i).
Step 1: We consider $\mathcal{F} : L_0^p \times \mathcal{H}^p(T,H) \to \mathcal{H}^p(T,H)$ and prove the existence of the directional derivatives.
(a) We fix $Y \in \mathcal{H}^p(T,H)$ and $\xi, \zeta \in L_0^p$ and prove that there exists the directional derivative $\partial_1 \mathcal{F}(\xi, Y; \zeta) = S(\cdot)\zeta \in \mathcal{H}^p(T,H)$:
This assertion is obvious since

$$\mathcal{F}(\xi + h\zeta, Y)(t) - \mathcal{F}(\xi, Y)(t) = hS(t)\zeta, \text{ for all } t \in [0,T] \text{ and } h \in \mathbb{R}$$

(b) We fix $\xi \in L_0^p$ and $Y, Z \in \mathcal{H}^p(T,H)$ and prove that there exists the directional derivative

$$\begin{aligned} \partial_2 \mathcal{F}(\xi, Y; Z) = (&\int_0^t S(t-s)DF(Y(s))Z(s)\, ds \\ &+ \int_0^t S(t-s)DB(Y(s))Z(s)\, dW(s))_{t \in [0,T]} \in \mathcal{H}^p(T,H) : \end{aligned}$$

Let $t \in [0,T]$ and $h \neq 0$. Then we get that

$$\begin{aligned}
&\|\frac{\mathcal{F}(\xi, Y+hZ)(t) - \mathcal{F}(\xi,Y)(t)}{h} \\
&- \int_0^t S(t-s)DF(Y(s))Z(s)\, ds - \int_0^t S(t-s)DB(Y(s))Z(s)\, dW(s)\|_{L^p} \\
\leq\ & \|\int_0^t S(t-s)\Big(\frac{F(Y(s)+hZ(s)) - F(Y(s))}{h} - DF(Y(s))Z(s)\Big)\, ds\|_{L^p} \\
&+ \|\int_0^t S(t-s)\Big(\frac{B(Y(s)+hZ(s)) - B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\, dW(s)\|_{L^p}
\end{aligned}$$

(1.) The first summand can be estimated independently of $t \in [0,T]$ in the following way

$$\begin{aligned}
&\Big(E\big(\|\int_0^t S(t-s)\Big(\frac{F(Y(s)+hZ(s)) - F(Y(s))}{h} - DF(Y(s))Z(s)\Big)\, ds\|^p\big)\Big)^{\frac{1}{p}} \\
\leq\ & M_T T^{\frac{p-1}{p}}\Big(E\big(\int_0^T \|\frac{F(Y(s)+hZ(s)) - F(Y(s))}{h} - DF(Y(s))Z(s)\|^p\, ds\big)\Big)^{\frac{1}{p}}
\end{aligned}$$

Since

$$\|\frac{F(Y(s,\omega)+hZ(s,\omega)) - F(Y(s,\omega))}{h} - DF(Y(s,\omega))Z(s,\omega)\|^p \underset{h\to 0}{\longrightarrow} 0$$

for all $(s,\omega) \in \Omega_T$ and is bounded by $2^p C^p \|Z\|^p \in L^1(\Omega_T, \mathcal{P}_T, P_T)$ we obtain that

$$\begin{aligned}
&\sup_{t\in[0,T]} \|\int_0^t S(t-s)\big(\frac{F(Y(s)+hZ(s)) - F(Y(s))}{h} - DF(Y(s))Z(s)\big)\, ds\|_{L^p} \\
\leq\ & M_T T^{\frac{p-1}{p}}\Big(E\big(\int_0^T \|\frac{F(Y(s)+hZ(s)) - F(Y(s))}{h} - DF(Y(s))Z(s)\|^p\, ds\big)\Big)^{\frac{1}{p}} \\
&\underset{h\to 0}{\longrightarrow} 0
\end{aligned}$$

by Lebesgue's dominated convergence theorem. In this way we get the desired convergence of the first summand in $\mathcal{H}^p(T,H)$.

(2.) Now we want to estimate the second summand. For that we fix $\lambda > 1$ and get by the Burkholder-Davis-Gundy inequality the following estimation

$$\begin{aligned}
&\Big(E\big(\|\int_0^t S(t-s)\frac{B(Y(s)+hZ(s)) - B(Y(s))}{h} \\
&\qquad - S(t-s)DB(Y(s))Z(s)\, dW(s)\|^p\big)\Big)^{\frac{1}{p}}
\end{aligned}$$

$$\le \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big(\int_0^t \Big(E\Big(\|S(t-s)\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - S(t-s)DB(Y(s))Z(s)\|_{L_2}^p\Big)\Big)^{\frac{2}{p}}\,ds\Big)^{\frac{1}{2}}$$

$$= \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big(\int_0^{\frac{t}{\lambda}} \Big(E\Big(\| S(t-s)\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - S(t-s)DB(Y(s))Z(s)\|_{L_2}^p\Big)\Big)^{\frac{2}{p}}\,ds + \int_{\frac{t}{\lambda}}^t \Big(E\Big(\| S(t-s)\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - S(t-s)DB(Y(s))Z(s)\|_{L_2}^p\Big)\Big)^{\frac{2}{p}}\,ds\Big)^{\frac{1}{2}}$$

By the semigroup property we obtain for the first integral that

$$\int_0^{\frac{t}{\lambda}} \Big[E\Big(\|S(t-s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big]^{\frac{2}{p}}\,ds$$

$$= \int_0^{\frac{t}{\lambda}} \Big[E\Big(\|S(t-\lambda s)S((\lambda-1)s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big]^{\frac{2}{p}}\,ds$$

$$\le M_T^2\int_0^T \Big[E\Big(\|S((\lambda-1)s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big]^{\frac{2}{p}} ds$$

If we fix $s \in]0,T]$ then we obtain by Lemma 4.5 that

$$\|S((\lambda-1)s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p \underset{h\to 0}{\longrightarrow} 0$$

P-a.s.. Moreover, by Hypothesis H.1 and Remark 4.2 (ii), the term is bounded by $2^pK^p((\lambda-1)s)\|Z(s)\|^p \in L^1(\Omega,\mathcal{F},P)$. Therefore we get by Lebesgue's dominated convergence theorem that

$$\Big(E\Big(\|S((\lambda-1)s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big)^{\frac{2}{p}} \underset{h\to 0}{\longrightarrow} 0$$

for all $s \in]0,T]$. In addition, for all $h \neq 0$ and $s \in]0,T]$ the above expectation is bounded by the function $4K^2((\lambda-1)\cdot)\|Z\|_{\mathcal{H}^p}^2 \in L^1([0,T],dx)$. Thus we obtain again by Lebesgue's dominated convergence theorem that

$$\int_0^T \Big[E\Big(\|S((\lambda-1)s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big]^{\frac{2}{p}}\, ds \underset{h\to 0}{\longrightarrow} 0$$

The second integral can be estimated independently of $h \neq 0$ and $t \in [0,T]$ as follows

$$\begin{aligned} &\int_{\frac{t}{\lambda}}^t \Big[E\Big(\|S(t-s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big]^{\frac{2}{p}}\, ds \\ &\leq \int_{\frac{t}{\lambda}}^t 4K^2(t-s)\big(E(\|Z(s)\|^p\big)^{\frac{2}{p}}\, ds \\ &\leq 4\|Z\|_{\mathcal{H}^p}^2 \int_0^{\frac{(\lambda-1)T}{\lambda}} K^2(s)\, ds \end{aligned}$$

where $\|Z\|_{\mathcal{H}^p} < \infty$ and $\int_0^{\frac{(\lambda-1)T}{\lambda}} K^2(s)\, ds \underset{\lambda\downarrow 1}{\longrightarrow} 0$, since $K \in L^2([0,T],dx)$. Altogether we are able to estimate the second summand independently of $t \in [0,T]$ and we have that

$$\begin{aligned} &\sup_{t\in[0,T]} \|\int_0^t S(t-s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\, dW(s)\|_{L^p} \\ &\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\Big(M_T^2 \int_0^T \Big[E\Big(\|S((\lambda-1)s)\Big(\frac{B(Y(s)+hZ(s))-B(Y(s))}{h} - DB(Y(s))Z(s)\Big)\|_{L_2}^p\Big)\Big]^{\frac{2}{p}} ds \\ &\qquad + 4\|Z\|_{\mathcal{H}^p}^2 \int_0^{\frac{(\lambda-1)T}{\lambda}} K^2(s)\, ds\Big)^{\frac{1}{2}} \end{aligned}$$

where the righthand side becomes small if λ is near to 1 and h tends to zero. So finally the desired convergence of the second summand in $\mathcal{H}^p(T,H)$ is proved.

Step 2: We prove the continuity of the directional derivatives

$$\begin{aligned}\partial_1\mathcal{F}:& L_0^p \times \mathcal{H}^p(T,H) \times L_0^p \to \mathcal{H}^p(T,H)\\ \partial_2\mathcal{F}:& L_0^p \times \mathcal{H}^p(T,H) \times \mathcal{H}^p(T,H) \to \mathcal{H}^p(T,H)\end{aligned}$$

(Notice that by Lemma D.4 this assertion especially implies that

$$\begin{aligned}&\partial_1\mathcal{F}(\xi,Y;\cdot) \in L(L_0^p, \mathcal{H}^p(T,H))\\ &\partial_2\mathcal{F}(\xi,Y;\cdot) \in L(\mathcal{H}^p(T,H))\end{aligned}$$

for all $\xi \in L_0^p$ and $Y \in \mathcal{H}^p(T,H)$. Hence Step 2 includes the result that the directional derivatives are the Gâteaux derivatives. Therefore then it is justified to write $\partial_1\mathcal{F}(\xi,Y)\zeta$ and $\partial_2\mathcal{F}(\xi,Y)Z$ instead of $\partial_1\mathcal{F}(\xi,Y;\zeta)$ and $\partial_2\mathcal{F}(\xi,Y;Z)$.):
(a) Since for all $\xi,\zeta \in L_0^p$ and $Y \in \mathcal{H}^p(T,H)$

$$\partial_1\mathcal{F}(\xi,Y;\zeta) = (S(t)\zeta)_{t\in[0,T]}$$

the continuity of $\partial_1\mathcal{F}$ follows immediately from the fact that for all $t \in [0,T]$ $\|S(t)\|_{L(H)} \leq M_T$.
(b) To analyze the continuity of

$$\begin{aligned}\partial_2\mathcal{F}:& L_0^p \times \mathcal{H}^p(T,H) \times \mathcal{H}^p(T,H) \to \mathcal{H}^p(T,H)\\ &(\xi,Y,Z) \mapsto \Big(\int_0^t S(t-s)DF(Y(s))Z(s)\,ds\\ &\qquad\qquad + \int_0^t S(t-s)DB(Y(s))Z(s)\,dW(s)\Big)_{t\in[0,T]}\end{aligned}$$

we take $\xi,\xi_n \in L_0^p$, $Y, Y_n, Z, Z_n \in \mathcal{H}^p(T,H)$, $n \in \mathbb{N}$, such that $\|\xi_n - \xi\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0$, $\|Y_n - Y\|_{\mathcal{H}^p} \underset{n\to\infty}{\longrightarrow} 0$ and $\|Z_n - Z\|_{\mathcal{H}^p} \underset{n\to\infty}{\longrightarrow} 0$. Then we have for $t \in [0,T]$

$$\begin{aligned}&\|\partial_2\mathcal{F}(\xi_n,Y_n;Z_n)(t) - \partial_2\mathcal{F}(\xi,Y;Z)(t)\|_{L^p}\\ &\leq \Big\|\int_0^t S(t-s)\big(DF(Y_n(s))Z_n(s) - DF(Y(s))Z(s)\big)\Big\|_{L^p} \qquad (1.)\\ &\quad + \Big\|\int_0^t S(t-s)\big(DB(Y_n(s))Z_n(s) - DB(Y(s))Z(s)\big)\Big\|_{L^p} \qquad (2.)\end{aligned}$$

(1.) At first we want to estimate the first summand independently of $t \in [0,T]$.

$$\Big(E\big(\Big\|\int_0^t S(t-s)(DF(Y_n(s))Z_n(s) - DF(Y(s))Z(s))\,ds\Big\|^p\big)\Big)^{\frac{1}{p}}$$

$$
\begin{aligned}
&\leq M_T T^{\frac{p-1}{p}} \Big[\Big(E\big(\int_0^T \|DF(Y_n(s))\|^p_{L(H)} \|Z_n(s) - Z(s)\|^p \, ds \big) \Big)^{\frac{1}{p}} \\
&\qquad + \Big(E\big(\int_0^T \|DF(Y_n(s))Z(s) - DF(Y(s))Z(s)\|^p \, ds \big) \Big)^{\frac{1}{p}} \Big] \\
&\leq M_T T^{\frac{p-1}{p}} \Big[C T^{\frac{1}{p}} \|Z_n - Z\|_{\mathcal{H}^p} \\
&\qquad + \Big(E\big(\int_0^T \|DF(Y_n(s))Z(s) - DF(Y(s))Z(s)\|^p \, ds \big) \Big)^{\frac{1}{p}} \Big]
\end{aligned}
$$

The first summand $CT^{\frac{1}{p}}\|Z_n - Z\|_{\mathcal{H}^p}$ converges to zero as $n \to \infty$ by assumption.
Concerning the second summand we obtain by Lemma 4.6 that

$$
\|DF(Y_n)Z - DF(Y)Z\| \underset{n\to\infty}{\longrightarrow} 0 \quad \text{in } P_T\text{-measure}
$$

since $(Y_n, Z) \longrightarrow (Y, Z)$ in P_T-measure and $DF : H \times H \to H$, $(y, z) \mapsto DF(y)z$, is continuous. Moreover

$$
\|DF(Y_n)Z - DF(Y)Z\|^p \leq 2^p C^p \|Z\|^p \in L^1(\Omega_T, \mathcal{P}_T, P_T)
$$

for all $n \in \mathbb{N}$. Hence $\|DF(Y_n)Z - DF(Y)Z\|^p$, $n \in \mathbb{N}$, is uniformly integrable and altogether we obtain that

$$
\Big(E\big(\int_0^T \|DF(Y_n)Z - DF(Y)Z\|^p \, ds \big) \Big)^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} 0
$$

Since this convergence does not depend on $t \in [0, T]$ we finally have that

$$
\sup_{t\in[0,T]} \Big\| \int_0^t S(t-s)\big(DF(Y_n(s))Z_n(s) - DF(Y(s))Z(s) \big) \, ds \Big\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0
$$

(2.) As next step we want to estimate the second summand independently of $t \in [0, T]$.
Using the Burkholder-Davis-Gundy inequality and the triangle inequality for $\|\cdot\|_{L^p}$ we obtain for fixed $t \in [0, T]$ and for arbitrary $\lambda > 1$ that

$$
\begin{aligned}
&\Big(E\big(\Big\| \int_0^t S(t-s)\big(DB(Y_n(s))Z_n(s) - DB(Y(s))Z(s) \big) \, dW(s) \Big\|^p \big) \Big)^{\frac{1}{p}} \\
&\leq \big(\frac{p}{2}(p-1) \big)^{\frac{1}{2}} \\
&\quad \Big(\int_0^t \Big(E\big(\|S(t-s)\big(DB(Y_n(s))Z_n(s) - DB(Y(s))Z(s) \big)\|^p_{L_2} \big) \Big)^{\frac{2}{p}} \, ds \Big)^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big[\, 2\int_0^t \Big(E\big(\|S(t-s)DB(Y_n(s))(Z_n(s)-Z(s))\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds \\
&\quad + 2\int_0^t \Big(E\big(\|S(t-s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds\,\Big]^{\frac{1}{2}} \\
&\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big[\, 2\int_0^t K^2(t-s)\Big(E\big(\|Z_n(s)-Z(s)\|^p\big)\Big)^{\frac{2}{p}}\, ds \\
&\quad + 2\int_0^{\frac{t}{\lambda}} \Big(E\big(\|S(t-s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds \\
&\quad + 2\int_{\frac{t}{\lambda}}^t \Big(E\big(\|S(t-s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds\,\Big]^{\frac{1}{2}} \\
&\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big[\, 2\int_0^T K^2(s)\, ds\, \|Z_n-Z\|_{\mathcal{H}^p}^2 \\
&\quad + 2\int_0^{\frac{t}{\lambda}} M_T^2\Big(E\big(\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds \\
&\quad + 8\int_{\frac{t}{\lambda}}^t K^2(t-s)\Big(E\big(\|Z(s)\|^p\big)\Big)^{\frac{2}{p}}\, ds\,\Big]^{\frac{1}{2}} \\
&\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big[\, 2\int_0^T K^2(s)\, ds\, \|Z_n-Z\|_{\mathcal{H}^p}^2 \\
&\quad + 2M_T^2\int_0^T \Big(E\big(\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds \\
&\quad + 8\int_0^{\frac{(\lambda-1)T}{\lambda}} K^2(s)\, ds\, \|Z\|_{\mathcal{H}^p}^2\,\Big]^{\frac{1}{2}}
\end{aligned}
$$

Since $\|Z_n - Z\|_{\mathcal{H}^p} \longrightarrow 0$ as $n \to \infty$ and since by Lebesgue's dominated convergence theorem $\int_0^{\frac{(\lambda-1)T}{\lambda}} K^2(s)\, ds$ becomes small for λ near to 1 it only remains to show that

$$
\int_0^T \Big(E\big(\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\, ds \underset{n\to\infty}{\longrightarrow} 0
$$

By Lemma 4.6 the inner term

$$
\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p \underset{n\to\infty}{\longrightarrow} 0
$$

in probability for all $s \in]0,T]$. Moreover it is dominated by $2^p K^p((\lambda-1)s)\|Z(s)\|^p \in L^1(\Omega,\mathcal{F},P)$ for all $n \in \mathbb{N}$ and $s \in]0,T]$. Therefore we obtain that

$$
\Big(E\big(\|S\big((\lambda-1)s\big)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}} \underset{n\to\infty}{\longrightarrow} 0
$$

for all $s \in]0,T]$.
Since the above expectation is bounded by the function $4K^2((\lambda-1)\cdot)\|Z\|^2_{\mathcal{H}^p} \in L^1([0,T],dx)$ we obtain by Lebesgue's dominated convergence theorem that

$$\int_0^T \Big(E\big(\|S\big((\lambda-1)s\big)[DBY_n(s)) - DB(Y(s))]Z(s)\|^p_{L_2}\big)\Big)^{\frac{2}{p}}\, ds \underset{n\to\infty}{\longrightarrow} 0$$

So finally we get the required convergence in $\mathcal{H}^p(T,H)$

$$\sup_{t\in[0,T]} \Big(E\big(\|\int_0^t S(t-s)\big(DB(Y_n(s))Z_n(s) - DB(Y(s))Z(s)\big)\, dW(s)\|^p\big)\Big)^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} 0.$$

Proof of (ii): Let $\xi,\zeta \in L^p_0$. Then by Theorem D.8 (i) we have the following representation of the Gâteaux derivative of X:

$$\partial X(\xi)\zeta = \big[I - \partial_2\mathcal{F}(\xi, X(\xi))\big]^{-1}\partial_1\mathcal{F}(\xi,X(\xi))\zeta$$

and therefore

$$\begin{aligned}\partial X(\xi)\zeta &= \partial_1\mathcal{F}(\xi,X(\xi))\zeta + \partial_2\mathcal{F}(\xi,X(\xi))\partial X(\xi)\zeta\\ &= \Big(S(t)\zeta + \int_0^t S(t-s)DF\big(X(\xi)(s)\big)\partial X(\xi)\zeta(s)\, ds\\ &\qquad + \int_0^t S(t-s)DB\big(X(\xi)(s)\big)\partial X(\xi)\zeta(s)\, ds\Big)_{t\in[0,T]}\end{aligned}$$

Proof of (iii): By Theorem 3.2 the mild solution $X : L^p_0 \to \mathcal{H}^p(T,H)$ is Lipschitz continuous with Lipschitz constant $L_{T,p}$. Hence we get that

$$\|\partial X(\xi)\zeta\|_{\mathcal{H}^p} = \lim_{h\to 0}\|\frac{X(\xi+h\zeta)-X(\xi)}{h}\|_{\mathcal{H}^p} \le L_{T,p}\|\zeta\|_{L^p} \quad \text{for all } \xi,\zeta\in L^p_0$$

Proof of (iv): To prove the Fréchet differentiability of the mild solution X it remains to verify the fifth condition of Theorem D.8.
(a) The mapping

$$\begin{aligned}\partial_1\mathcal{F} : L^q_0 \times \mathcal{H}^q(T,H) &\to L(L^q_0, \mathcal{H}^p(T,H))\\ (\xi,Y) &\mapsto \partial_1\mathcal{F}(\xi,Y)\end{aligned}$$

is continuous since it is constant.
(b) To prove the continuity of the mapping

$$\begin{aligned}\partial_2\mathcal{F} : L^q_0 \times \mathcal{H}^q(T,H) &\to L(\mathcal{H}^q(T,H), \mathcal{H}^p(T,H))\\ (\xi,Y) &\mapsto \partial_2\mathcal{F}(\xi,Y)\end{aligned}$$

let $\xi, \xi_n \in L_0^q$ and $Y, Y_n, Z \in \mathcal{H}^q(T, H)$, $n \in \mathbb{N}$, such that $\|\xi_n - \xi\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0$ and $\|Y_n - Y\|_{\mathcal{H}^q} \underset{n\to\infty}{\longrightarrow} 0$. Then we have to show the existence of a sequence of positive real numbers c_n, $n \in \mathbb{N}$, independent of $t \in [0, T]$ such that $c_n \underset{n\to\infty}{\longrightarrow} 0$ and

$$
\begin{aligned}
&\|\partial_2 \mathcal{F}(\xi_n, Y_n)Z(t) - \partial_2 \mathcal{F}(\xi, Y)Z(t)\|_{L^p} \\
&\le \|\int_0^t S(t-s)\big(DF(Y_n(s))Z(s) - DF(Y(s))Z(s)\big)\, ds\|_{L^p} \qquad (1.)\\
&\quad + \|\int_0^t S(t-s)\big(DB(Y_n(s))Z(s) - DB(Y(s))Z(s)\big)\, ds\|_{L^p} \qquad (2.)\\
&\le c_n \|Z\|_{\mathcal{H}^q}
\end{aligned}
$$

for all $t \in [0, T]$.

(1.) The first summand can be estimated independently of $t \in [0, T]$ in the following way

$$
\begin{aligned}
&\Big(E\big(\|\int_0^t S(t-s)\big(DF(Y_n(s))Z(s) - DF(Y(s))Z(s)\big)\, ds\|^p\big)\Big)^{\frac{1}{p}} \\
&\le M_T T^{\frac{p-1}{p}} \Big(E\big(\int_0^T \|DF(Y_n(s))Z(s) - DF(Y(s))Z(s)\|^p\, ds\big)\Big)^{\frac{1}{p}} \\
&\le M_T T^{\frac{p-1}{p}} \Big(\int_0^T E\big(\|DF(Y_n(s)) - DF(Y(s))\|_{L(H)}^p \|Z(s)\|^p\big)\, ds\Big)^{\frac{1}{p}} \\
&\le M_T T^{\frac{p-1}{p}} \Big(\int_0^T E\big(\|DF(Y_n(s)) - DF(Y(s))\|_{L(H)}^{pr'}\big)\, ds\Big)^{\frac{1}{pr'}} \\
&\qquad \Big(\int_0^T E\big(\|Z(s)\|^q\big)\, ds\Big)^{\frac{1}{q}} \\
&\quad \text{(Hölder inequality for } r = \frac{q}{p} > 1 \text{ and } r' = \frac{q}{q-p}) \\
&\le M_T T^{\frac{p-1}{p}} \Big(\int_0^T E\big(\|DF(Y_n(s)) - DF(Y(s))\|_{L(H)}^{pr'}\big)\, ds\Big)^{\frac{1}{pr'}} \\
&\quad T^{\frac{1}{q}} \|Z\|_{\mathcal{H}^q}
\end{aligned}
$$

As in part (i) Step 2 (b) we get that

$$
a_n := M_T T^{\frac{p-1}{p}} T^{\frac{1}{q}} \Big(\int_0^T E\big(\|DF(Y_n(s)) - DF(Y(s))\|_{L(H)}^{pr'}\big)\, ds\Big)^{\frac{1}{pr'}} \underset{n\to\infty}{\longrightarrow} 0
$$

(2.) To estimate the second summand we fix $\lambda > 1$ and use the Burkholder-Davis-Gundy inequality.

$$
\Big(E\big(\|\int_0^t S(t-s)\big(DB(Y_n(s))Z(s) - DB(Y(s))Z(s)\big)\, dW(s)\|^p\big)\Big)^{\frac{1}{p}}
$$

$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}$$
$$\Big[\int_0^{\frac{t}{\lambda}}\Big(E\big(\|S(t-s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds$$
$$+\int_{\frac{t}{\lambda}}^t\Big(E\big(\|S(t-s)[DB(Y_n(s))-DB(Y(s))]Z(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds\quad\Big]^{\frac{1}{2}}$$
$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}$$
$$\Big[M_T^2\int_0^T\Big(E\big(\,\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]\|_{L(H,L_2)}^p$$
$$\|Z(s)\|^p\big)\Big)^{\frac{2}{p}}\,ds$$
$$+\int_{\frac{t}{\lambda}}^t\Big(E\big(\,\|S(t-s)[DB(Y_n(s))-DB(Y(s))]\|_{L(H,L_2)}^p$$
$$\|Z(s)\|^p\big)\Big)^{\frac{2}{p}}\,ds\quad\Big]^{\frac{1}{2}}$$
$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\sup_{t\in[0,T]}\Big(E\big(\|Z(t)\|^q\big)\Big)^{\frac{1}{q}}$$
$$\Big[M_T^2\int_0^T\Big(E\big(\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]\|_{L(H,L_2)}^{pr'}\big)\Big)^{\frac{2}{pr'}}$$
$$+\int_{\frac{t}{\lambda}}^t\Big(E\big(\|S(t-s)[DB(Y_n(s))-DB(Y(s))]\|_{L(H,L_2)}^{pr'}\big)\Big)^{\frac{2}{pr'}}\quad\Big]^{\frac{1}{2}}$$

(Hölder inequality for $r=\dfrac{q}{p}>1$ and $r'=\dfrac{q}{q-p}$)

$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\|Z\|_{\mathcal{H}^q}$$
$$\Big[M_T^2\int_0^T\Big(E\big(\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]\|_{L(H,L_2)}^{pr'}\big)\Big)^{\frac{2}{pr'}}\,ds$$
$$+4\int_0^{\frac{(\lambda-1)T}{\lambda}}K^2(s)\,ds\quad\Big]^{\frac{1}{2}}$$

As in part (i) Step 2 (b) we get that

$$\int_0^T\Big(E\big(\|S((\lambda-1)s)[DB(Y_n(s))-DB(Y(s))]\|_{L(H,L_2)}^{pr'}\big)\Big)^{\frac{2}{pr'}}\,ds\underset{n\to\infty}{\longrightarrow}0$$

(This is the moment where we use the full assumption that $S(t)DB: H\to L(H,L_2(U,H))$ is continuous.)

Besides $\displaystyle\int_0^{\frac{(\lambda-1)T}{\lambda}}K^2(s)\,ds\underset{\lambda\downarrow 1}{\longrightarrow}0$ so that we get the existence of a sequence

b_n, $n \in \mathbb{N}$, such that $b_n \underset{n\to\infty}{\longrightarrow} 0$ and

$$\|\int_0^{\cdot} S(\cdot - s)\big(DB(Y_n(s))Z(s) - DB(Y(s))Z(s)\big)\, dW(s)\|_{\mathcal{H}^p} \le b_n \|Z\|_{\mathcal{H}^q}$$

Altogether we have that

$$\|\partial_2\mathcal{F}(\xi_n, Y_n)Z - \partial_2\mathcal{F}(\xi, Y)Z\|_{\mathcal{H}^p} \le \underbrace{(a_n + b_n)}_{=:\, c_n}\|Z\|_{\mathcal{H}^q}$$

where $c_n \underset{n\to\infty}{\longrightarrow} 0$. □

Now we are interested in finding conditions under which the mapping $DX : L_0^q \to \mathcal{H}^p(T, H)$ is even uniformly continuous. This result will be relevant in chapter 6.

Hypothesis H.1′

- $DF : H \to L(H)$ is uniformly continuous.

- There exists a mapping $\tilde{K} : [0, T] \to [0, \infty[$, $\tilde{K} \in L^2([0,T], dx)$ such that

  $$\|S(t)\big(DB(x) - DB(y)\big)\|_{L(H,L_2)} \le \tilde{K}(t)\|x - y\|$$

  for all $t \in]0, T]$ and $x, y \in H$.

Corollary 4.7. *Assume that the mappings A, F and B satisfy the Hypotheses H.0, H.1 and H.1' and let* $4 \le 2p \le q < \infty$. *Then*

$$DX : L_0^q \to L\big(L_0^q, \mathcal{H}^p(T, H)\big)$$

is uniformly continuous.

Proof. By Corollary D.12 we only have to check that the mappings

$$D_1\mathcal{F} : L_0^q \times \mathcal{H}^q(T,H) \to L\big(L_0^q, \mathcal{H}^p(T,H)\big)$$
$$\text{and } D_2\mathcal{F} : L_0^q \times \mathcal{H}^q(T,H) \to L\big(\mathcal{H}^q(T,H), \mathcal{H}^p(T,H)\big)$$

are uniformly continuous since we already know that the derivative $DX : L_0^q \to L(L_0^q, \mathcal{H}^p(T,H))$ is bounded (see Theorem 4.3 (iii)).
Since $D_1\mathcal{F}$ is constant we only have to prove the uniform continuity of $D_2\mathcal{F}$. For that let $\xi, \tilde{\xi} \in L_0^q$, $Y, \tilde{Y}, Z \in \mathcal{H}^q(T,H)$. Then we get for all $t \in [0,T]$:

$$\begin{aligned}
&\|D_2\mathcal{F}(\xi,Y)Z(t)-D_2\mathcal{F}(\tilde{\xi},\tilde{Y})Z(t)\|_{L^p}\\
&\le \|\int_0^t S(t-s)[DF\big(Y(s)\big)-DF\big(\tilde{Y}(s)\big)]Z(s)\,ds\|_{L^p} && (1.)\\
&\quad+\|\int_0^t S(t-s)[DB\big(Y(s)\big)-DB\big(\tilde{Y}(s)\big)]Z(s)\,ds\|_{L^p} && (2.)
\end{aligned}$$

(1.) Let $\varepsilon>0$. Since $DF:H\to L(H)$ is uniformly continuous there exists $\delta>0$ such that

$$\|DF(x)-DF(y)\|_{L(H)}\le\varepsilon \quad \text{for all } x,y\in H \text{ with } \|x-y\|\le\delta.$$

Using this fact we obtain that

$$\begin{aligned}
&\|\int_0^t S(t-s)[DF\big(Y(s)\big)-DF\big(\tilde{Y}(s)\big)]Z(s)\,ds\|_{L^p}\\
&\le M_T T^{\frac{p-1}{p}}\Big(\int_0^T E\big(\|[DF\big(Y(s)\big)-DF\big(\tilde{Y}(s)\big)]Z(s)\|^p\big)\,ds\Big)^{\frac{1}{p}}\\
&\le M_T T^{\frac{p-1}{p}}\Big(\int_0^T \Big(E\big(\|DF\big(Y(s)\big)-DF\big(\tilde{Y}(s)\big)\|_{L(H)}^{2p}\big)\Big)^{\frac{1}{2}}\Big(E\big(\|Z(s)\|^{2p}\big)\Big)^{\frac{1}{2}}\,ds\Big)^{\frac{1}{p}}\\
&\le M_T T^{\frac{p-1}{p}}\Big(\int_0^T \Big(E\big(\|DF\big(Y(s)\big)-DF\big(\tilde{Y}(s)\big)\|_{L(H)}^{2p}1_{\{\|Y(s)-\tilde{Y}(s)\|\le\delta\}}\big)\Big)^{\frac{1}{2}}\\
&\qquad\qquad+\Big(E\big(\|DF\big(Y(s)\big)-DF\big(\tilde{Y}(s)\big)\|_{L(H)}^{2p}1_{\{\|Y(s)-\tilde{Y}(s)\|>\delta\}}\big)\Big)^{\frac{1}{2}}\,ds\Big)^{\frac{1}{p}}\\
&\quad\|Z\|_{\mathcal{H}^{2p}(T,H)}\\
&\le M_T T^{\frac{p-1}{p}}\Big(\int_0^T \varepsilon^p+(2C)^p\big(P(\|Y(s)-\tilde{Y}(s)\|>\delta)\big)^{\frac{1}{2}}\,ds\Big)^{\frac{1}{p}}\|Z\|_{\mathcal{H}^{2p}(T,H)}\\
&\le M_T T^{\frac{p-1}{p}}\Big(\int_0^T \varepsilon^p+(2C)^p\frac{1}{\delta}\big(E(\|Y(s)-\tilde{Y}(s)\|^2)\big)^{\frac{1}{2}}\,ds\Big)^{\frac{1}{p}}\|Z\|_{\mathcal{H}^{2p}(T,H)}\\
&\le M_T T\Big(\varepsilon^p+(2C)^p\frac{1}{\delta}\|Y-\tilde{Y}\|_{\mathcal{H}^2(T,H)}\Big)^{\frac{1}{p}}\|Z\|_{\mathcal{H}^{2p}(T,H)}
\end{aligned}$$

(2.) By the Burkholder-Davis-Gundy inequality we get that

$$\begin{aligned}
&\|\int_0^t S(t-s)[DB\big(Y(s)\big)-DB\big(\tilde{Y}(s)\big)]Z(s)\,ds\|_{L^p}\\
&\le \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\Big(\int_0^t \big(E\big(\|S(t-s)[DB(Y(s))-DB(\tilde{Y}(s))]Z(s)\|_{L_2}^p\big)\big)^{\frac{2}{p}}\,ds\Big)^{\frac{1}{2}}\\
&\le \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\Big(\int_0^t \tilde{K}^2(t-s)\big(E\big(\|Y(s)-\tilde{Y}(s)\|^p\|Z(s)\|^p\big)\big)^{\frac{2}{p}}\,ds\Big)^{\frac{1}{2}}
\end{aligned}$$

$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big(\int_0^t \tilde{K}^2(t-s)\big(E\big(\|Y(s)-\tilde{Y}(s)\|^{2p}\big)^{\frac{1}{p}}\big(E\big(\|Z(s)\|^{2p}\big)\big)^{\frac{1}{p}}\,ds\Big)^{\frac{1}{2}}$$
$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big(\int_0^T \tilde{K}^2(s)\,ds\Big)^{\frac{1}{2}}\|Y-\tilde{Y}\|_{\mathcal{H}^{2p}(T;H)}\|Z\|_{\mathcal{H}^{2p}(T;H)}$$

Since the estimates (1.) and (2.) are independent of $t \in [0,T]$ the mapping

$$D_2\mathcal{F} : L_0^q \times \mathcal{H}^q(T,H) \to L(\mathcal{H}^q(T,H), \mathcal{H}^p(T,H))$$

is uniformly continuous. □

In the second part of chapter 6 we analyze the existence of a strict solution of the Kolmogorov equation which is related to problem (2.1). For that it is important to know assumptions under which $DX(x)y(t)$, $t \in [0,T]$, $x, y \in H$, has a pathwise continuous version.

Proposition 4.8. *We assume that the Hypotheses H.0 and H.1 are fulfilled and in addition we demand that there exists $\alpha \in]0,\frac{1}{2}[$ such that*

$$\int_0^T s^{-2\alpha}K^2(s)\,ds < \infty$$

Then we get that there is a continuous version of

$$\partial X(\xi)\zeta(t), \quad t \in [0,T],$$

for all $\xi, \zeta \in L_0^p$, $p > \frac{1}{\alpha}$.

Proof. By Theorem 4.3 (ii) we know that

$$\partial X(\xi)\zeta(t) =; S(t)\zeta + \int_0^t S(t-s)DF(X(\xi)(s))\partial X(\xi)\zeta(s)\,ds$$
$$+ \int_0^t S(t-s)DB(X(\xi)(s))\partial X(\xi)\zeta(s)\,dW(s) \quad P\text{-a.s.}$$

where:
1. It is clear that $t \mapsto S(t)\zeta$ is P-a.s. continuous since $S(t)$, $t \in [0,T]$, is a C_0- semigroup.
2. Since the process $DF(X(\xi)(s))\partial X(\xi)\zeta(s)$, $s \in [0,T]$, is predictable and P-a.s. Bochner integrable Lemma 3.9 provides that

$$t \mapsto \int_0^t S(t-s)DF(X(\xi)(s))\partial X(\xi)\zeta(s)\,ds$$

is P-a.s. continuous on $[0,T]$.
3. To prove that there is a continuous version of

$$\int_0^t S(t-s)\underbrace{DB(X(\xi)(s))\partial X(\xi)\zeta(s)}_{=:\,\Phi(s)}\,dW(s),\quad t\in[0,T],$$

we first want to use the factorization formula 3.12 to get that

$$\int_0^t S(t-s)\Phi(s)\,dW(s)=\frac{\sin\alpha\pi}{\pi}\int_0^t (t-s)^{\alpha-1}S(t-s)Y_\alpha^\Phi(s)\,ds\quad P\text{-a.s.}$$

for all $t\in[0,T]$ where $Y_\alpha^\Phi(s)$, $s\in[0,T]$, is a $\mathcal{B}([0,T])\otimes\mathcal{F}_T$-measurable version of

$$\int_0^s (s-u)^{-\alpha}S(s-u)\Phi(u)\,dW(u),\quad s\in[0,T].$$

To this end we have to know that

$$\int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha}E(\|S(s-u)\Phi(s)\|_{L_2}^2)\,du\Big]^{\frac{1}{2}}\,ds<\infty \tag{4.1}$$

for all $t\in[0,T]$. But this is true since the additional assumption concerning $K:[0,T]\to[0,\infty[$ implies that

$$\begin{aligned}&\int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha}E(\|S(s-u)DB(X(\xi)(s))\\ &\qquad\qquad\qquad\qquad DX(\xi)\zeta(s)\|_{L_2}^2)\,du\Big]^{\frac{1}{2}}\,ds\\ &\le \int_0^t (t-s)^{\alpha-1}\Big[\int_0^s (s-u)^{-2\alpha}K^2(s-u)E(\|DX(\xi)\zeta(s)\|^2)\,du\Big]^{\frac{1}{2}}\,ds\\ &\le \|DX(\xi)\zeta\|_{\mathcal{H}^2(T,H)}\int_0^t (t-s)^{\alpha-1}\Big[\int_0^T u^{-2\alpha}K^2(u)\,du\Big]^{\frac{1}{2}}\,ds<\infty\end{aligned}$$

To prove now that

$$\int_0^t (t-s)^{\alpha-1}S(t-s)Y_\alpha^\Phi(s)\,ds,\quad t\in[0,T],$$

has a continuous version we only have to check that $Y_\alpha^\Phi\in L^p([0,T];H)$ P-a.s. (see Proposition 3.13). But by the help of the Burkholder-Davis-Gundy inequality we can conclude that

$$
\begin{aligned}
&E(\|Y_\alpha^\Phi(t)\|^p) \\
&= E(\|\int_0^t (t-s)^{-\alpha} S(t-s) DB(X(\xi)(s)) DX(\xi)\zeta(s)\, ds\|^p) \\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{p}{2}} \Big(\int_0^t (t-s)^{-2\alpha} \big(E(\|S(t-s)DB(X(\xi)(s))DX(\xi)\zeta(s)\|_{L_2}^p)\big)^{\frac{2}{p}}\, ds\Big)^{\frac{p}{2}} \\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{p}{2}} \Big(\int_0^t (t-s)^{-2\alpha} K^2(t-s) \big(E(\|DX(\xi)\zeta(s)\|^p)\big)^{\frac{2}{p}}\, ds\Big)^{\frac{p}{2}} \\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{p}{2}} \big(\int_0^t (t-s)^{-2\alpha} K^2(t-s) \|DX(\xi)\zeta\|_{\mathcal{H}^p}^2\, ds\big)^{\frac{p}{2}} \\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{p}{2}} \|DX(\xi)\zeta\|_{\mathcal{H}^p}^p \big(\int_0^T s^{-2\alpha} K^2(s)\, ds\big)^{\frac{p}{2}} < \infty
\end{aligned}
$$

Since the process $\varphi : \Omega_T \times [0,T] \to L_2(U,H)$ given by

$$\varphi(u,\omega,s) := 1_{[0,s[}(u)(s-u)^{-\alpha} S(s-u)\Phi(u)$$

fulfills the conditions of the stochastic Fubini Theorem 1.37 (see (4.1)) the process Y_α^Φ has a $\mathcal{B}([0,T]) \otimes \mathcal{F}$- measurable version (see proof of [DaPrZa 92, Theorem 4.18, p.109]). So finally we get by the real Fubini theorem that

$$E(\int_0^T \|Y_\alpha^\Phi(t)\|^p\, dt) = \int_0^T E(\|Y_\alpha^\Phi(t)\|^p)\, dt < \infty$$

□

Chapter 5

Second Order Differentiability of the Mild Solution

As in the previous chapter let $X(\xi)(t)$, $t \in [0,T]$, be the mild solution of problem (2.1) with initial condition ξ. To get the second order differentiability of the mapping $\xi \mapsto X(\xi)$ we make the following assumptions concerning the coefficients F and B.
(For details about the different concepts of differentiability see Appendix D.)

Hypothesis H.2

- F and B are twice Fréchet differentiable where the derivatives

$$D^2F : H \to L(H, L(H))$$
$$D^2B : H \to L(H, L(H, L(U, H)))$$

  are continuous.

- There exists $C_1 > 0$ such that $\|D^2F(x)\|_{L(H,L(H))} \leq C_1$ for all $x \in H$.

- There exists a square integrable mapping $K_1 : [0,T] \to [0, \infty[$ such that

$$\|S(t)D^2B(x)(y)z\|_{L_2} \leq K_1(t)\|y\|\|z\|$$

  for all $t \in]0,T]$ and $x, y, z \in H$.

- The mapping $S(t)D^2B : H \to L(H, L(H, L_2(U,H)))$ is continuous for all $t \in]0,T]$.

Remark 5.1. These conditions differ from those made in [DaPrZa 96, Theorem 5.4.1 (ii), p.69] as we do not demand that the derivative D^2B itself is bounded while the last assumption is additional.

Remark 5.2. If the mappings A, F and B satisfy Hypothesis H.2 they even fulfill Hypothesis H.1' by the fundamental theorem for Bochner integrals.

Theorem 5.3. *Let $q > 2p \geq 4$. Assume that the mappings A, F and B satisfy the Hypotheses H.0, H.1 and H.2.. Then the following statements hold.*

(i) The Fréchet derivative of X

$$DX : L_0^q \to L\big(L_0^q, \mathcal{H}^p(T,H)\big)$$

is Gâteaux differentiable.

(ii) The Gâteaux derivative of $DX : L_0^q \to L(L_0^q, \mathcal{H}^p(T,H))$ fulfills the following equation for all $\xi, \zeta_1, \zeta_2 \in L_0^q$ and $t \in [0,T]$

$$\begin{aligned} &\partial DX(\xi)(\zeta_1)\zeta_2(t) \\ = &\int_0^t S(t-s)DF(X(\xi)(s))\partial DX(\xi)(\zeta_1)\zeta_2(s)\, ds \\ &+ \int_0^t S(t-s)DB(X(\xi)(s))\partial DX(\xi)(\zeta_1)\zeta_2(s)\, dW(s) \\ &+ \int_0^t S(t-s)D^2F(X(\xi)(s))(DX(\xi)\zeta_1(s))DX(\xi)\zeta_2(s)\, ds \\ &+ \int_0^t S(t-s)D^2B(X(\xi)(s))(DX(\xi)\zeta_1(s))DX(\xi)\zeta_2(s)\, dW(s) \quad P\text{-a.s.} \end{aligned}$$

(iii) There exists a constant $\tilde{C}_{T,p,q} > 0$ such that the following inequality holds for all $\zeta_1, \zeta_2 \in L_0^q$

$$\|\partial DX(\xi)(\zeta_1)\zeta_2\|_{\mathcal{H}^p} \leq \tilde{C}_{T,p,q}\|\zeta_1\|_{L^q}\|\zeta_2\|_{L^q}$$

(iv) If $q > 4p \geq 8$ the mapping

$$X : L_0^q \to \mathcal{H}^p(T,H)$$

is twice continuously Fréchet differentiable.
In particular, the mapping

$$\begin{aligned} X : H &\to \mathcal{H}^p(T,H) \\ x &\mapsto X(x) \end{aligned}$$

is twice continuously Fréchet differentiable for all $p \geq 2$.

Before we can prove Theorem 5.3 we need the following lemma

Lemma 5.4. *Assume that the mapping B satisfies the Hypotheses H.0, H.1 and H.2. Then we have for all $t \in]0,T]$ and $x, y \in H$ that*

$$\begin{aligned}&\|\frac{S(t)DB(x+hy)-S(t)DB(x)}{h} - S(t)D^2B(x)(\cdot)y\|_{L(H,L_2)}\\ &\le \frac{1}{h}\int_0^h \|S(t)D^2B(x+sy)(\cdot)y - S(t)D^2B(x)(\cdot)y\|_{L(H,L_2)}\,ds\\ &\le 2K_1(t)\|y\|\end{aligned}$$

In particular

$$\|\frac{S(t)DB(x+hy)-S(t)DB(x)}{h} - S(t)D^2B(x)(\cdot)y\|_{L(H,L_2)} \underset{h\to 0}{\longrightarrow} 0$$

Proof. The proof is analogous to the proof of Lemma 4.5. □

Remark 5.5. The idea of the proof of Theorem 5.3 can be found in [DaPrZa 96, Theorem 5.4.1 (ii), p.69] where the abstract Theorem D.13 (i) (see [Za 98, Theorem 10.4, p.208]) is applied. In this way the authors get the existence of the directional derivatives of second order of the mapping $X : H \to \mathcal{H}^p(T,H)$. The proof is not explained in detail but if one is interested in it it is not hard to modify the proof of our Theorem 5.3 (i) in such a way that it provides the existence of the directional derivatives of second order even in the case that X is considered as a mapping from L_0^{2p} to $\mathcal{H}^p(T,H)$. Since we are interested in twice Fréchet differentiability of the mild solution we added part (ii) and (iii) to Theorem D.13. Especially we consider the case that the initial condition may even be random.

Proof of Theorem 5.3
Idea of the proof of (i): Since $q > 2p \ge 4$ there exists a $q' \in]2p, q[$. To prove the Gâteaux differentiability of

$$DX : L_0^q \to L(L_0^q, \mathcal{H}^p(T,H))$$

we will apply Theorem D.13 (ii) to the mapping $G := \mathcal{F}$ and the spaces $\Lambda_1 := L_0^q$, $\Lambda_0 := L_0^{q'}$, $\Lambda := L_0^p$, $E_0 := \mathcal{H}^{q',\lambda(q')}(T,H)$ and $E := \mathcal{H}^{p,\lambda(p)}(T,H)$ where $\lambda(r) > 0$ such that

$$\mathcal{F} : L_0^r \times \mathcal{H}^{r,\lambda(r)}(T,H) \to \mathcal{H}^{r,\lambda(r)}(T,H), \quad r \ge 2,$$

is a contraction in the second variable.
In this way we get the Gâteaux differentiability of the $L(L_0^q, \mathcal{H}^p(T,H))$-valued mapping DX by the equivalence of the two norms $\| \, \|_{\mathcal{H}^r}$ and $\| \, \|_{r,\lambda(r),T}$.
By the proof of Theorem 4.3 we already know that the conditions 1. and 4. of Theorem D.13 are fulfilled. Therefore it remains to verify that

$$\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q',\lambda(q')}(T,H) \to \mathcal{H}^{p,\lambda(p)}(T,H)$$

is twice continuously Fréchet differentiable in each variable. For simplicity we prove that

$$\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to \mathcal{H}^{p}(T,H)$$

is twice continuously Fréchet differentiable in each variable.
Proof of (i):
Step 1: Let $(q >)q' \geq 2p \geq 4$. We prove the existence of the directional derivatives of $D_1\mathcal{F}$ and $D_2\mathcal{F}$:
(a) Since

$$\begin{aligned} D_1\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q'}(T,H) &\to L\big(L_0^{q'}, \mathcal{H}^p(T,H)\big) \\ (\xi, Y) &\mapsto (S(t))_{t\in[0,T]} \end{aligned}$$

is constant we obtain that

- $\partial_1 D_1\mathcal{F}(\xi, Y; \zeta) \equiv 0 \in L\big(L_0^{q'}, \mathcal{H}^p(T,H)\big)$ for all $\zeta \in L_0^{q'}$
- $\partial_2 D_1\mathcal{F}(\xi, Y; Z_2) \equiv 0 \in L\big(L_0^{q'}, \mathcal{H}^p(T,H)\big)$ for all $Z_2 \in \mathcal{H}^{q'}(T,H)$

(b) Since

$$\begin{aligned} D_2\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q'}(T,H) &\to L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H)) \\ (\xi, Y) &\mapsto \Big(\int_0^t S(t-s)DF(Y(s))\,ds \\ &\qquad + \int_0^t S(t-s)DB(Y(s))\,dW(s)\Big)_{t\in[0,T]} \end{aligned}$$

does not depend on $\xi \in L_0^{q'}$ we get that

$$\partial_1 D_2\mathcal{F}(\xi, Y; \zeta) \equiv 0 \in L\big(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H)\big) \text{ for all } \zeta \in L_0^{q'}.$$

Concerning the differentiability of $D_2\mathcal{F}$ in the second variable $Y \in \mathcal{H}^{q'}(T,H)$ we obtain that

$$
\begin{aligned}
\partial_2 D_2 \mathcal{F}(\xi, Y; Z_2) = \Big(\int_0^t & S(t-s) D^2 F(Y(s))(\cdot) Z_2(s)\, ds \\
& + \int_0^t S(t-s) D^2 B(Y(s))(\cdot) Z_2(s)\, dW(s) \Big)_{t \in [0,T]} \\
& \in L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H)),
\end{aligned}
$$

where the integrals on the right hand side are pointwisely well defined as strong integrals:
Let $Y, Z_1, Z_2 \in \mathcal{H}^{q'}(T,H)$. Then we have for $t \in [0,T]$ and $h \neq 0$ that

$$
\begin{aligned}
& \| \frac{D_2\mathcal{F}(\xi, Y + hZ_2) Z_1(t) - D_2\mathcal{F}(\xi, Y) Z_1(t)}{h} \\
& \quad - \int_0^t S(t-s) D^2 F(Y(s))(Z_1(s)) Z_2(s)\, ds \\
& \quad - \int_0^t S(t-s) D^2 B(Y(s))(Z_1(s)) Z_2(s)\, dW(s) \|_{L^p} \\
& \le \| \int_0^t S(t-s) \Big(\frac{DF(Y(s) + hZ_2(s)) Z_1(s) - DF(Y(s)) Z_1(s)}{h} \\
& \qquad - D^2 F(Y(s))(Z_1(s)) Z_2(s) \Big)\, ds \|_{L^p} \qquad (1.) \\
& \quad + \| \int_0^t S(t-s) \Big(\frac{DB(Y(s) + hZ_2(s)) Z_1(s) - DB(Y(s)) Z_1(s)}{h} \\
& \qquad - D^2 B(Y(s))(Z_1(s)) Z_2(s) \Big)\, dW(s) \|_{L^p} \qquad (2.)
\end{aligned}
$$

(1.) The first summand can be estimated independently of $t \in [0,T]$ in the following way

$$
\begin{aligned}
& \Big(E\big(\| \int_0^t S(t-s) \Big(\frac{DF(Y(s) + hZ_2(s)) Z_1(s) - DF(Y(s)) Z_1(s)}{h} \\
& \qquad - D^2 F(Y(s))(Z_1(s)) Z_2(s) \Big)\, ds \|^p \big) \Big)^{\frac{1}{p}} \\
& \le M_T T^{\frac{p-1}{p}} \Big(\int_0^T E\big(\| \frac{DF(Y(s) + hZ_2(s)) Z_1(s) - DF(Y(s)) Z_1(s)}{h} \\
& \qquad - D^2 F(Y(s))(Z_1(s)) Z_2(s) \|^p \big)\, ds \Big)^{\frac{1}{p}}
\end{aligned}
$$

$$\leq M_T T^{\frac{p-1}{p}} \Big(\int_0^T \Big(E\big(\| \frac{DF(Y(s)+hZ_2(s)) - DF(Y(s))}{h} - D^2F(Y(s))(\cdot)Z_2(s)\|_{L(H)}^{2p} \big) \Big)^{\frac{1}{2}} \Big(E\big(\|Z_1(s)\|^{2p} \big) \Big)^{\frac{1}{2}} \, ds \Big)^{\frac{1}{p}}$$

$$\leq M_T T^{\frac{p-1}{p}} \Big(\int_0^T \Big(E\big(\| \frac{DF(Y(s)+hZ_2(s)) - DF(Y(s))}{h} - D^2F(Y(s))(\cdot)Z_2(s)\|_{L(H)}^{2p} \big) \Big)^{\frac{1}{2}} \, ds \Big)^{\frac{1}{p}} \|Z_1\|_{\mathcal{H}^{q'}}$$

Since F is twice Fréchet differentiable we get that

$$\| \frac{DF(Y(s)+hZ_2(s)) - DF(Y(s))}{h} - D^2F(Y(s))(\cdot)Z_2(s)\|_{L(H)} \underset{h\to 0}{\longrightarrow} 0$$

P-a.s. for all $s \in [0,T]$. In addition, by the fundamental theorem for Bochner integrals, Theorem A.7, we have that

$$\| \frac{DF(x+hy)z - DF(x)z}{h} \| \leq C_1 \|z\| \|y\|,$$

for all $x, y, z \in H$ and therefore we get that

$$\| \frac{DF(Y(s)+hZ_2(s)) - DF(Y(s))}{h} - D^2F(Y(s))(\cdot)Z_2(s)\|_{L(H)}^{2p} \leq (2C_1)^{2p} \|Z_2(s)\|^{2p} \in L^1(\Omega, \mathcal{F}, P)$$

for all $s \in [0,T]$. Hence we obtain by Lebesgue's dominated convergence theorem that

$$\Big(E\big(\| \frac{DF(Y(s)+hZ_2(s)) - DF(Y(s))}{h} - D^2F(Y(s))(\cdot)Z_2(s)\|_{L(H)}^{2p} \big) \Big)^{\frac{1}{2}} \underset{h\to 0}{\longrightarrow} 0$$

for all $s \in [0,T]$. Moreover the above expectation is dominated by $(2C_1)^p \|Z_2\|_{\mathcal{H}^{q'}}^p < \infty$ for all $h \neq 0$ and $s \in [0,T]$. Thus we get again by Lebesgue's dominated convergence theorem that

$$c_h := M_T T^{\frac{p-1}{p}} \Big(\int_0^T \Big(E\big(\| \frac{DF(Y(s)+hZ_2(s)) - DF(Y(s))}{h} - D^2F(Y(s))(\cdot)Z_2(s)\|_{L(H)}^{2p} \big) \Big)^{\frac{1}{2}} \, ds \Big)^{\frac{1}{p}} \underset{h\to 0}{\longrightarrow} 0$$

(2.) Now we want to estimate the second summand independently of $t \in [0,T]$. For that we fix $\lambda > 1$ and by the Burkholder-Davis-Gundy inequality we get the following estimation

$$
\begin{aligned}
&\Big(E\big(\|\int_0^t S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))Z_1(s)-DB(Y(s))Z_1(s)}{h}\\
&\qquad - D^2B(Y(s))(Z_1(s))Z_2(s)\Big)\,dW(s)\|^p\big)\Big)^{\frac{1}{p}}\\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\\
&\quad\Big[\int_0^{\frac{t}{\lambda}}\Big(E\big(\|S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))Z_1(s)-DB(Y(s))Z_1(s)}{h}\\
&\qquad - D^2B(Y(s))(Z_1(s))Z_2(s)\Big)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds\\
&\quad+\int_{\frac{t}{\lambda}}^t\Big(E\big(\|S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))Z_1(s)-DB(Y(s))Z_1(s)}{h}\\
&\qquad - D^2B(Y(s))(Z_1(s))Z_2(s)\Big)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds\Big]^{\frac{1}{2}}\\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\\
&\quad\Big[M_T^2\int_0^T\Big(E\big(\|S((\lambda-1)s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h}\\
&\qquad - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|_{L(H,L_2)}^p\|Z_1(s)\|^p\big)\Big)^{\frac{2}{p}}\,ds\\
&\quad+\int_{\frac{t}{\lambda}}^t\Big(E\big(\|S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h}\\
&\qquad - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|_{L(H,L_2)}^p\|Z_1(s)\|^p\big)\Big)^{\frac{2}{p}}\,ds\Big]^{\frac{1}{2}}\\
&\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\|Z_1\|_{\mathcal{H}^{q'}}\\
&\quad\Big[M_T^2\int_0^T\Big(E\big(\|S((\lambda-1)s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h}\\
&\qquad - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|_{L(H,L_2)}^{2p}\big)\Big)^{\frac{1}{p}}\,ds\\
&\quad+\int_{\frac{t}{\lambda}}^t\Big(E\big(\|S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h}\\
&\qquad - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|_{L(H,L_2)}^{2p}\big)\Big)^{\frac{1}{p}}\,ds\Big]^{\frac{1}{2}}
\end{aligned}
$$

With regard to the first summand we get by Lemma 5.4 that

$$\|S((\lambda-1)s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h} - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|^{2p}_{L(H,L_2)} \underset{h\to 0}{\longrightarrow} 0 \quad P\text{-a.s.}$$

for all $s \in]0,T]$. Moreover, again by Lemma 5.4, the term is dominated by $2^{2p}K_1^{2p}((\lambda-1)s)\|Z_2(s)\|^{2p} \in L^1(\Omega,\mathcal{F},P)$. Therefore we obtain by Lebesgue's dominated convergence theorem that

$$\Big(E\Big(\|S((\lambda-1)s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h} - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|^{2p}_{L(H,L_2)}\Big)\Big)^{\frac{1}{p}} \underset{h\to 0}{\longrightarrow} 0$$

For all $h \neq 0$ and $s \in]0,T]$ the above expectation is bounded by the function $4K_1^2((\lambda-1)\cdot)\|Z_2\|^2_{\mathcal{H}^{q'}} \in L^1([0,T],dx)$. Therefore we can conclude again by Lebesgue's dominated convergence theorem that

$$\int_0^T \Big(E\Big(\|S((\lambda-1)s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h} - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|^{2p}_{L(H,L_2)}\Big)\Big)^{\frac{1}{p}}\, ds \underset{h\to 0}{\longrightarrow} 0$$

For λ near to 1 the second summand becomes small independently of $h \neq 0$ and $t \in [0,T]$ as we have by Lemma 5.4 that

$$\begin{aligned}
&\int_{\frac{t}{\lambda}}^t \Big(E\Big(\|S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))-DB(Y(s))}{h} - D^2B(Y(s))(\cdot)Z_2(s)\Big)\|^{2p}_{L(H,L_2)}\Big)\Big)^{\frac{1}{p}}\, ds \\
&\leq \int_{\frac{t}{\lambda}}^t 4K_1^2(t-s)\Big(E\Big(\|Z_2(s)\|^{2p}\Big)\Big)^{\frac{1}{p}}\, ds \\
&\leq 4\int_0^{\frac{(\lambda-1)T}{\lambda}} K_1^2(s)\, ds\, \|Z_2\|^2_{\mathcal{H}^{q'}} \underset{\lambda\downarrow 1}{\longrightarrow} 0
\end{aligned}$$

Therefore we get that

$$\Big(E\Big(\|\int_0^t S(t-s)\Big(\frac{DB(Y(s)+hZ_2(s))Z_1(s)-DB(Y(s))Z_1(s)}{h} - D^2B(Y(s))(Z_1(s))Z_2(s)\Big)\, dW(s)\|^p\Big)\Big)^{\frac{1}{p}} \leq \tilde{c}_h\|Z_1\|_{\mathcal{H}^{q'}}$$

where $\tilde{c}_h \underset{h\to 0}{\longrightarrow} 0$. Altogether this implies that

$$\sup_{t\in[0,T]} \| \frac{D_2\mathcal{F}(\xi, Y+hZ_2)Z_1(t) - D_2\mathcal{F}(\xi,Y)Z_1(t)}{h} \\ - \int_0^t S(t-s)D^2F(Y(s))(Z_1(s))Z_2(s)\,ds \\ - \int_0^t S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\,dW(s)\|_{L^p} \\ \le (c_h+\tilde{c}_h)\|Z_1\|_{\mathcal{H}^{q'}} \text{ where } c_h+\tilde{c}_h \underset{h\to 0}{\longrightarrow} 0.$$

Step 2: We prove that the directional derivatives are the Fréchet derivatives in the case that $(q>)q' > 2p \ge 4$:
Since

$$\partial_1 D_1\mathcal{F}(\xi,Y;\zeta) \equiv 0 \in L(L_0^{q'}, \mathcal{H}^p(T,H)), \\ \partial_2 D_1\mathcal{F}(\xi,Y;Z_2) \equiv 0 \in L(L_0^{q'}, \mathcal{H}^p(T,H)) \\ \text{and } \partial_1 D_2\mathcal{F}(\xi,Y;\zeta) \equiv 0 \in L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))$$

for all $\xi,\zeta \in L_0^{q'}$ and $Y, Z_2 \in \mathcal{H}^{q'}(T,H)$ it remains to verify that $\partial_2 D_2\mathcal{F} = D_2^2\mathcal{F}$. Proposition D.6 provides that it is enough to check that

(a) $\partial_2 D_2\mathcal{F}(\xi,Y;\cdot) \in L(\mathcal{H}^{q'}(T,H), L(\mathcal{H}^{q'}(T,H),\mathcal{H}^p(T,H)))$ for $\xi \in L_0^{q'}$ and $Y \in \mathcal{H}^{q'}(T,H)$

(b) $\partial_2 D_2\mathcal{F}(\xi,\cdot) : \mathcal{H}^{q'}(T,H) \to L(\mathcal{H}^{q'}(T,H), L(\mathcal{H}^{q'}(T,H),\mathcal{H}^p(T,H)))$ is continuous for $\xi \in L_0^{q'}$.

(a) Since $\partial D_2\mathcal{F}(\xi,Y;\cdot)$ is given pointwisely by

$$\partial_2 D_2\mathcal{F}(\xi,Y;Z_2)Z_1 = (\int_0^t S(t-s)D^2F(Y(s))(Z_1(s))Z_2(s)\,ds \\ + \int_0^t S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\,dW(s))_{t\in[0,T]},$$

for all $Z_1, Z_2 \in \mathcal{H}^{q'}(T,H)$ the linearity in Z_2 is obvious.
Moreover we have for $\xi \in L_0^{q'}$, $Y, Z_1, Z_2 \in \mathcal{H}^{q'}(T,H)$ and $t\in[0,T]$

$$\|\partial_2 D_2\mathcal{F}(\xi,Y;Z_2)Z_1\|_{L^p} \\ \le M_T T^{\frac{p-1}{p}} \Big(E\big(\int_0^T \|D^2F(Y(s))(Z_1(s))Z_2(s)\|^p\,ds\big)\Big)^{\frac{1}{p}}$$

$$+\left(\frac{p}{2}(p-1)\right)^{\frac{1}{2}}\Big[\int_0^t \Big(E\big(\|S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds\Big]^{\frac{1}{2}}$$

$$\leq M_T T^{\frac{p-1}{p}} C_1\Big(\int_0^T \Big(E\big(\|Z_1(s)\|^{2p}\big)\Big)^{\frac{1}{2}}\Big(E\big(\|Z_2(s)\|^{2p}\big)\Big)^{\frac{1}{2}}\,ds\Big)^{\frac{1}{p}}$$

$$+\left(\frac{p}{2}(p-1)\right)^{\frac{1}{2}}\Big[\int_0^t K_1^2(t-s)\Big(E\big(\|Z_1(s)\|^{2p}\big)\Big)^{\frac{1}{p}}\Big(E\big(\|Z_2(s)\|^{2p}\big)\Big)^{\frac{1}{p}}\,ds\Big]^{\frac{1}{2}}$$

$$\leq \Big[M_T T C_1+\left(\frac{p}{2}(p-1)\right)^{\frac{1}{2}}\Big(\int_0^T K_1^2(s)\,ds\Big)^{\frac{1}{2}}\Big]\|Z_1\|_{\mathcal{H}^{q'}}\|Z_2\|_{\mathcal{H}^{q'}}$$

Hence the mapping $Z_2 \mapsto \partial_2 D_2\mathcal{F}(\xi, Y; Z_2) \in L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))$ is linear and continuous. Therefore we write $\partial_2 D_2\mathcal{F}(\xi,Y)(\cdot)Z_2$ instead of $\partial_2 D_2\mathcal{F}(\xi,Y;Z_2)$ where $\partial_2 D_2\mathcal{F}(\xi,Y)$ is the Gâteaux derivative of $D_2\mathcal{F}(\xi,\cdot) : \mathcal{H}^{q'}(T,H) \to L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))$.

(b) Let now $\xi \in L_0^{q'}$ and $Y, Y_n, Z_1, Z_2 \in \mathcal{H}^{q'}(T,H)$ such that $Y_n \longrightarrow Y$ in $\mathcal{H}^{q'}(T,H)$ as $n \to \infty$. Then we have for $t \in [0,T]$ that

$$\|\partial_2 D_2\mathcal{F}(\xi,Y_n)(Z_1)Z_2(t) - \partial_2 D_2\mathcal{F}(\xi,Y)(Z_1)Z_2(t)\|_{L^p}$$

$$\leq \|\int_0^t S(t-s)D^2F(Y_n(s))(Z_1(s))Z_2(s) - S(t-s)D^2F(Y(s))(Z_1(s))Z_2(s)\,ds\|_{L^p} \qquad (1.)$$

$$+\|\int_0^t S(t-s)D^2B(Y_n(s))(Z_1(s))Z_2(s) - S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\,dW(s)\|_{L^p} \qquad (2.)$$

(1.) At first we want to estimate the first summand independently of $t \in [0,T]$:

$$\Big(E\big(\|\int_0^t S(t-s)(D^2F(Y_n(s))(Z_1(s))Z_2(s) - D^2F(Y(s))(Z_1(s))Z_2(s))\,ds\|^p\big)\Big)^{\frac{1}{p}}$$

$$\leq M_T T^{\frac{p-1}{p}}\Big[E\Big(\int_0^T \|D^2F(Y_n(s))(Z_1(s))Z_2(s) - D^2F(Y(s))(Z_1(s))Z_2(s)\|^p\,ds\Big)\Big]^{\frac{1}{p}}$$

$$\leq M_T T^{\frac{p-1}{p}}\Big[\int_0^T E\big(\|D^2F(Y_n(s)) - D^2F(Y(s))\|_{L(H,L(H))}^p\|Z_1(s)\|^p\|Z_2(s)\|^p\big)\,ds\Big]^{\frac{1}{p}}$$

$$\leq M_T T^{\frac{p-1}{p}}\Big[\int_0^T \Big(E\big(\|D^2F(Y_n(s)) - D^2F(Y(s))\|_{L(H,L(H))}^{\frac{q'p}{q'-2p}}\big)\Big)^{\frac{q'-2p}{q'}}\Big(E\big(\|Z_1(s)\|^{\frac{q'}{2}}\|Z_2(s)\|^{\frac{q'}{2}}\big)\Big)^{\frac{2p}{q'}}\,ds\Big]^{\frac{1}{p}}$$

(Hölder inequality for $\frac{q'}{2p} > 1$ and $\frac{q'}{q'-2p}$)

$$\leq M_T T^{\frac{p-1}{p}} \Big[\int_0^T \Big(E\big(\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))}^{\frac{q'p}{q'-2p}} \big) \Big)^{\frac{q'-2p}{q'}} \Big(E\big(\|Z_1(s)\|^{q'} \big) \Big)^{\frac{p}{q'}} \Big(E\big(\|Z_2(s)\|^{q'} \big) \Big)^{\frac{p}{q'}} \, ds \Big]^{\frac{1}{p}}$$

(Hölder inequality for 2)

$$\leq M_T T^{\frac{p-1}{p}} \Big[\int_0^T \Big(E\big(\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))}^{\frac{q'p}{q'-2p}} \big) \Big)^{\frac{q'-2p}{q'}} \, ds \Big]^{\frac{1}{p}} \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}}$$

The sequence $Y_n(s)$, $n \in \mathbb{N}$, converges to $Y(s)$ in probability for all $s \in [0,T]$ and $D^2F : H \to L\big(H, L(H)\big)$ is continuous.
Therefore we get by Lemma 4.6 that

$$\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))} \underset{n\to\infty}{\longrightarrow} 0 \quad \text{in probability}$$

Moreover

$$\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))}^{\frac{q'p}{q'-2p}} \leq (2C_1)^{\frac{q'p}{q'-2p}} \text{ for all } n \in \mathbb{N}$$

which implies that the family $\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))}^{\frac{q'p}{q'-2p}}$, $n \in \mathbb{N}$, is uniformly integrable. Hence we obtain that

$$\Big(E\big(\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))}^{\frac{q'p}{q'-2p}} \big) \Big)^{\frac{q'-2p}{q'}} \underset{n\to\infty}{\longrightarrow} 0 \text{ for all } s \in [0,T].$$

Since the above expectation is bounded by $(2C_1)^p$ we get by Lebesgue's dominated convergence theorem that

$$a_n := M_T T^{\frac{p-1}{p}} \Big[\int_0^T \Big(E\big(\| D^2F(Y_n(s)) - D^2F(Y(s)) \|_{L(H,L(H))}^{\frac{q'p}{q'-2p}} \big) \Big)^{\frac{q'-2p}{q'}} \, ds \Big]^{\frac{1}{p}} \underset{n\to\infty}{\longrightarrow} 0$$

(2.) Now we want to estimate the second summand independently of $t \in [0,T]$. To this end we fix $\lambda > 0$ and obtain by the Burkholder-Davis-Gundy inequality that

$$\Big(E\big(\Big\| \int_0^t S(t-s) D^2B(Y_n(s))(Z_1(s))Z_2(s) - S(t-s) D^2B(Y(s))(Z_1(s))Z_2(s) \, dW(s) \Big\|^p \big) \Big)^{\frac{1}{p}}$$

$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}\Big[\int_0^{\frac{t}{\lambda}}\Big(E\big(\| S(t-s)D^2B(Y_n(s))(Z_1(s))Z_2(s) - S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds$$
$$+\int_{\frac{t}{\lambda}}^t\Big(E\big(\| S(t-s)D^2B(Y_n(s))(Z_1(s))Z_2(s) - S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds\Big]^{\frac{1}{2}}$$
$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}$$
$$\Big[M_T^2\int_0^T\Big(E\big(\| S((\lambda-1)s)D^2B(Y_n(s))(Z_1(s))Z_2(s) - S((\lambda-1)s)D^2B(Y(s))(Z_1(s))Z_2(s)\|_{L_2}^p\big)\Big)^{\frac{2}{p}}\,ds$$
$$+\int_{\frac{t}{\lambda}}^t 4K_1^2(t-s)\Big(E\big(\|Z_1(s)\|^p\|Z_2(s)\|^p\big)\Big)^{\frac{2}{p}}\,ds\Big]^{\frac{1}{2}}$$
$$\leq \Big(\frac{p}{2}(p-1)\Big)^{\frac{1}{2}}(M_T^2+1)^{\frac{1}{2}}\|Z_1\|_{\mathcal{H}^{q'}}\|Z_2\|_{\mathcal{H}^{q'}}$$
$$\Big[\int_0^T\Big(E\big(\|S((\lambda-1)s)[D^2B(Y_n(s))-D^2B(Y(s))]\|_{L(H,L(H,L_2))}^{\frac{q'p}{q'-2p}}\big)\Big)^{\frac{2(q'-2p)}{q'p}}\,ds$$
$$+\int_0^{\frac{(\lambda-1)T}{\lambda}} 4K_1^2(s)\,ds\Big]^{\frac{1}{2}}$$

where we used the Hölder inequality for $\frac{q'}{2p}$ and then for 2 as in (1.).
For each $s \in [0,T]$ the sequence $Y_n(s)$, $n \in \mathbb{N}$, converges to $Y(s)$ in probability and the mapping $S((\lambda-1)s)D^2B : H \to L\big(H, L(H, L_2(U,H))\big)$ is continuous for all $s \in]0,T]$ by our additional assumption. Therefore we can conclude by Lemma 4.6 that

$$\|S((\lambda-1)s)[D^2B(Y_n(s)) - D^2B(Y(s))]\|_{L(H,L(H,L_2))} \underset{n\to\infty}{\longrightarrow} 0$$

in probability for all $s \in]0,T]$. Moreover we have for all $s \in]0,T]$ and all $n \in \mathbb{N}$ that

$$\|S((\lambda-1)s)[D^2B(Y_n(s)) - D^2B(Y(s))]\|_{L(H,L(H,L_2))}^{\frac{q'p}{q'-2p}} \leq (2K_1((\lambda-1)s))^{\frac{q'p}{q'-2p}}$$

Hence the family $\|S((\lambda-1)s)[D^2B(Y_n(s)) - D^2B(Y(s))]\|_{L(H,L(H,L_2))}^{\frac{q'p}{q'-2p}}$, $n \in \mathbb{N}$, is uniformly integrable which implies that

$$\Big(E\big(\|S((\lambda-1)s)[D^2B(Y_n(s)) - D^2B(Y(s))]\|_{L(H,L(H,L_2))}^{\frac{q'p}{q'-2p}}\big)\Big)^{\frac{2(q'-2p)}{q'p}} \underset{n\to\infty}{\longrightarrow} 0$$

for all $s \in]0,T]$. Since for all $n \in \mathbb{N}$ the above expectation is dominated by the function $4K_1^2((\lambda-1)\cdot) \in L^1([0,T],dx)$ we get by Lebesgue's dominated convergence theorem that

$$\int_0^T \Big(E\big(\| S((\lambda-1)s)[D^2B(Y_n(s)) - D^2B(Y(s))]\|_{L(H,L(H,L_2))}^{\frac{q'p}{q'-2p}}\big)\Big)^{\frac{2(q'-2p)}{q'p}} ds \underset{n\to\infty}{\longrightarrow} 0$$

Moreover we know that $\int_0^{\frac{(\lambda-1)T}{\lambda}} K_1^2(s)\, ds \underset{\lambda\downarrow 1}{\longrightarrow} 0$ so that we get the existence of a sequence b_n, $n \in \mathbb{N}$, such that $b_n \longrightarrow 0$ as $n \to \infty$ and

$$\begin{aligned} &\Big(E\big(\|\int_0^t S(t-s)D^2B(Y_n(s))(Z_1(s))Z_2(s) \\ &\qquad - S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\, dW(s)\|^p\big)\Big)^{\frac{1}{p}} \\ &\le b_n \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}} \end{aligned}$$

for all $t \in [0,T]$. Altogether we have that

$$\|\partial_2 D_2\mathcal{F}(\xi,Y_n)(Z_1)Z_2 - \partial_2 D_2\mathcal{F}(\xi,Y)(Z_1)Z_2\|_{\mathcal{H}^p} \le (a_n+b_n)\|Z_1\|_{\mathcal{H}^{q'}}\|Z_2\|_{\mathcal{H}^{q'}}$$

where $a_n + b_n \longrightarrow 0$ as $n \to \infty$. That implies that the Gâteaux derivative of $D_2\mathcal{F}(\xi,\cdot)$ is the Fréchet derivative and therefore it is justified to write $D_2^2\mathcal{F}$ instead of $\partial_2 D_2\mathcal{F}$.

Proof of (ii): Let $\xi, \zeta_1, \zeta_2 \in L_0^q$. Then by Theorem D.13 (i) we have the following representation for the Gâteaux derivative of DX:

$$\partial DX(\xi)(\zeta_1)\zeta_2 = [I - \partial_2\mathcal{F}(\xi, X(\xi))]^{-1} D_2^2\mathcal{F}(\xi, X(\xi))(DX(\xi)\zeta_1)DX(\xi)\zeta_2$$

and therefore

$$\begin{aligned} &\partial DX(\xi)(\zeta_1)\zeta_2 \\ &= \partial_2\mathcal{F}(\xi,X(\xi))\partial DX(\xi)(\zeta_1)\zeta_2 + D_2^2\mathcal{F}(\xi,X(\xi))(DX(\xi)\zeta_1)DX(\xi)\zeta_2 \\ &= \Big(\int_0^t S(t-s)DF(X(\xi)(s))\partial DX(\xi)(\zeta_1)\zeta_2(s)\, ds \\ &\quad + \int_0^t S(t-s)DB(X(\xi)(s))\partial DX(\xi)(\zeta_1)\zeta_2(s)\, dW(s) \\ &\quad + \int_0^t S(t-s)D^2F(X(\xi)(s))(DX(\xi)\zeta_1(s))DX(\xi)\zeta_2(s)\, ds \\ &\quad + \int_0^t S(t-s)D^2B(X(\xi)(s))(DX(\xi)\zeta_1(s))DX(\xi)\zeta_2(s)\, dW(s)\Big)_{t\in[0,T]} \end{aligned}$$

Proof of (iii): Let $q' \in]2p, q[$. We will apply Corollary D.14 (i) to the spaces $\Lambda_1 := L_0^q$, $\Lambda_0 := L_0^{q'}$, $\Lambda := L_0^p$, $E_0 := \mathcal{H}^{q',\lambda(q')}(T,H)$ and $E := \mathcal{H}^{p,\lambda(p)}(T,H)$. By Theorem 4.3 (iii) we know that $DX : L_0^q \to L(L_0^q, \mathcal{H}^{q'}(T,H))$ is bounded. Therefore it remains to show that

$$D_2^2\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L\big(\mathcal{H}^{q'}(T,H), L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))\big)$$

is bounded.
To this end let $\xi \in L_0^{q'}$ and $Y, Z_1, Z_2 \in \mathcal{H}^{q'}(T,H)$ then we have for all $t \in [0,T]$

$$\begin{aligned}
&\|D_2^2\mathcal{F}(\xi,Y)(Z_1)Z_2(t)\|_{L^p} \\
\le\ &\|\int_0^t S(t-s)D^2F(Y(s))(Z_1(s))Z_2(s)\ ds\|_{L^p} && (1.)\\
&+\|\int_0^t S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\ dW(s)\|_{L^p} && (2.)
\end{aligned}$$

(1.) The first summand can be estimated independently of $t \in [0,T]$ as follows

$$\|\int_0^t S(t-s)D^2F(Y(s))(Z_1(s))Z_2(s)\ ds\|_{L^p} \le M_T T C_1 \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}}$$

(2.) By the Burkholder-Davis-Gundy inequality we have for all $t \in [0,T]$ that

$$\begin{aligned}
&\|\int_0^t S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\ dW(s)\|_{L^p} \\
\le\ &\big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\Big(\int_0^t \Big(E\big(\|S(t-s)D^2B(Y(s))(Z_1(s))Z_2(s)\|^p\big)\Big)^{\frac{2}{p}}\ ds\Big)^{\frac{1}{2}} \\
\le\ &\big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\Big(\int_0^T K_1^2(s)\ ds\Big)^{\frac{1}{2}}\|Z_1\|_{\mathcal{H}^{q'}}\|Z_2\|_{\mathcal{H}^{q'}}
\end{aligned}$$

(1.) and (2.) provide the result

$$\begin{aligned}
&\|D_2^2\mathcal{F}(\xi,Y)(Z_1)Z_2\|_{\mathcal{H}^p} \\
\le\ &\underbrace{\Big(M_T T C_1 + \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}}\Big(\int_0^T K_1^2(s)\ ds\Big)^{\frac{1}{2}}\Big)}_{<\infty}\|Z_1\|_{\mathcal{H}^{q'}}\|Z_2\|_{\mathcal{H}^{q'}}
\end{aligned}$$

Proof of (iv): If $q > 4p \ge 8$ then there exists $q' \in]2p, \frac{q}{2}[$.

We apply Theorem D.13 (iii) to the spaces $\Lambda_1 := L_0^q$, $\Lambda_0 := L_0^{q'}$, $\Lambda := L_0^p$, $E_0 := \mathcal{H}^{q',\lambda(q')}(T,H)$ and $E := \mathcal{H}^{p,\lambda(p)}(T,H)$.
Since $q > q' > 2p$ all conditions of Theorem D.13 (ii) are fulfilled. Moreover $q > 2q'$. Therefore we obtain by (i) that $DX : L_0^q \to L(L_0^q, \mathcal{H}^{q',\lambda(q')}(T,H))$ is Gâteaux differentiable with derivative

$$\partial DX : L_0^q \to L(L_0^q, L(L_0^q, \mathcal{H}^{q',\lambda(q')}(T,H))).$$

Hence we get by Theorem D.13 (iii) that

$$X : L_0^q \to \mathcal{H}^p(T,H)$$

is twice Fréchet differentiable. □

As in the previous chapter where we gave conditions under which DX is uniformly continuous we are now interested in conditions under which D^2X is uniformly continuous. We will use these results in chapter 6.

Hypothesis H.2′

- $D^2F : H \to L(H, L(H))$ is uniformly continuous.
- There exists a mapping $\tilde{K}_1 : [0,T] \to [0,\infty[$, $\tilde{K}_1 \in L^2([0,T], dx)$, such that

  $$\|S(t)\big(D^2B(x) - D^2B(y)\big)\|_{L(H,L(H,L_2))} \le \tilde{K}_1(t)\|x-y\|$$

  for all $t \in]0,T]$ and $x, y \in H$.

Corollary 5.6. *Assume that the mappings A, F and B satisfy the Hypotheses H.0, H.1, H.2 and H.2' and let $q > 6p \geq 12$.*
Then

$$D^2X : L_0^q \to L(L_0^q, L(L_0^q, \mathcal{H}^p(T,H)))$$

is uniformly continuous.

Proof. Let $q' = 3p$. We want to apply Corollary D.14 to the spaces $\Lambda_1 := L_0^q$, $\Lambda_0 := L_0^{q'}$, $\Lambda := L_0^p$, $E_0 := \mathcal{H}^{q',\lambda(q')}(T,H)$ and $E := \mathcal{H}^{p,\lambda(p)}(T,H)$.
Since $q' \geq 2p$ we already know by the proof of Corollary 4.7 that

$$D_2\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))$$

is uniformly continuous. Moreover we have by Theorem 4.3 (iii) and by Corollary 4.7 that

$$DX : L_0^q \to L(L_0^q, \mathcal{H}^{q'}(T,H))$$

is bounded and uniformly continuous since $q \geq 2q'$. Moreover $\partial DX : L_0^q \to L(L_0^q, L(L_0^q, \mathcal{H}^{q'}(T,H)))$ is bounded by Theorem 5.3 (iii) since $q > 2q'$.
Therefore it remains to check the boundedness and uniform continuity of the mappings

$$\begin{aligned} D_1^2\mathcal{F} &: L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L(L_0^{q'}, L(L_0^{q'}, \mathcal{H}^p(T,H))) \\ D_1D_2\mathcal{F} &: L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L(L_0^{q'}, L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))) \\ D_2D_1\mathcal{F} &: L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L(\mathcal{H}^{q'}(T,H), L(L_0^{q'}, \mathcal{H}^p(T,H))) \\ D_2^2\mathcal{F} &: L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L(\mathcal{H}^{q'}(T,H), L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))) \end{aligned}$$

Since $D_1^2\mathcal{F}$, $D_1D_2\mathcal{F}$ and $D_2D_1\mathcal{F}$ are constant we only have to prove the boundedness and uniform continuity of $D_2^2\mathcal{F}$. Since $q' > 2p$ the boundedness is clear by the proof of Theorem 5.3 (iii).
To prove the uniform continuity let $\xi, \tilde{\xi} \in L_0^{q'}$ and $Y, \tilde{Y}, Z_1, Z_2 \in \mathcal{H}^{q'}(T,H)$. Then we have for all $t \in [0,T]$:

$$\begin{aligned} &\|D_2^2\mathcal{F}(\xi,Y)(Z_1)Z_2(t) - D_2^2\mathcal{F}(\tilde{\xi},\tilde{Y})(Z_1)Z_2(t)\|_{L^p} \\ &\leq \|\int_0^t S(t-s)\big[D^2F(Y(s)) - D^2F(\tilde{Y}(s))\big](Z_1(s))Z_2(s)\,ds\|_{L^p} \qquad (1.) \\ &\quad + \|\int_0^t S(t-s)\big[D^2B(Y(s)) - D^2B(\tilde{Y}(s))\big](Z_1(s))Z_2(s)\,ds\|_{L^p} \qquad (2.) \end{aligned}$$

(1.) Let $\varepsilon > 0$. Since $D^2F : H \to L(H, L(H))$ is uniformly continuous there exists $\delta > 0$ such that

$$\|D^2F(x) - D^2F(y)\|_{L(H,L(H))} \leq \varepsilon \ \text{ for all } x, y \in H \text{ with } \|x-y\| \leq \delta.$$

Using this fact we obtain for all $t \in [0,T]$ that

$$\begin{aligned} &\|\int_0^t S(t-s)\big[D^2F(Y(s)) - D^2F(\tilde{Y}(s))\big](Z_1(s))Z_2(s)\,ds\|_{L^p} \\ &\leq M_T T^{\frac{p-1}{p}}\Big(\int_0^T E\big(\|D^2F(Y(s)) - D^2F(\tilde{Y}(s))\|^p_{L(H,L(H))} \\ &\qquad \|Z_1(s)\|^p\|Z_2(s)\|^p\big)\,ds\Big)^{\frac{1}{p}} \end{aligned}$$

$$\leq M_T T^{\frac{p-1}{p}} \Big(\int_0^T \Big(E\big(\|D^2F(Y(s)) - D^2F(\tilde{Y}(s))\|^{3p}_{L(H,L(H))}\big)\Big)^{\frac{1}{3}} \, ds \Big)^{\frac{1}{p}}$$
$$\|Z_1\|_{\mathcal{H}^{3p}} \|Z_2\|_{\mathcal{H}^{3p}}$$
$$\leq M_T T^{\frac{p-1}{p}} \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}}$$
$$\Big(\int_0^T \Big(E\big(\|D^2F(Y(s)) - D^2F(\tilde{Y}(s))\|^{3p}_{L(H,L(H))} 1_{\{\|Y(s)-\tilde{Y}(s)\| \leq \delta\}}\big)\Big)^{\frac{1}{3}}$$
$$+ \Big(E\big(\|D^2F(Y(s)) - D^2F(\tilde{Y}(s))\|^{3p}_{L(H,L(H))} 1_{\{\|Y(s)-\tilde{Y}(s)\| > \delta\}}\big)\Big)^{\frac{1}{3}} \, ds\Big)^{\frac{1}{p}}$$
$$\leq M_T T^{\frac{p-1}{p}} \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}} \Big(\int_0^T \varepsilon^p + (2C_1)^p \Big(\frac{1}{\delta^3} E\big(\|Y(s) - \tilde{Y}(s)\|^3\big)\Big)^{\frac{1}{3}} \, ds\Big)^{\frac{1}{p}}$$
$$\leq M_T T \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}} \Big(\varepsilon^p + (2C_1)^p \frac{1}{\delta} \|Y - \tilde{Y}\|_{\mathcal{H}^3}\Big)^{\frac{1}{p}}$$

(2.) By the Burkholder-Davis-Gundy inequality we get for all $t \in [0,T]$ that

$$\Big(E\big(\| \int_0^t S(t-s) \big[D^2B(Y(s)) - D^2B(\tilde{Y}(s))\big] (Z_1(s)) Z_2(s) \, dW(s)\|^p\big)\Big)^{\frac{1}{p}}$$
$$\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}} \Big(\int_0^t \tilde{K}_1^2(t-s) \Big(E\big(\|Y(s) - \tilde{Y}(s)\|^p \|Z_1(s)\|^p \|Z_2(s)\|^p\big)\Big)^{\frac{2}{p}} \, ds \Big)^{\frac{1}{2}}$$
$$\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}} \Big(\int_0^t \tilde{K}_1^2(t-s) \, \Big(E\big(\|Y(s) - \tilde{Y}(s)\|^{3p}\big)\Big)^{\frac{2}{3p}}$$
$$\Big(E\big(\|Z_1(s)\|^{3p}\big)\Big)^{\frac{2}{3p}} \Big(E\big(\|Z_2(s)\|^{3p}\big)\Big)^{\frac{2}{3p}} \, ds\Big)^{\frac{1}{2}}$$
$$\leq \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}} \Big(\int_0^T \tilde{K}_1^2(s) \, ds\Big)^{\frac{1}{2}} \|Y - \tilde{Y}\|_{\mathcal{H}^{q'}} \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}}$$

Since these estimates are independent of $t \in [0,T]$ we finally have that

$$\|D_2^2\mathcal{F}(\xi, Y)(Z_1)Z_2 - D_2^2\mathcal{F}(\tilde{\xi}, \tilde{Y})(Z_1)Z_2\|_{\mathcal{H}^p}$$
$$\leq \Big[M_T T \Big(\varepsilon^p + (2C_1)^p \frac{1}{\delta} \|Y - \tilde{Y}\|_{\mathcal{H}^3}\Big)^{\frac{1}{p}}$$
$$+ \big(\frac{p}{2}(p-1)\big)^{\frac{1}{2}} \Big(\int_0^T \tilde{K}_1^2(s) \, ds\Big)^{\frac{1}{2}} \|Y - \tilde{Y}\|_{\mathcal{H}^{q'}} \Big] \|Z_1\|_{\mathcal{H}^{q'}} \|Z_2\|_{\mathcal{H}^{q'}}$$

and therefore the uniform continuity of the mapping

$$D_2^2\mathcal{F} : L_0^{q'} \times \mathcal{H}^{q'}(T,H) \to L\big(\mathcal{H}^{q'}(T,H), L(\mathcal{H}^{q'}(T,H), \mathcal{H}^p(T,H))\big)$$

□

To prove the existence of a strict solution of the Kolmogorov equation 6.1 related to problem (2.1) (see chapter 6) we need the pathwise continuity of $X(x)(t)$, $DX(x)y(t)$ and $D^2X(x)(y)z(t)$, $t \in [0,T]$, $x, y, z \in H$. In chapter 3 and 4 we already gave conditions under which there exist continuous versions of $X(x)(t)$ and $DX(x)y(t)$, $t \in [0,T]$. The assumption we make to get a continuous version of $D^2X(x)(y)z(t)$, $t \in [0,T]$, is analogous to those made in the previous cases.

Proposition 5.7. *Assume that the mappings A, F and B satisfy the Hypotheses H.0, H.1 and H.2.*
If there exists $\alpha \in]0, \frac{1}{2}[$ such that

$$\int_0^T s^{-2\alpha} K_1^2(s)\, ds < \infty$$

then the process $\partial DX(\xi)(\zeta_1)\zeta_2(t)$, $t \in [0,T]$, has a continuous version for all $\xi, \zeta_1, \zeta_2 \in L_0^q$, $q > 2p \geq 4$.

Proof. If $\xi, \zeta_1, \zeta_2 \in L_0^q$ we know by Theorem 5.3 (ii) that the following equation holds for all $t \in [0,T]$.

$$\begin{aligned}
&\partial DX(\xi)(\zeta_1)\zeta_2(t)\\
&= \int_0^t S(t-s)DF(X(\xi)(s))\partial DX(\xi)(\zeta_1)\zeta_2(s)\, ds\\
&\quad + \int_0^t S(t-s)DB(X(\xi)(s))\partial DX(\xi)(\zeta_1)\zeta_2(s)\, dW(s)\\
&\quad + \int_0^t S(t-s)D^2F(X(\xi)(s))(DX(\xi)\zeta_1(s))DX(\xi)\zeta_2(s)\, ds\\
&\quad + \int_0^t S(t-s)D^2B(X(\xi)(s))(DX(\xi)\zeta_1(s))DX(\xi)\zeta_2(s)\, dW(s) \ P\text{-a.s.}.
\end{aligned}$$

The appearing Bochner integrals have P-a.s. continuous trajectories by Lemma 3.9.
To prove that there is a continuous version of the stochastic integrals one proceeds analogously to the case of the first derivative (see Proposition 4.8). □

Chapter 6

Application: Feller Property of the Transition Semigroup and Kolmogorov Equation

We assume that Hypothesis H.0 is fulfilled. Let $X(x)(t)$, $t \in [0,T]$, be the mild solution of problem (2.1)

$$\begin{cases} dX(t) & = [AX(t) + F(X(t))]\, dt + B(X(t))\, dW(t) \\ X(0) & = x \in H \end{cases}$$

For any

$$\varphi \in \mathcal{B}_b(H) := \{f : H \to \mathbb{R} \mid f \text{ bounded and Borel measurable}\}$$

we define

$$p_t\varphi(x) := E(\varphi(X(x)(t)), \quad t \in [0,T],\ x \in H.$$

Then p_t, $t \in [0,T]$, has the semigroup property, i.e. $p_t(p_s\varphi)(x) = p_{t+s}\varphi(x)$ for all $x \in H$ and $s,t \in [0,T]$ with $s+t \in [0,T]$. (This can be proved in the same way as [DaPrZa 92, Theorem 9.8, p.249; Corollary 9.9, p.251; Corollary 9.10, p.251].) Hence p_t, $t \in [0,T]$, is called the transition semigroup corresponding to the mild solution X.
In addition, we set for $k \in \{0,1,2\}$

$$C_b^k(H) := \{f : H \to \mathbb{R} \mid f \text{ is } k \text{ times Fréchet differentiable with bounded and continuous derivatives up to order } k\}$$

and

$$UC_b^k(H) := \{f \in C_b^k(H) \mid f \text{ has uniformly continuous derivatives}\}$$

6.1 Feller property of the transition semigroup

Actually, with regard to the existence of a strict solution of the Kolmogorov equation (6.1), it is important to know that $p_t : UC_b^2(H) \to UC_b^2(H)$. Since for example the Feller property of p_t, $t \in [0,T]$, is also of independent interest we will split the above assertion.

Theorem 6.1. *If we assume that we are in the setting described above we get the following results.*

(i) Under Hypothesis H.0 we have for all $t \in [0,T]$ that $p_t : C_b(H) \to C_b(H)$ and $p_t : UC_b(H) \to UC_b(H)$ respectively.

(ii) Under the Hypotheses H.0 and H.1 we have for all $t \in [0,T]$ that $p_t : C_b^1(H) \to C_b^1(H)$. If we assume in addition that Hypothesis H.1' is fulfilled we even get that $p_t : UC_b^1(H) \to UC_b^1(H)$.

(iii) Under the Hypotheses H.0, H.1 and H.2 we have for all $t \in [0,T]$ that $p_t : C_b^2(H) \to C_b^2(H)$. If we assume in addition that Hypothesis H.2' is fulfilled we even get that $p_t : UC_b^2(H) \to UC_b^2(H)$.

To prove this theorem we need the following lemma.

Lemma 6.2. *Let $p \geq 2$ and $X : H \to \mathcal{H}^p(T,H)$ once Fréchet differentiable. Moreover let E be a Banach space and we assume that $G : H \to E$ is a continuously Fréchet differentiable function with bounded derivative ($\|DG\|_{L(H,E)} \leq c$). Then we get for all $t \in [0,T]$ and $x, y \in H$ that*

$$\frac{G(X(x+hy)(t)) - G(X(x)(t))}{h} \underset{h\to 0}{\longrightarrow} DG(X(x)(t))DX(x)y(t)$$

in $L^p(\Omega, \mathcal{F}_t, P; E)$.

Proof. We fix $t \in [0,T]$ and $x, y \in H$. Then we get by the fundamental Theorem A.7 that

$$\begin{aligned}
&\Big(E\big(\|\frac{G(X(x+hy)(t)) - G(X(x)(t))}{h} - DG(X(x)(t))DX(x)y(t)\|_E^p\big)\Big)^{\frac{1}{p}} \\
&= \Big(E\big(\| \int_0^1 DG\big(X(x)(t) + \sigma\,(X(x+hy)(t) - X(x)(t))\big) \\
&\qquad\qquad\qquad (\frac{X(x+hy)(t) - X(x)(t)}{h}) \\
&\qquad -DG(X(x)(t))DX(x)y(t)\; d\sigma\; \|_E^p\big)\Big)^{\frac{1}{p}}
\end{aligned}$$

$$\leq \Big(E\big(\int_0^1 \|DG\big(X(x)(t)+\sigma\,(X(x+hy)(t)-X(x)(t))\big)\Big(\frac{X(x+hy)(t)-X(x)(t)}{h}\Big) - DG(X(x)(t))DX(x)y(t)\|_E^p\, d\sigma\,\big)\Big)^{\frac{1}{p}}$$

$$\leq \Big(E\big(\int_0^1 \|DG\big(X(x)(t)+\sigma(X(x+hy)(t)-X(x)(t))\big)\|_{L(H,E)}^p \|\frac{X(x+hy)(t)-X(x)(t)}{h} - DX(x)y(t)\|^p\, d\sigma\big)\Big)^{\frac{1}{p}}$$

$$+ \Big(E\big(\int_0^1 \|DG\big(X(x)(t)+\sigma(X(x+hy)(t)-X(x)(t))\big) - DG(X(x)(t))\|_{L(H,E)}^p \|DX(x)y(t)\|^p\, d\sigma\big)\Big)^{\frac{1}{p}}$$

$$\leq c\Big(E\big(\|\frac{X(x+hy)(t)-X(x)(t)}{h} - DX(x)y(t)\|^p\big)\Big)^{\frac{1}{p}}$$

$$+ \Big(\int_0^1 E\big(\|\,DG(X(x)(t)+\sigma(X(x+hy)(t)-X(x)(t))) - DG(X(x)(t))\|_{L(H,E)}^{2p}\big)\, d\sigma\Big)^{\frac{1}{2p}} \Big(E\big(\|DX(x)y(t)\|^{2p}\big)\Big)^{\frac{1}{2p}}$$

As $X : H \to \mathcal{H}^p(T,H)$ is Fréchet differentiable it is easy to see that the first summand converges to zero as h converges to zero.
Because of the continuity of $X(\cdot)(t) : H \to L^p(\Omega, \mathcal{F}_t, P; H)$ we obtain that

$$X(x)(t) + \cdot(X(x+hy)(t) - X(x)(t)) \underset{h\to 0}{\longrightarrow} X(x)(t)$$

in P_1-probability. Therefore we get by Lemma 4.6 that

$$\|DG\big(X(x)(t) + \cdot(X(x+hy)(t) - X(x)(t))\big) - DG(X(x)(t))\|_{L(H,E)} \underset{h\to 0}{\longrightarrow} 0$$

in P_1-probability. In addition, we have by assumption that the norm of the difference is dominated by $2c$ and therefore the second summand also converges to zero as $h \to 0$. □

Proof of Theorem 6.1. Proof of (i): Let $\varphi \in C_b(H)$. Since the mapping $x \mapsto X(x)(t)$ is especially stochastically continuous we get by the help of Lemma 4.6 that $p_t\varphi : H \to \mathbb{R}$ is continuous. The boundedness is obvious.
Now let $\varphi \in UC_b(H)$ with $|\varphi(x)| \leq c$ for all $x \in H$ and let $\varepsilon > 0$. Then there exists $\delta > 0$ such that $|\varphi(z) - \varphi(\tilde{z})| < \varepsilon$ for all $z, \tilde{z} \in H$ with $\|z - \tilde{z}\| < \delta$. Thus we obtain by Theorem 3.2 for $x, \tilde{x} \in H$ that

$$\begin{aligned}
&E(|\varphi(X(x)(t)) - \varphi(X(\tilde{x})(t))|) \\
&\leq \int_{\{\|X(x)(t)-X(\tilde{x})(t))\|<\delta\}} |\varphi(X(x)(t)) - \varphi(X(\tilde{x})(t))| \, dP \\
&\quad + \int_{\{\|X(x)(t)-X(\tilde{x})(t))\|\geq\delta\}} |\varphi(X(x)(t)) - \varphi(X(\tilde{x})(t))| \, dP \\
&\leq \varepsilon + 2c \, P(\|X(x)(t) - X(\tilde{x})(t)\| \geq \delta) \\
&\leq \varepsilon + \frac{2c}{\delta^2} \|X(x) - X(\tilde{x})\|^2_{\mathcal{H}^2} \\
&\leq \varepsilon + \frac{2c}{\delta^2} L^2_{T,2} \|x - \tilde{x}\|^2
\end{aligned}$$

Therefore the assertion follows.

Proof of (ii): Let $\varphi \in C^1_b(H)$ with $|\varphi(x)| + \|D\varphi(x)\|_{L(H,\mathbb{R})} \leq c$ for all $x \in H$.

Claim 1: The mapping $x \mapsto p_t\varphi(x)$ is once Gâteaux differentiable with $\partial(p_t\varphi)(x)y = E(D\varphi(X(x)(t))DX(x)y(t))$.

This result is easy to see because

$$\begin{aligned}
&|\frac{p_t\varphi(x + hy) - p_t\varphi(x)}{h} - E(D\varphi(X(x)(t))DX(x)y(t))| \\
\leq\ & E(|\frac{\varphi(X(x + hy)(t) - \varphi(X(x)(t))}{h} - D\varphi(X(x)(t))DX(x)y(t))|) \longrightarrow 0
\end{aligned}$$

as $h \to 0$ by Lemma 6.2 and it is obvious that $\partial(p_t\varphi)(x) \in L(H, \mathbb{R})$.

Claim 2: Concerning the boundedness we even have that

$$\sup_{(t,x)\in[0,T]\times H} \|\partial(p_t\varphi)(x)\|_{L(H,\mathbb{R})} < \infty.$$

For that let $x, y \in H$. Then we get that

$$\begin{aligned}
|E(D\varphi(X(x)(t))DX(x)y(t))| &\leq E(|D\varphi(X(x)(t))DX(x)y(t)|) \\
&\leq c\|DX(x)y\|_{\mathcal{H}^2(T,H)} \\
&\leq c \, L_{T,2}\|y\|
\end{aligned}$$

as $X : H \to \mathcal{H}^2(T, H)$ is Lipschitz continuous and the claim is proved.

Claim 3: The mapping $x \mapsto \partial(p_t\varphi)(x)$ is continuous from H to $L(H, \mathbb{R})$.

To this end let $x_0, x, y \in H$. Then we get that

$$\begin{aligned}
&|\partial(p_t\varphi)(x)y - \partial(p_t\varphi)(x_0)y| \\
&= |E(D\varphi(X(x)(t))DX(x)y(t) - D\varphi(X(x_0)(t))DX(x_0)y(t))|
\end{aligned}$$

$$\begin{aligned}
&\leq E(|D\varphi(X(x)(t))\big[DX(x)y(t) - DX(x_0)y(t)\big]| \\
&\quad + |\big[D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\big]DX(x_0)y(t)|) \\
&\leq c\,E(\|DX(x)y(t) - DX(x_0)y(t)\|) \\
&\quad + E(\|D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\|_{L(H,\mathbb{R})}\|DX(x_0)y(t)\|) \\
&\leq c\,\|DX(x)y - DX(x_0)y\|_{\mathcal{H}^2(T,H)} \\
&\quad + \big(E(\|D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}\|DX(x_0)y\|_{\mathcal{H}^2(T,H)} \\
&\leq c\,\|DX(x) - DX(x_0)\|_{L(H,\mathcal{H}^2(T,H))}\|y\| \\
&\quad + \big(E(\|D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}\|DX(x_0)\|_{L(H,\mathcal{H}^2(T,H))}\|y\|
\end{aligned}$$

Under Hypotheses H.0 and H.1 we know by Theorem 4.3 (iv) that $DX : H \to L(H, \mathcal{H}^2(T,H))$ is continuous. Moreover we get by the help of Lemma 4.6 that

$$E(\|D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\|^2_{L(H,\mathbb{R})}) \longrightarrow 0 \quad \text{as } x \to x_0$$

since the mapping $x \mapsto X(x)(t)$ is especially stochastically continuous by Theorem 3.2 and $D\varphi : H \to L(H,\mathbb{R})$ is continuous and bounded.

Actually Claim 3 implies by Proposition D.6 that the mapping $x \mapsto p_t\varphi(x)$ is even (continuously) Fréchet differentiable with $D(p_t\varphi)(x) = \partial(p_t\varphi)(x)$.

Let now $\varphi \in UC_b^1(H)$.
Claim 4: If we assume that the additional Hypothesis H.1' is fulfilled then the mapping $x \mapsto D(p_t\varphi)(x)$ is uniformly continuous from H to $L(H,\mathbb{R})$.
Let $\varepsilon > 0$ and $x, \tilde{x}, y \in H$. Then we get that

$$\begin{aligned}
&|D(p_t\varphi)(x)y - D(p_t\varphi)(\tilde{x})y| \\
&\leq c\,\|DX(x) - DX(\tilde{x})\|_{L(H,\mathcal{H}^2(T,H))}\|y\| \\
&\quad + \big(E(\|D\varphi(X(x)(t)) - D\varphi(X(\tilde{x})(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}\|DX(\tilde{x})\|_{L(H,\mathcal{H}^2(T,H))}\|y\|
\end{aligned}$$

By Theorem 4.3 (iii) we have that $DX : H \to L(H, \mathcal{H}^2(T,H))$ is bounded and by Corollary 4.7 that there exists a $\delta > 0$ such that

$$\|DX(x) - DX(\tilde{x})\|_{L(H,\mathcal{H}^2(T,H))} \leq \varepsilon \quad \text{if } \|x - \tilde{x}\| < \delta$$

Besides, as in the second part of the proof of (i), we get the existence of $\tilde{\delta} > 0$ such that

$$\big(E(\|D\varphi(X(x)(t)) - D\varphi(X(\tilde{x})(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}} \leq \varepsilon$$

for all $x, \tilde{x} \in H$ with $\|x - \tilde{x}\| < \tilde{\delta}$. In this way (ii) is proved.

Proof of (iii): Let $\varphi \in C_b^2(H)$ with
$|\varphi(x)| + \|D\varphi(x)\|_{L(H,\mathbb{R})} + \|D^2\varphi(x)\|_{L(H,L(H,\mathbb{R}))} \leq c$ for all $x \in H$.
Claim 1: The mapping $x \mapsto D(p_t\varphi)(x)$ from H to $L(H,\mathbb{R})$ is Gâteaux differentiable where

$$\begin{aligned}\partial D(p_t\varphi)(x)(y)z = \; & E\big(D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\ & + D\varphi(X(x)(t))D^2X(x)(y)z(t)\big)\end{aligned}$$

for all $x, y, z \in H$.
If we take x, y, z from H we get that

$$\begin{aligned}
&|\,E\Big(\frac{D\varphi(X(x+hz)(t))DX(x+hz)y(t) - D\varphi(X(x)(t))DX(x)y(t)}{h}\\
&\quad - D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\
&\quad - D\varphi(X(x)(t))D^2X(x)(y)z(t)\Big)|\\
\leq\; & E\Big(|\frac{D\varphi(X(x+hz)(t))DX(x+hz)y(t) - D\varphi(X(x+hz)(t))DX(x)y(t)}{h}\\
&\quad - D\varphi(X(x)(t))D^2X(x)(y)z(t)|\Big)\\
&+ E\Big(|\frac{D\varphi(X(x+hz)(t))DX(x)y(t) - D\varphi(X(x)(t))DX(x)y(t)}{h}\\
&\quad - D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)|\Big)
\end{aligned}$$

1. The first summand can be estimated in the following way.

$$\begin{aligned}
&E\Big(|D\varphi(X(x+hz)(t))\frac{DX(x+hz)y(t) - DX(x)y(t)}{h}\\
&\quad - D\varphi(X(x)(t))D^2X(x)(y)z(t)|\Big)\\
\leq\; & E\Big(|D\varphi(X(x+hz)(t))\big(\frac{DX(x+hz)y(t) - DX(x)y(t)}{h}\\
&\qquad - D^2X(x)(y)z(t)\big)|\Big)\\
&+ E\big(|\big[D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))\big]D^2X(x)(y)z(t)|\big)\\
\leq\; & c\,\|\frac{DX(x+hz)y - DX(x)y}{h} - D^2X(x)(y)z\|_{\mathcal{H}^2(T,H)}\\
&+ \big(E(\|D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}\\
&\qquad \|D^2X(x)(y)z\|_{\mathcal{H}^2(T,H)}\\
\leq\; & c\,\|\frac{DX(x+hz) - DX(x)}{h} - D^2X(x)(\cdot)z\|_{L(H,\mathcal{H}^2(T,H))}\|y\|
\end{aligned}$$

$$+ \ \big(E(\|D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}$$
$$\|D^2X(x)(\cdot)z\|_{L(H,\mathcal{H}^2(T,H))}\|y\|$$
$$= c_h \ \|y\|$$

where $c_h \longrightarrow 0$ as $h \to 0$ because $X : H \to \mathcal{H}^2(T,H)$ is twice Fréchet differentiable by Theorem 5.3 (iv) and

$$E(\|D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))\|^2_{L(H,\mathbb{R})}) \longrightarrow 0 \quad \text{as } h \to 0$$

as we have already seen in the proof of (ii), Claim 3.
2. Concerning the second summand we obtain that

$$E\Big(|\frac{D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))}{h}DX(x)y(t)$$
$$- D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)|\Big)$$
$$\leq \Big(\ E\Big(\|\frac{D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))}{h}$$
$$- D^2\varphi(X(x)(t))(\cdot)DX(x)z(t)\|^2_{L(H,\mathbb{R})}\Big)\Big)^{\frac{1}{2}}$$
$$\big(E(\|DX(x)y(t)\|^2)\big)^{\frac{1}{2}}$$
$$\leq \Big(\ E\Big(\|\frac{D\varphi(X(x+hz)(t)) - D\varphi(X(x)(t))}{h}$$
$$- D^2\varphi(X(x)(t))(\cdot)DX(x)z(t)\|^2_{L(H,\mathbb{R})}\Big)\Big)^{\frac{1}{2}}$$
$$\|DX(x)\|_{L(H,\mathcal{H}^2(T,H))}\|y\|$$
$$= \tilde{c}_h \ \|y\|$$

where $\tilde{c}_h \longrightarrow 0$ as $h \to 0$ because of Lemma 6.2. Moreover it is obvious that $\partial D(p_t\varphi)(x) \in L(H, L(H,\mathbb{R}))$ for all $x \in H$.
Claim 2: Concerning the boundedness we even obtain that

$$\sup_{(t,x)\in[0,T]\times H} \|\partial D(p_t\varphi)(x)\|_{L(H,L(H,\mathbb{R}))} < \infty.$$

For that let $x, y, z \in H$. We obtain by Theorem 4.3 (iii) and Theorem 5.3 (iii) that

$$|E\big(D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)$$
$$+ D\varphi(X(x)(t))D^2X(x)(y)z(t)\big)|$$

$$\begin{aligned}
&\leq E\big(|D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)|\big)\\
&\quad + E\big(|D\varphi(X(x)(t))(D^2X(x)(y)z(t))|\big)\\
&\leq (\tilde{c}\, L_{T,2}^2 + c\, \tilde{C}_{T,2})\|y\|\|z\|
\end{aligned}$$

where $\tilde{C}_{T,2} := \inf_{q>4} \tilde{C}_{T,2,q}$ (see Theorem 5.3 (iii) and Theorem 4.3(iii)).

Claim 3: The mapping $x \mapsto \partial D(p_t\varphi)(x)$ from H to $L(H, L(H,\mathbb{R}))$ is continuous.

Let $x, x_0, y, z \in H$. Then we obtain that

$$\begin{aligned}
&|\,E\big(D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\
&\qquad + D\varphi(X(x)(t))D^2X(x)(y)z(t)\big)\\
&\quad - E\big(D^2\varphi(X(x_0)(t))(DX(x_0)y(t))DX(x_0)z(t)\\
&\qquad\quad + D\varphi(X(x_0)(t))D^2X(x_0)(y)z(t)\big)|\\
&\leq E\big(|D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\
&\qquad - D^2\varphi(X(x_0)(t))(DX(x_0)y(t))DX(x_0)z(t)|\big)\\
&\quad + E\big(|D\varphi(X(x)(t))D^2X(x)(y)z(t) - D\varphi(X(x_0)(t))D^2X(x_0)(y)z(t)|\big)
\end{aligned}$$

1. The first summand can be estimated in the following way:

$$\begin{aligned}
&E\big(|D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\
&\qquad - D^2\varphi(X(x_0)(t))(DX(x_0)y(t))DX(x_0)z(t)|\big)\\
&\leq E\big(|D^2\varphi(X(x)(t))(DX(x)y(t))\big(DX(x)z(t) - DX(x_0)z(t)\big)|\big)\\
&\quad + E\big(|D^2\varphi(X(x)(t))\big(DX(x)y(t) - DX(x_0)y(t)\big)DX(x_0)z(t)|\big)\\
&\quad + E\big(|\big(D^2\varphi(X(x)(t)) - D^2\varphi(X(x_0)(t))\big)(D(X(x_0)y(t))DX(x_0)z(t)|\big)\\
&\leq \Big(2c\, L_{T,2}\, \|DX(x) - DX(x_0)\|_{L(H,\mathcal{H}^2(T,H))}\\
&\quad + L_{T,4}^2\big(E(\|D^2\varphi(X(x)(t)) - D^2\varphi(X(x_0)(t))\|^2_{L(H,L(H,\mathbb{R}))})\big)^{\frac{1}{2}}\Big)\|y\|\|z\|
\end{aligned}$$

where we know by the same arguments as in the proof of (ii), Claim 3, that for all $\varepsilon > 0$ there exists $\delta = \delta(x_0) > 0$ such that

$$\begin{aligned}
&2c\, L_{T,2}\, \|DX(x) - DX(x_0)\|_{L(H,\mathcal{H}^2(T,H))}\\
&\quad + L_{T,4}^2\big(E(\|D^2\varphi(X(x)(t)) - D^2\varphi(X(x_0)(t))\|^2_{L(H,L(H,\mathbb{R}))})\big)^{\frac{1}{2}} \leq \varepsilon
\end{aligned}$$

for all $x \in H$ with $\|x - x_0\| \leq \delta$.

2. The second summand can be estimated in a similar way. We get that for all $\varepsilon > 0$ there exists $\delta = \delta(x_0) > 0$ such that

$$E(|D\varphi(X(x)(t))D^2X(x)(y)z(t) - D\varphi(X(x_0)(t))D^2X(x_0)(y)z(t)|)$$

$$
\begin{aligned}
&\le E\Big(|D\varphi(X(x)(t))\big(D^2X(x)(y)z(t) - D^2X(x_0)(y)z(t)\big)|\Big)\\
&\quad + E\Big(|\big(D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\big)D^2X(x_0)(y)z(t)|\Big)\\
&\le c\,E\Big(\|D^2X(x)(y)z(t) - D^2X(x_0)(y)z(t)\|\Big)\\
&\quad + \big(E(\|D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}\big(E(\|D^2X(x_0)(y)z(t)\|^2)\big)^{\frac{1}{2}}\\
&\le \Big(c\,\|D^2X(x) - D^2X(x_0)\|_{L(H,L(H,\mathcal{H}^2(T,H)))}\\
&\quad + C_{T,2}\big(E(\|D\varphi(X(x)(t)) - D\varphi(X(x_0)(t))\|^2_{L(H,\mathbb{R})})\big)^{\frac{1}{2}}\Big)\|y\|\|z\|\\
&\le \varepsilon\|y\|\|z\|
\end{aligned}
$$

for all $x \in H$ with $\|x - x_0\| \le \delta$.
In particular, we know by Proposition D.6 that $\partial D(p_t\varphi) = D^2(p_t\varphi)$.

Let now $\varphi \in UC_b^2(H)$.
Claim 4: If we assume that the additional Hypothesis H.2' is fulfilled we get that the mapping $D^2(p_t\varphi) : H \to L(H, L(H,\mathbb{R}))$ is uniformly continuous.
This is clear as we know by Corollary 4.7 and Corollary 5.6 that the derivatives DX and D^2X of $X : H \to \mathcal{H}^2(T,H)$ are uniformly continuous. Using the same estimates as in Claim 3 the assertion follows by inspection. □

6.2 Kolmogorov Equation

Now we have all instruments to study the following Kolmogorov equation associated with the problem (2.1) (see [DaPrZa 96, p.71])

$$
(6.1)\qquad \begin{cases} \frac{\partial u}{\partial t}(t,x) &= \frac{1}{2}\,\mathrm{tr}\,\big[D^2u(t,x)B(x)\big(B(x)\big)^*\big]\\ &\quad +\langle Ax + F(x), Du(t,x)\rangle, \quad t \in [0,T],\ x \in D(A)\\ u(0,x) &= \varphi(x), \quad x \in H \end{cases}
$$

Definition 6.3 (Strict solution). A continuous function
$u : [0,T] \times H \to \mathbb{R}$ is called strict solution of (6.1) on $[0,T]$ if

(i) $u(t,\cdot) \in C_b^2(H)$ for all $t \in [0,T]$

(ii) $u(\cdot,x) \in C^1([0,T])$ for all $x \in D(A)$ and

(iii) the Kolmogorov equation (6.1) is fulfilled.

To prove the existence and uniqueness of solutions for equation (6.1) we have to introduce a new condition on B (see [DaPrZa 96, Hypothesis 5.2 (ii), p.71]).

Hypothesis H.3

- $B : H \to L_2(U, H)$ and there exists a constant $K > 0$ such that

$$\|B(x) - B(y)\|_{L_2} \le K \, \|x - y\|, \quad \text{for all } x, y \in H.$$

Theorem 6.4. *(i) Let $p > 2$. We assume that the mapping $F : H \to H$ is Lipschitz continuous with Lipschitz constant $C > 0$, $B : H \to L_2(U, H)$ fulfills Hypothesis H.3 and $A \in L(H)$.*
In addition, we require that there is a strong solution $X(x)(t)$, $t \in [0, T]$, of problem (2.1) for each initial condition $x \in H$ such that $X : H \to \mathcal{H}^p(T, H)$ is twice continuously Fréchet differentiable.
If there are continuous versions of $X(x)(t)$, $DX(x)y(t)$ and $D^2X(x)(y)z(t)$, $t \in [0, T]$, for all $x, y, z \in H$ and if we know that $p_t\varphi = E(\varphi(X(\cdot)(t)) \in UC_b^2(H)$ for all $\varphi \in UC_b^2(H)$ then the mapping $u : [0, T] \times H \to \mathbb{R}$ given by $u(t, x) = p_t\varphi(x)$ is a strict solution of problem (6.1) with initial condition $\varphi \in UC_b^2(H)$.

(ii) We assume that $X(x)(t)$, $t \in [0, T]$, is a strong solution of problem (2.1) with initial condition $x \in H$. If $v : [0, T] \times H \to \mathbb{R}$ is a strict solution of problem (6.1) with initial condition $\varphi \in UC_b^2(H)$ such that v and the derivatives $\frac{\partial v}{\partial t}, Dv$ and D^2v are uniformly continuous on bounded subsets of $[0, T] \times H$ and bounded on $[0, T] \times H$ then we have that

$$v(t, x) = p_t\varphi(x) \quad \text{for all } (t, x) \in [0, T] \times H.$$

Remark 6.5. (i) Let $A \in L(H)$, assume that the Hypotheses H.0, H.1, H.2, H.2' and H.3 are fulfilled and require in addition that the condition of Proposition 5.7 is satisfied.
Then there is a mild solution X of problem (2.1) (see Theorem 3.2) such that the mapping $X : H \to \mathcal{H}^p(T, H)$, $p \ge 2$, is twice continuously Fréchet differentiable (see Theorem 5.3). Proposition 3.15, Proposition 4.8 and Proposition 5.7 provides that there exist pathwise continuous versions of the derivatives. Moreover we know by Theorem 6.1 that $p_t\varphi \in UC_b^2(H)$ for all $\varphi \in UC_b^2(H)$. By Proposition 2.10 and Proposition 2.9 we know that it is even a strong solution. Hence all conditions of Theorem 6.4 are satisfied.
These assumptions are corresponding to those made by Da Prato and Zabczyk in [DaPrZa 96, Theorem 5.4.2, p.71] where we added the assumptions providing the pathwise continuity of the second order derivative of the mild solution.

(ii) In contrast to [DaPrZa 96, Theorem 5.4.2, p.71] we do not get the uniqueness of the strict solution $p.\varphi : [0,T] \times H \to \mathbb{R}$ in Theorem 6.4 (i). This uniqueness can only be proved under stronger assumptions on the solution (see Theorem 6.4 (ii)). In general we do not know that the mapping $p.\varphi$ fulfills these conditions. For example the mapping itself is not uniformly continuous on bounded subsets of $[0,T] \times H$ in general.
(iii) The Yosida approximation provides a possibility to generalize this result to the case that A is not necessarily in $L(H)$ (see [DaPrZa 92, Theorem 9.17, p.261]).

To prove Theorem 6.4 we need the following preparations.

Lemma 6.6. *Under the assumptions of Theorem 6.4 (i) we get that the mappings*

$$\text{(i)} \qquad t \mapsto E\big(\langle AX(x)(t) + F(X(x)(t)), D\varphi(X(x)(t))\rangle\big) + E\big(\mathrm{tr}\big[D^2\varphi(X(x)(t))B(X(x)(t))\big(B(X(x)(t))\big)^*\big]\big)$$

and

$$\text{(ii)} \qquad t \mapsto \mathrm{tr}\big[D^2(p_t\varphi)(x)B(x)\big(B(x)\big)^*\big] + \langle Ax + F(x), D(p_t\varphi)(x)\rangle$$

are continuous from $[0,T]$ to $\mathbb{R}$ for all $x \in H$.

Proof. (i): By assumption we know that the mapping $t \mapsto X(x)(t)$ is P-a.s. continuous. Therefore we get that

$$t \mapsto \langle AX(x)(t) + F(X(x)(t)), D\varphi(X(x)(t))\rangle$$

is P-a.s. continuous because of the continuity of A, F and $D\varphi$. In addition, the family $\langle AX(x)(t) + F(X(x)(t)), D\varphi(X(x)(t))\rangle$, $t \in [0,T]$, is uniformly integrable since it is dominated by $\tilde{c}\|X(x)(t)\|$, $t \in [0,T]$, and $X(x) \in \mathcal{H}^p(T,H)$ for $p > 1$. Therefore the continuity of the first summand is clear. To prove the continuity of the second summand we fix $t_0 \in [0,T]$ and a sequence t_n, $n \in \mathbb{N}$, in $[0,T]$ such that $t_n \longrightarrow t_0$ as $n \to \infty$. Then we get by Lemma B.8 that

$$\begin{aligned}
&\Big|\mathrm{tr}\Big[D^2\varphi(X(x)(t_n))B(X(x)(t_n))\big(B(X(x)(t_n))\big)^*\Big] \\
&\quad - \mathrm{tr}\Big[D^2\varphi(X(x)(t_0))B(X(x)(t_0))\big(B(X(x)(t_0))\big)^*\Big]\Big| \\
&\le \Big|\mathrm{tr}\Big[D^2\varphi(X(x)(t_n))B(X(x)(t_n))\Big(\big(B(X(x)(t_n))\big)^* - \big(B(X(x)(t_0))\big)^*\Big)\Big]\Big| \\
&\quad + \Big|\mathrm{tr}\Big[D^2\varphi(X(x)(t_n))\big(B(X(x)(t_n)) - B(X(x)(t_0))\big)\big(B(X(x)(t_0))\big)^*\Big]\Big| \\
&\quad + \Big|\mathrm{tr}\Big[\big(D^2\varphi(X(x)(t_n)) - D^2\varphi(X(x)(t_0))\big)B(X(x)(t_0))\big(B(X(x)(t_0))\big)^*\Big]\Big|
\end{aligned}$$

$$\leq c\,K^2\,(1+\|X(x)(t_n)\|)\|X(x)(t_n)-X(x)(t_0)\|$$
$$+\,c\,K^2\,(1+\|X(x)(t_0)\|)\|X(x)(t_n)-X(x)(t_0)\|$$
$$+\,\|D^2\varphi(X(x)(t_n))-D^2\varphi(X(x)(t_0))\|_{L(H)}\,K^2(1+\|X(x)(t_0)\|)^2\longrightarrow 0$$

P-a.s. as $n\to\infty$. In addition, the family is uniformly integrable as $X(x)\in\mathcal{H}^p(T,H)$, $p>2$, and therefore the continuity of the second summand is also proved.
(ii): To verify the continuity of the first summand we first have to notice that by Proposition B.10

$$\operatorname{tr}\big[D^2(p_t\varphi)(x)B(x)\big(B(x)\big)^*\big]=\operatorname{tr}\big[\big(B(x)\big)^*D^2(p_t\varphi)(x)B(x)\big]$$

Moreover we have that

$$\begin{aligned}t\mapsto & D^2(p_t\varphi)(x)(y)z\Big(=\langle\underbrace{D^2(p_t\varphi)(x)z}_{\in H\cong L(H,\mathbb{R})},y\rangle\Big)\\ =&E(\,D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\ &\quad+D\varphi(X(x)(t))D^2X(x)(y)z(t))\end{aligned}$$

is continuous from $[0,T]$ to $\mathbb{R}$ for all $x,y,z\in H$ as we first know that

$$\begin{aligned}t\mapsto & D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\ &+D\varphi(X(x)(t))D^2X(x)(y)z(t)\end{aligned}$$

is P-a.s. continuous because of the P-a.s. continuity of $t\mapsto X(x)(t)$, $t\mapsto DX(x)y(t)$ and $t\mapsto D^2X(x)(y)z(t)$. Secondly, we have that the family

$$\begin{aligned}&D^2\varphi(X(x)(t))(DX(x)y(t))DX(x)z(t)\\ &+D\varphi(X(x)(t))D^2X(x)(y)z(t),\quad t\in[0,T],\end{aligned}$$

is uniformly integrable.
Hence let e_k, $k\in\mathbb{N}$, be an orthonormal basis of U. Then we get that

$$t\mapsto\langle(B(x))^*D^2(p_t\varphi)(x)B(x)e_k,e_k\rangle=D^2(p_t\varphi)(x)\big(B(x)e_k\big)B(x)e_k$$

is continuous from $[0,T]$ to H for all $k\in\mathbb{N}$ and

$$\begin{aligned}&|\langle(B(x))^*D^2(p_t\varphi)(x)B(x)e_k,e_k\rangle|\\ \leq\;&\tilde{c}\,\|B(x)e_k\|^2\end{aligned}$$

as $\sup\limits_{t\in[0,T]}\|D^2(p_t\varphi)(x)\|_{L(H)}\leq c(\|DX(x)\|^2_{L(H,\mathcal{H}^2)}+\|D^2X(x)\|_{L(H,L(H,\mathcal{H}^2))})$ $<\infty$. But $\sum\limits_{k\in\mathbb{N}}\|B(x)e_k\|^2=\|B(x)\|^2_{L_2}<\infty$ and therefore the first summand

is continuous by Lemma 3.10.
To prove that the second summand is also continuous we simply show that

$$t \mapsto \langle \underbrace{D(p_t\varphi)(x)}_{\in H \cong L(H,\mathbb{R})}, y\rangle = D(p_t\varphi)(x)y = E\big(D\varphi(X(x)(t))DX(x)y(t)\big)$$

is continuous from $[0, T]$ to $\mathbb{R}$ for all $x, y \in H$. This is true as

$$t \mapsto D\varphi(X(x)(t))DX(x)y(t)$$

is P-a.s. continuous because of the P-a.s. continuity of $t \mapsto X(x)(t)$ and $t \mapsto DX(x)y(t)$. The family $D\varphi(X(x)(t))DX(x)y(t)$, $t \in [0, T]$, is in addition uniformly integrable as $DX(x)y \in \mathcal{H}^p(T, H)$ for $p > 1$ and therefore the assertion finally follows. □

Lemma 6.7. *Let $f : [0, T] \to \mathbb{R}$ be a continuous function which is differentiable on the right on $[0, T[$ such that the derivative $f'_+ : [0, T[\to \mathbb{R}$ is continuous.*
Then we get for all $0 \leq s < t \leq T$ that there exists $\xi \in [s, t[$ such that

$$\frac{f(t) - f(s)}{t - s} = f'_+(\xi)$$

Proof. Case 1: $\dfrac{f(t) - f(s)}{t - s} = 0$
If f is constant the assertion follows immediately. Otherwise there exists an element $t_0 \in]s, t[$ such that $f(t_0) \neq f(s)$ and $f(u) \leq f(t_0)$ for all $U \in [s, t]$ or $f(u) \geq f(t_0)$ for all $u \in [s, t]$. Without loss of generality, we consider the case that there exists a $t_0 \in]s, t[$ such that $f(s) < f(t_0)$ and $f(u) \leq f(t_0)$ for all $u \in [s, t]$. Then it is clear that

$$f'_+(t_0) = \lim_{u \downarrow t_0} \frac{f(u) - f(t_0)}{u - t_0} \leq 0$$

Now we set $c := \sup\{u \geq s \mid f(v) \leq f(s) \text{ for all } v \in [s, u]\}$. It is obvious that $c < t_0$. In addition, $f'_+(c) \geq 0$ since if we assume that $f'_+(c) < 0$ it follows that there is an $\varepsilon > 0$ such that

$$\frac{f(u) - f(c)}{u - c} < 0 \quad \text{for all } u \in]c, c + \varepsilon[$$

This implies that $f(u) \leq f(c) = f(s)$ for all $u \in [s, c + \varepsilon[$ which leads to a contradiction to the definition of c.
Hence, because of the continuity of f'_+ on $[c, t_0]$, we get that there is an

element $\xi \in [c, t_0]$ such that $f'_+(\xi) = 0$.

Case 2: $\dfrac{f(t) - f(s)}{t - s} = a \in \mathbb{R}$

In this case we define $g : [0, T] \to \mathbb{R}$ by $g(u) := f(u) - au$. Then it is clear that g still fulfills all assumptions of the lemma and we obtain that

$$\frac{g(t) - g(s)}{t - s} = \frac{f(t) - f(s)}{t - s} - \frac{a(t - s)}{t - s} = 0$$

Using the result of Case 1 this implies that there exists $\xi \in [s, t[$ such that $g'_+(\xi) = f'_+(\xi) - a = 0$. □

Lemma 6.8. *Let $f : [0, T] \to \mathbb{R}$ be a continuous function which is differentiable on the right on $[0, T[$. We assume that the derivative $f'_+ : [0, T[\to \mathbb{R}$ is also continuous with a continuous extension on $[0, T]$ also denoted by f'_+. Then the mapping f is even continuously differentiable with $f'_+ = f'$.*

Proof. We show that $f(t) - f(0) = \int_0^t f'_+(s)\, ds$ for all $t \in [0, T]$:

Let $\varepsilon > 0$. Because of the continuity of $f'_+ : [0, T] \to \mathbb{R}$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$

$$\left| \int_0^t f'_+(s)\, ds - \sum_{k=0}^{n-1} f'_+\left(\frac{kt}{n}\right)\frac{1}{n} \right| < \varepsilon$$

and

$$|f'_+(s) - f'_+(u)| < \varepsilon \quad \text{for all } s, u \in [0, T] \text{ with } |s - u| < \frac{1}{n}$$

Therefore we get that

$$\begin{aligned}
&|f(t) - f(0) - \int_0^t f'_+(s)\, ds| \\
&\leq |f(t) - f(0) - \sum_{k=0}^{n-1} f'_+\left(\frac{kt}{n}\right)\frac{1}{n}| + |\sum_{k=0}^{n-1} f'_+\left(\frac{kt}{n}\right)\frac{1}{n} - \int_0^t f'_+(s)\, ds| \\
&\leq |f(t) - f(0) - \sum_{k=0}^{n-1}\Big[f'_+\left(\frac{kt}{n}\right) - \frac{f\left(\frac{(k+1)t}{n}\right) - f\left(\frac{kt}{n}\right)}{\frac{1}{n}}\Big]\frac{1}{n} \\
&\qquad - \sum_{k=0}^{n-1}\Big[f\left(\frac{(k+1)t}{n}\right) - f\left(\frac{kt}{n}\right)\Big]| + \varepsilon \\
&\leq \sum_{k=0}^{n-1}\Big[\underbrace{|f'_+\left(\frac{kt}{n}\right) - \frac{f\left(\frac{(k+1)t}{n}\right) - f\left(\frac{kt}{n}\right)}{\frac{1}{n}}|}_{< \varepsilon \text{ because of Lemma 6.7}}\frac{1}{n}\Big] + \varepsilon < 2\varepsilon
\end{aligned}$$

□

Proof of Theorem 6.4. Proof of (i): We first prove the existence of a strict solution.
Since $X(x)(t)$, $t \in [0,T]$, is a strong solution it is an Itô process given by

$$X(x)(t) = x + \int_0^t AX(x)(s) + F(X(x)(s))\, ds + \int_0^t B(X(x)(s))\, dW(s)$$

P-a.s. for all $x \in H$ and $t \in [0,T]$. This allows to apply the Itô formula 1.33 to $\varphi(X(x))$, $\varphi \in UC_b^2(H)$, which provides that

$$\begin{aligned}\varphi(X(x)(t)) = \varphi(x) &+ \int_0^t \langle D\varphi(X(x)(s)), B(X(x)(s))\rangle\, dW(s)\\ &+ \int_0^t \langle D\varphi(X(x)(s)), AX(x)(s) + F(X(x)(s))\rangle\\ &\qquad + \frac{1}{2}\operatorname{tr}\big[\underbrace{D^2\varphi(X(x)(s))}_{\in L(H)}\underbrace{B(X(x)(s))}_{\in L_2(U,H)}\underbrace{\big(B(X(x)(s))\big)^*}_{L_2(H,U)}\big]\, ds\end{aligned}$$

Since $\langle D\varphi(X(x)(s)), B(X(x)(s))\rangle$, $s \in [0,T]$, is in $\mathcal{N}_W^2(0,T;\mathbb{R})$ we obtain that

$$E\Big(\int_0^t \langle D\varphi(X(x)(s)), B(X(x)(s))\rangle\, dW(s)\Big) = 0$$

On the other hand we know that

$$\begin{aligned}(s,\omega) \mapsto &\langle D\varphi(X(x)(s,\omega)), AX(x)(s,\omega) + F(X(x)(s,\omega))\rangle\\ &+ \frac{1}{2}\operatorname{tr}\big[D^2\varphi(X(x)(s,\omega))B(X(x)(s,\omega))\big(B(X(x)(s,\omega))\big)^*\big]\end{aligned}$$

is $\mathcal{P}_T$- and therefore in particular $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable. In addition, we have that

$$\begin{aligned}&E\Big(\int_0^T |\langle D\varphi(X(x)(s)), AX(x)(s) + F(X(x)(s))\rangle|\, ds\Big)\\ &\le E\Big(\int_0^T c\big[\|A\|_{L(H)}\|X(x)(s)\| + C(1 + \|X(x)(s)\|)\big]\, ds\Big) < \infty\end{aligned}$$

and by Lemma B.8 and Remark B.6

$$\begin{aligned}
&E\Big(\int_0^T \big|\text{tr}\,\big[D^2\varphi(X(x)(s))B(X(x)(s))\big(B(X(x)(s))\big)^*\big]\big|\,ds\Big)\\
&\le E\Big(\int_0^T \|D^2\varphi(X(x)(s))B(X(x)(s))\|_{L_2(U,H)}\|B(X(x)(s))\|_{L_2(U,H)}\,ds\Big)\\
&\le c\,E\Big(\int_0^T \|B(X(x)(s))\|^2_{L_2(U,H)}\,ds\Big) < \infty
\end{aligned}$$

Therefore we can use the real Fubini theorem to get that

$$\begin{aligned}
&E\Big(\int_0^t \langle D\varphi(X(x)(s)), AX(x)(s) + F(X(x)(s))\rangle\\
&\qquad + \frac{1}{2}\text{tr}\,\big[D^2\varphi(X(x)(s))B(X(x)(s))\big(B(X(x)(s))\big)^*\big]\,ds\Big)\\
&= \int_0^t E\big(\langle D\varphi(X(x)(s)), AX(x)(s) + F(X(x)(s))\rangle\\
&\qquad + \frac{1}{2}\text{tr}\,\big[D^2\varphi(X(x)(s))B(X(x)(s))\big(B(X(x)(s))\big)^*\big]\big)\,ds
\end{aligned}$$

Altogether this provides that

$$\begin{aligned}
p_t\varphi(x) =& E(\varphi(X(x)(t)))\\
=& \varphi(x) + \int_0^t E\big(\langle D\varphi(X(x)(s)), AX(x)(s) + F(X(x)(s))\rangle\\
&\qquad + \frac{1}{2}\,\text{tr}\big[D^2\varphi(X(x)(s))B(X(x)(s))\big(B(X(x)(s))\big)^*\big]\big)\,ds
\end{aligned}$$

and since

$$\begin{aligned}
s \mapsto & E\big(\langle D\varphi(X(x)(s)), AX(x)(s) + F(X(x)(s))\rangle\big)\\
&+ \frac{1}{2}E\big(\text{tr}\,\big[D^2\varphi(X(x)(s))B(X(x)(s))\big(B(X(x)(s))\big)^*\big]\big)
\end{aligned}$$

is continuous by Lemma 6.6 (i) the fundamental Theorem A.7 implies that

$$\begin{aligned}
\frac{\partial^+}{\partial t}p_t\varphi(x)_{|t=0} &:= \lim_{t\downarrow 0}\frac{p_t\varphi(x) - p_0\varphi(x)}{t}\\
&= E\big(\langle D\varphi(X(x)(0)), AX(x)(0) + F(X(x)(0))\rangle\\
&\qquad + \frac{1}{2}\text{tr}\big[D^2\varphi(X(x)(0))B(X(x)(0))\big(B(X(x)(0))\big)^*\big]\big)\\
&= \langle D\varphi(x), Ax + F(x)\rangle + \frac{1}{2}\text{tr}\big[D^2\varphi(x)B(x)\big(B(x)\big)^*\big]
\end{aligned}$$

Now we fix $t_0 > 0$ and we want to calculate the derivative at this point. That is the moment where we will use the fact that $p_{t_0}\varphi = \tilde{\varphi} \in UC_b^2(H)$ if $\varphi \in UC_b^2(H)$. Together with the semigroup property of p_t, $t \in [0,T]$, this provides that

$$\begin{aligned}\frac{\partial^+}{\partial t} p_t\varphi(x)_{|t=t_0} &= \lim_{t\downarrow 0} \frac{p_{t_0+t}\varphi(x) - p_{t_0}\varphi(x)}{t} \\ &= \lim_{t\downarrow 0} \frac{p_t\tilde{\varphi}(x) - p_0\tilde{\varphi}(x)}{t} \\ &= \langle D\tilde{\varphi}(x), Ax + F(x)\rangle + \frac{1}{2}\mathrm{tr}\big[D^2\tilde{\varphi}(x)B(x)\big(B(x)\big)^*\big] \\ &= \langle D(p_{t_0}\varphi)(x), Ax + F(x)\rangle + \frac{1}{2}\mathrm{tr}\big[D^2(p_{t_0}\varphi)(x)B(x)\big(B(x)\big)^*\big]\end{aligned}$$

As the right hand side is a continuous function in t by Lemma 6.6 (ii) we get by Lemma 6.8 that $p.\varphi(x) \in C^1([0,T])$ with

$$\frac{\partial}{\partial t}(p_t\varphi)(x)_{|t=t_0} = \langle D(p_{t_0}\varphi)(x), Ax + F(x)\rangle + \frac{1}{2}\,\mathrm{tr}\big[D^2(p_{t_0}\varphi)(x)B(x)\big(B(x)\big)^*\big]$$

As it is obvious that $p_0\varphi(x) = E(\varphi(X(x)(0))) = \varphi(x)$ and since we know by Theorem 6.1 that $p_t\varphi \in UC_b^2(H) \subset C_b^2(H)$ it only remains to prove that

$$\begin{aligned} p.\varphi : [0,T] \times H &\to \mathbb{R} \\ (t,x) &\mapsto p_t\varphi(x)\end{aligned}$$

is continuous to get that $p.\varphi$ is a strict solution of the Kolmogorov equation (6.1).
For that let $(t,x), (t_n, x_n) \in [0,T] \times H$ such that $t_n \longrightarrow t$ and $x_n \longrightarrow x$ as $n \to \infty$. Then we obtain that

$$\begin{aligned} &|p_t\varphi(x) - p_{t_n}\varphi(x_n)| \\ &\le E(|\varphi(X(x)(t)) - \varphi(X(x)(t_n))|) + E(|\varphi(X(x)(t_n)) - \varphi(X(x_n)(t_n))|) \\ &\le E(|\varphi(X(x)(t)) - \varphi(X(x)(t_n))|) + c\|X(x) - X(x_n)\|_{\mathcal{H}^2} \underset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

Proof of (ii): To prove that $v(s,x) = p_s\varphi(x)$ for all $(s,x) \in [0,T] \times H$ we apply the Itô formula to the process

$$v(s-u, X(x)(u)), \quad u \in [0,s],$$

and get that

$$
\begin{aligned}
&v(s-u, X(x)(u)) \\
&= v(s, X(x)(0)) + \int_0^u \langle Dv(s-r, X(x)(r)), B(X(x)(r))\rangle \, dW(r) \\
&\quad + \int_0^u -\frac{\partial v}{\partial t}(s-r, X(x)(r)) \, dr \\
&\quad + \int_0^u \langle Dv(s-r, X(x)(r)), AX(x)(r) + F(X(x)(r))\rangle \, dr \\
&\quad + \int_0^u \frac{1}{2} \operatorname{tr}\left[D^2 v(s-r, X(x)(r)) B(X(x)(r)) \big(B(X(x)(r))\big)^*\right] dr \\
&= v(s,x) + \int_0^u \langle Dv(s-r, X(x)(r)), B(X(x)(r))\rangle \, dW(r)
\end{aligned}
$$

as v is a strict solution of problem (6.1). If we choose now $u = s$ and take then the expectation we obtain that

$$
\begin{aligned}
&p_s\varphi(x) = E(\varphi(X(x)(s))) = E(v(0, X(x)(s))) \\
&= v(s,x) + E\Big(\int_0^s \langle Dv(s-r, X(x)(r)), B(X(x)(r))\rangle \, dW(r)\Big) = v(s,x)
\end{aligned}
$$

as the process $\langle Dv(s-r, X(x)(r)), B(X(x)(r))\rangle$, $r \in [0,s]$, is an element of $\mathcal{N}_W^2(0,s;\mathbb{R})$.
Therefore the uniqueness is proved. □

Remark 6.9. Section 2 of [AlKoRö 95] deals with the problem of essential self-adjointness of Dirichlet operators. There a special symmetric differential operator H_μ is considered. [AlKoRö 95, Theorem 1, p.107] provides conditions under which this operator fulfills Berezansky's abstract parabolic criterium of essential self-adjointness (see [Be 86, Theorem 6.13], [BeKo 88, Theorem 1.10]). By this criterium the question of essential self-adjointness is traced back to the question of the existence of a solution of the Kolmogorov equation associated with a stochastic differential equation of the type

$$
\text{(6.2)} \qquad \begin{cases} dX(t) &= \frac{1}{2}F(X(t))\,dt + dW(t) \quad t \in [0,1] \\ X(0) &= x \in U_1 \end{cases}
$$

To get this existence it is used that there is a solution of the stochastic equation (6.2) which depends regularly on initial data. Concerning the latter the authors refer to [Dal 67] which, as explained, does not contain proofs. But to justify the above statement one could use our Theorem 4.3 (iv) and

Theorem 5.3 (iv) which we have proved in detail. In fact they can be applied as the stochastic equation is considered within the following framework:
$(U, \langle\ ,\ \rangle_U)$ and $(U_1, \langle\ ,\ \rangle_1)$ are two separable Hilbert spaces such that $U \subset U_1$ and the embedding by the identity function $I : U \to U_1$ is Hilbert-Schmidt. $W(t)$, $t \in [0,1]$, is a cylindrical $Q = I_U$-Wiener process in U (with values in U_1, see subsection 1.5.1). Besides, in contrast to our setting, F has not bounded derivatives of first and second order but lies in

$$C^2_{pol}(U_1, U_1) := \{f : U_1 \to U_1 \mid f \text{ is twice continuously Fréchet differentiable and the derivatives } f^{(l)} \text{ are polynomially bounded}\}$$

such that there exists a constant $C \geq 0$ with $\langle F(x), x\rangle_1 \leq C(1 + \|x\|_1)^2$ and $\langle DF(x)y, y\rangle_1 \leq C\|y\|_1^2$ for all $x, y \in U_1$ (see [AlKoRö 95, Theorem 1, p.107]).
Instead of a strict solution which solves the Kolmogorov equation pointwisely it is asked for a strong solution as a function $u : [0,1] \to L^2(\mu)$. But to answer this question it is also used that there exists a strong solution $X(x)(t)$, $t \in [0,1]$, of problem (6.2) which is twice Fréchet differentiable with respect to the initial data $x \in U_1$ (see proof of Theorem 6.4 (i)).
This problem is reduced to our case that F is twice continuously Fréchet differentiable with bounded first and second order derivatives (see [AlKoRö 95, p.110] and for more details [Zü 95, Theorem 6.4, p.30]). Then it is easy to see that the conditions of Theorem 5.3 are fulfilled. Hence we first get the existence of a mild solution $X(x)(t)$, $t \in [0,1]$, of equation (6.2) which is twice continuously Fréchet differentiable with respect to the initial data. Secondly, our Proposition 2.10 and Proposition 2.9 provide that the mild solution is even a strong one.
Hence it is verified that the statement about the existence and the differentiability of X used in the proof of [AlKoRö 95, Theorem 1, p.107] is really true in this setting. For example this result plays an important role in the proof of [AlKoRö 95, Lemma 4, p.112] which says that $p_t\varphi \in C^2_{pol}(U_1, \mathbb{R})$ for all $\varphi \in C^2_{pol}(U_1, \mathbb{R})$. It is proved in a similar way as our Theorem 6.1.

Appendix A

The Bochner Integral

Let $(X, \| \ \|)$ be a Banach space, $\mathcal{B}(X)$ the Borel σ-field of X and $(\Omega, \mathcal{F}, \mu)$ a measure space with finite measure μ.

A.1 Definition of the Bochner integral

Step 1:
As first step we want to define the integral for simple functions which are defined as follows. Set

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

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

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

To get that $(\mathcal{E}, \| \ \|_{\mathcal{E}})$ is a normed vector space we consider equivalence classes with respect to $\| \ \|_{\mathcal{E}}$. For simplicity we will not change the notations.
For $f \in \mathcal{E}$ we define now the Bochner integral to be

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

In this way we get a mapping

$$\begin{aligned} \text{int} : (\mathcal{E}, \| \ \|_{\mathcal{E}}) &\to (X, \| \ \|) \\ f &\mapsto \int f \, d\mu \end{aligned}$$

which is linear and uniformly continuous since $\|\int f\, d\mu\| \leq \int \|f\|\, d\mu$ for all $f \in \mathcal{E}$.
Therefore we can extend the mapping int to the abstract completion of $\mathcal{E}$ with respect to $\| \|_{\mathcal{E}}$ which we denote by $\overline{\mathcal{E}}$.

Step 2: We give an explicit representation of $\overline{\mathcal{E}}$.

Definition A.1. A function $f : \Omega \to X$ is called strongly measurable if it is Borel measurable and $f(\Omega) \subset X$ is separable.

Definition A.2. Let $1 \leq p < \infty$. Then we define

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

and the semi-norm $\|f\|_{L^p} := \left(\int \|f\|^p\, d\mu\right)^{\frac{1}{p}}$, $f \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu; X)$. The space of all equivalence classes in $\mathcal{L}^p(\Omega, \mathcal{F}, \mu; X)$ with respect to $\| \|_{L^p}$ is denoted by $L^p(\Omega, \mathcal{F}, \mu; X) := L^p(\mu; X)$.

Claim: $L^1(\Omega, \mathcal{F}, \mu; X) = \overline{\mathcal{E}}$.
Step 1: $(L^1(\Omega, \mathcal{F}, \mu; X), \| \|_{L^1})$ is complete.
The proof is just a modification of the proof of the Fischer-Riesz theorem by the help of the following proposition.

Proposition A.3. *Let $(\Omega, \mathcal{F})$ be a measurable space and let X be a Banach space. Then*

(i) the set of Borel measurable functions from Ω to X is closed under the formation of pointwise limits, and

(ii) the set of strongly measurable functions from Ω to X is closed under the formation of pointwise limits.

Proof. [Co 80, Proposition E.1., p.350] □

Step 2: $\mathcal{E}$ is as dense subset of $L^1(\Omega, \mathcal{F}, \mu; X)$ with respect to $\| \|_{L^1}$.
This can be shown by the help of the following lemma.

Lemma A.4. *Let E be a metric space with metric d and let $f : \Omega \to E$ be strongly measurable. Then there exists a sequence f_n, $n \in \mathbb{N}$, of simple E-valued functions (i.e. f_n is $\mathcal{F}/\mathcal{B}(E)$-measurable and takes only a finite number of values) such that for arbitrary $\omega \in \Omega$ the sequence $d(f_n(\omega), f(\omega))$, $n \in \mathbb{N}$, is monotonely decreasing to zero.*

Proof. [DaPrZa 92, Lemma 1.1, p.16] Let $\{e_k | k \in \mathbb{N}\}$ be a countable dense subset of $f(\Omega)$. For $m \in \mathbb{N}$ define

$$\begin{aligned} d_m(\omega) &:= min\{d(f(\omega), e_k) | k \leq m\} \\ k_m(\omega) &:= min\{k \leq m | d_m(\omega) = d(f(\omega), e_k)\} \\ f_m(\omega) &:= e_{k_m(\omega)} \end{aligned}$$

Obviously f_m, $m \in \mathbb{N}$, are simple function since

$$f_m(\Omega) \subset \{e_1, e_2, ..., e_m\}$$

Moreover, by the density of $\{e_k | k \in \mathbb{N}\}$, the sequence $d_m(\omega)$, $m \in \mathbb{N}$, is monotonically decreasing to zero for arbitrary $\omega \in \Omega$. Since
$d(f_m(\omega), f(\omega)) = d_m(\omega)$ the assertion follows. □

Let now $f \in L^1(\mu; X)$. By the above Lemma A.4 we get the existence of a sequence of simple functions f_n, $n \in \mathbb{N}$, such that

$$\|f_n(\omega) - f(\omega)\| \downarrow 0 \text{ for all } \omega \in \Omega \text{ as } n \to \infty$$

Hence $f_n \underset{n\to\infty}{\longrightarrow} f$ in $\| \ \|_{L^1}$ by Lebesgue's dominated convergence theorem.

A.2 Properties of the Bochner integral

Proposition A.5 (Bochner inequality). *Let $f \in L^1(\Omega, \mathcal{F}, \mu; X)$. Then*

$$\| \int f \, d\mu \| \leq \int \|f\| \, d\mu$$

Proof. If $f \in \mathcal{E}$ the assertion is obvious.
Otherwise there exists a sequence of simple functions f_n, $n \in \mathbb{N}$, such that $\|f_n - f\|_{L^1} \underset{n\to\infty}{\longrightarrow} 0$.
Since $int : L^1(\mu; X) \to X$ and $\| \ \|_{L^1} : L^1(\mu; X) \to \mathbb{R}$ are continuous we get

$$\| \int f \, d\mu \| = \lim_{n\to\infty} \| \int f_n \, d\mu \| \leq \lim_{n\to\infty} \int \|f_n\| \, d\mu = \int \|f\| \, d\mu$$

□

Proposition A.6. *Let $f \in L^1(\Omega, \mathcal{F}, \mu; X)$. Then*

$$\int \varphi \circ f \, d\mu = \varphi\Big(\int f \, d\mu\Big)$$

holds for all $\varphi \in X^* = L(X, \mathbb{R})$.

Proof. [Co 80, Proposition E.11, p.356] □

Proposition A.7 (Fundamental theorem). *Let* $-\infty < a < b < \infty$ *and* $f \in C^1([a,b];X)$. *Then*

$$f(t) - f(s) = \int_s^t f'(u)\, du := \begin{cases} \int 1_{[s,t]}(u) f'(u)\, du & \text{if } s \le t \\ -\int 1_{[t,s]}(u) f'(u)\, du & \text{otherwise} \end{cases}$$

for all $s, t \in [a,b]$ *where* du *denotes the Lebesgue measure on* $\mathcal{B}(\mathbb{R})$.

Proof. Claim 1: If we set $F(t) = \int_s^t f'(u)\, du$, $t \in [a,b]$, we get that $F'(t) = f'(t)$ for all $t \in [a,b]$.
For that we have to prove that

$$\|\frac{1}{h}(F(t+h) - F(t)) - f'(t)\|_E \underset{h \to 0}{\longrightarrow} 0$$

For this end we fix $t \in [a,b]$ and take an arbitrary $\varepsilon > 0$. Since f' is continuous on $[a,b]$ there exists $\delta > 0$ such that $\|f'(u) - f'(t)\|_E < \varepsilon$ for all $u \in [a,b]$ with $|u-t| < \delta$. Then we obtain that

$$\|\frac{1}{h}(F(t+h) - F(t)) - f'(t)\|_E = \|\frac{1}{h} \int_t^{t+h} f'(u) - f'(t)\, du\|_E$$
$$\le \frac{1}{h} \int_t^{t+h} \|f'(u) - f'(t)\|_E\, du < \varepsilon$$

if $t + h \in [a,b]$ and $|h| < \delta$.
Claim 2: If $\tilde{F} \in C^1([a,b];E)$ is a further function with $\tilde{F}' = F' = f'$ then there exists a constant c such that $F - \tilde{F} = c$.
For all $L \in E^* = L(E,\mathbb{R})$ we define $g_L := L(F - \tilde{F})$. Then $g_L' = 0$ and therefore g_L is constant. Since E^* separates the points of E by the Hahn-Banach theorem (see [Al 92, 4.2 Satz, p.114]) this implies that $F - \tilde{F}$ itself is constant. □

Appendix B

Nuclear and Hilbert-Schmidt Operators

Let $(U, \langle\ ,\ \rangle_U)$ and $(H, \langle\ ,\ \rangle)$ be two separable Hilbert spaces. The space of all bounded linear operators from U to H is denoted by $L(U,H)$ for simplicity we write $L(U)$ instead of $L(U,U)$. If we speak of the adjoint operator of $L \in L(U,H)$ we write $L^* \in L(H,U)$. An element $L \in L(U)$ is called symmetric if $\langle Lu, v\rangle_U = \langle u, Lv\rangle_U$ for all $u, v \in U$. In addition, $L \in L(U,H)$ is called nonnegative if $\langle Lu, x\rangle \geq 0$ for all $u \in U$ and $x \in H$.

Definition B.1 (Nuclear operator). An element $T \in L(U,H)$ is said to be a nuclear operator if there exists a sequence $(a_j)_{j\in\mathbb{N}}$ in H and a sequence $(b_j)_{j\in\mathbb{N}}$ in U such that

$$Tx = \sum_{j\in\mathbb{N}} a_j \langle b_j, x\rangle_U \quad \text{for all } x \in U$$

and

$$\sum_{j\in\mathbb{N}} \|a_j\| \|b_j\|_U < \infty$$

The space of all nuclear operators from U to H is denoted by $L_1(U,H)$.

Proposition B.2. *The space $L_1(U,H)$ endowed with the norm*

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

is a Banach space.

Proof. [MeVo 92, 16.25 Corollar, p.154] □

Definition B.3. Let $T \in L(U)$ and let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U. Then we define

$$\operatorname{tr} T := \sum_{k \in \mathbb{N}} \langle T e_k, e_k \rangle_U$$

if the series is convergent.

One has to notice that this definition could depend on the choice of the orthonormal basis. But there is the following result concerning nuclear operators.

Remark B.4. If $T \in L_1(U)$ then $\operatorname{tr} T$ is well defined independently of the choice of the orthonormal basis e_k, $k \in \mathbb{N}$. Moreover we have that $|\operatorname{tr} T| \leq \|T\|_{L_1(U)}$.

Proof. Let $(a_j)_{j \in \mathbb{N}}$ and $(b_j)_{j \in \mathbb{N}}$ be sequences in U such that $Tx = \sum_{j \in \mathbb{N}} a_j \langle b_j, x \rangle_U$ for all $x \in U$ and $\sum_{j \in \mathbb{N}} \|a_j\|_U \|b_j\|_U < \infty$.
Then we get for any orthonormal basis e_k, $k \in \mathbb{N}$, of U that

$$\langle T e_k, e_k \rangle_U = \sum_{j \in \mathbb{N}} \langle e_k, a_j \rangle_U \langle e_k, b_j \rangle_U$$

and therefore

$$\begin{aligned}
\sum_{k \in \mathbb{N}} |\langle T e_k, e_k \rangle_U| &\leq \sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{N}} |\langle e_k, a_j \rangle_U \langle e_k, b_j \rangle_U| \\
&\leq \sum_{j \in \mathbb{N}} \Big(\sum_{k \in \mathbb{N}} |\langle e_k, a_j \rangle_U|^2 \Big)^{\frac{1}{2}} \Big(\sum_{k \in \mathbb{N}} |\langle e_k, b_j \rangle_U|^2 \Big)^{\frac{1}{2}} \\
&= \sum_{j \in \mathbb{N}} \|a_j\|_U \|b_j\|_U < \infty
\end{aligned}$$

This implies that we can exchange the summation to get that

$$\begin{aligned}
\sum_{k \in \mathbb{N}} \langle T e_k, e_k \rangle_U &= \sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{N}} \langle e_k, a_j \rangle_U \langle e_k, b_j \rangle_U \\
&= \sum_{j \in \mathbb{N}} \langle a_j, b_j \rangle_U
\end{aligned}$$

and the assertion follows. □

Definition B.5 (Hilbert-Schmidt operator). A bounded linear operator $T : U \to H$ is called Hilbert-Schmidt if

$$\sum_{k \in \mathbb{N}} \|Te_k\|^2 < \infty$$

where e_k, $k \in \mathbb{N}$, is an orthonormal basis of U.
The space of all Hilbert-Schmidt operators from U to H is denoted by $L_2(U, H)$.

Remark B.6. (i) The definition of Hilbert-Schmidt operator and the number $\|T\|^2_{L_2(U,H)} := \sum_{k \in \mathbb{N}} \|Te_k\|^2$ does not depend on the choice of the orthonormal basis e_k, $k \in \mathbb{N}$, and we have that $\|T\|_{L_2(U,H)} = \|T^*\|_{L_2(H,U)}$. For simplicity we also write $\|T\|_{L_2}$ instead of $\|T\|_{L_2(U,H)}$.
(ii) $\|T\|_{L(U,H)} \leq \|T\|_{L_2(U,H)}$

Proof. (i): If e_k, $k \in \mathbb{N}$, is an orthonormal basis of U and f_k, $k \in \mathbb{N}$, is an orthonormal basis of H we obtain by the Parseval identity that

$$\sum_{k \in \mathbb{N}} \|Te_k\|^2 = \sum_{k \in \mathbb{N}} \sum_{j \in \mathbb{N}} |\langle Te_k, f_j \rangle|^2 = \sum_{j \in \mathbb{N}} \|T^* f_j\|^2_U$$

and therefore the assertion follows.
(ii): Let $x \in U$ and f_k, $k \in \mathbb{N}$, be an orthonormal basis of H. Then we get that

$$\|Tx\|^2 = \sum_{k \in \mathbb{N}} \langle Tx, f_k \rangle^2 \leq \|x\|^2_U \sum_{k \in \mathbb{N}} \|T^* f_k\|^2_U = \|T\|^2_{L_2(U,H)} \|x\|^2$$

□

Proposition B.7. *Let $S, T \in L_2(U, H)$ and let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U. If we define*

$$\langle T, S \rangle_{L_2} := \sum_{k \in \mathbb{N}} \langle Se_k, Te_k \rangle$$

we obtain that $(L_2(U, H), \langle\ ,\ \rangle_{L_2})$ is a separable Hilbert space.
If f_k, $k \in \mathbb{N}$, is an orthonormal basis of H we get that $f_j \otimes e_k := f_j \langle e_k, \cdot \rangle_U$, $j, k \in \mathbb{N}$, is an orthonormal basis of $L_2(U, H)$.

Proof. We have to prove the completeness and the separability.
1. $L_2(U, H)$ is complete:
Let T_n, $n \in \mathbb{N}$, be a Cauchy sequence in $L_2(U, H)$. Then it is clear that it is

also a Cauchy sequence in $L(U,H)$. Because of the completeness of $L(U,H)$ there exists an element $T \in L(U,H)$ such that $\|T_n - T\|_{L(U,H)} \longrightarrow 0$ as $n \to \infty$. But by the lemma of Fatou we also have for any orthonormal basis e_k, $k \in \mathbb{N}$, of U that

$$\begin{aligned}\|T_n - T\|_{L_2}^2 &= \sum_{k\in\mathbb{N}} \langle (T_n - T)e_k, (T_n - T)e_k\rangle \\ &= \sum_{k\in\mathbb{N}} \liminf_{m\to\infty} \|(T_n - T_m)e_k\|^2 \\ &\le \liminf_{m\to\infty} \sum_{k\in\mathbb{N}} \|(T_n - T_m)e_k\|^2 = \liminf_{m\to\infty} \|T_n - T_m\|_{L_2}^2 < \varepsilon\end{aligned}$$

for all $n \in \mathbb{N}$ big enough. Therefore the assertion follows.

2. $L_2(U,H)$ is separable:

If we define $f_j \otimes e_k := f_j \langle e_k, \cdot\rangle_U$, $j,k \in \mathbb{N}$, then it is clear that $f_j \otimes e_k \in L_2(U,H)$ for all $j,k \in \mathbb{N}$ and for arbitrary $T \in L_2(U,H)$ we get that

$$\langle f_j \otimes e_k, T\rangle_{L_2} = \sum_{n\in\mathbb{N}} \langle e_k, e_n\rangle_U \langle f_j, Te_n\rangle = \langle f_j, Te_k\rangle$$

Therefore it is obvious that $f_j \otimes e_k$, $j,k \in \mathbb{N}$, is an orthonormal system. In addition, $T = 0$ if $\langle f_j \otimes e_k, T\rangle_{L_2} = 0$ for all $j,k \in \mathbb{N}$, and therefore span $(f_j \otimes e_k \mid j,k \in \mathbb{N})$ is a dense subspace of $L_2(U,H)$. □

Proposition B.8. *Let $(G, \langle\ ,\ \rangle_G)$ be a further separable Hilbert space. If $T \in L_2(U,H)$ and $S \in L_2(H,G)$ then $ST \in L_1(U,G)$ and*

$$\|ST\|_{L_1(U,G)} \le \|S\|_{L_2} \|T\|_{L_2}$$

Proof. Let f_k, $k \in \mathbb{N}$, be an orthonormal basis of H. Then we have that

$$STx = \sum_{k\in\mathbb{N}} \langle Tx, f_k\rangle S f_k, \quad x \in U$$

and therefore

$$\begin{aligned}\|ST\|_{L_1(U,G)} &\le \sum_{k\in\mathbb{N}} \|T^* f_k\|_U \|S f_k\|_G \\ &\le \Big(\sum_{k\in\mathbb{N}} \|T^* f_k\|_U^2\Big)^{\frac{1}{2}} \Big(\sum_{k\in\mathbb{N}} \|S f_k\|_G^2\Big)^{\frac{1}{2}} \\ &= \|S\|_{L_2} \|T\|_{L_2}\end{aligned}$$

□

Remark B.9. Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U. If $T \in L(U)$ is symmetric, nonnegative with $\sum_{k\in\mathbb{N}} \langle Te_k, e_k\rangle_U < \infty$ then $T \in L_1(U)$.

Proof. The result is obvious by the previous proposition and the fact that there exists $T^{\frac{1}{2}} \in L(U)$ nonnegative and symmetric such that $T = T^{\frac{1}{2}}T^{\frac{1}{2}}$ (see Proposition 1.24). Then $T^{\frac{1}{2}} \in L_2(U)$. □

Proposition B.10. *Let $L \in L(H)$ and $B \in L_2(U,H)$. Then $LBB^* \in L_1(H)$, $B^*LB \in L_1(U)$ and we have that*

$$\operatorname{tr} LBB^* = \operatorname{tr} B^*LB$$

Proof. Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of U and let f_k, $k \in \mathbb{N}$, be an orthonormal basis of H. Then the Parseval identity provides that

$$\begin{aligned}
&\sum_{k\in\mathbb{N}}\sum_{n\in\mathbb{N}} |\langle f_k, Be_n\rangle\langle f_k, LBe_n\rangle| \\
&\le \sum_{n\in\mathbb{N}} \Big(\sum_{k\in\mathbb{N}} |\langle f_k, Be_n\rangle|^2\Big)^{\frac{1}{2}} \Big(\sum_{k\in\mathbb{N}} |\langle f_k, LBe_n\rangle|^2\Big)^{\frac{1}{2}} \\
&= \|Be_n\|\|LBe_n\| \le \|L\|_{L(H)}\|B\|_{L_2}
\end{aligned}$$

Therefore it is allowed to exchange the summation to obtain that

$$\begin{aligned}
&\operatorname{tr} LBB^* \\
&= \sum_{k\in\mathbb{N}} \langle LBB^* f_k, f_k\rangle = \sum_{k\in\mathbb{N}} \langle B^* f_k, B^*L^* f_k\rangle_U \\
&= \sum_{k\in\mathbb{N}}\sum_{n\in\mathbb{N}} \langle B^* f_k, e_n\rangle_U \langle B^*L^* f_k, e_n\rangle_U \\
&= \sum_{n\in\mathbb{N}}\sum_{k\in\mathbb{N}} \langle f_k, Be_n\rangle\langle f_k, LBe_n\rangle \\
&= \sum_{n\in\mathbb{N}} \langle Be_n, LBe_n\rangle = \sum_{n\in\mathbb{N}} \langle e_n, B^*LBe_n\rangle_U \\
&= \operatorname{tr} B^*LB
\end{aligned}$$

□

Appendix C

Pseudo Inverse of Linear Operators

Let $(U, \langle\ ,\ \rangle_U)$ and $(H, \langle\ ,\ \rangle)$ be two Hilbert spaces.

Definition C.1 (Pseudo inverse). Let $T \in L(U, H)$ and $\text{Ker}\,(T) := \{x \in U | Tx = 0\}$. The pseudo inverse of T is defined as

$$T^{-1} := (T_{|_{\text{Ker}\,(T)^\perp}})^{-1} : T(\text{Ker}\,(T)^\perp) = T(U) \to \text{Ker}\,(T)^\perp$$

Remark C.2. (i) There is an equivalent way of defining the pseudo inverse of a linear operator $T \in L(U, H)$. For $x \in T(U)$ one sets $T^{-1}x \in U$ to be the solution of minimal norm of the equation $Ty = x$, $y \in U$.
(ii) If $T \in L(U, H)$ then $T^{-1} : T(U) \to U$ is linear.

Proposition C.3. *Let $T \in L(U)$ and T^{-1} the pseudo inverse of T.*

(i) If we define an inner product on $T(U)$ by

$$\langle x, y\rangle_{T(U)} := \langle T^{-1}x, T^{-1}y\rangle_U \quad \textit{for all } x, y \in T(U),$$

then $(T(U), \langle\ ,\ \rangle_{T(U)})$ is a Hilbert space.

(ii) Let e_k, $k \in \mathbb{N}$, be an orthonormal basis of $(\text{Ker}\, T)^\perp$. Then Te_k, $k \in \mathbb{N}$, is an orthonormal basis of $(T(U), \langle\ ,\ \rangle_{T(U)})$.

Proof. (i): Let x_n, $n \in \mathbb{N}$, be a Cauchy sequence in $(T(U), \langle\ ,\ \rangle_{T(U)})$. This implies that $T^{-1}x_n$, $n \in \mathbb{N}$, is a Cauchy sequence in U and therefore there exists $u \in U$ such that $T^{-1}x_n \longrightarrow u$ as $n \to \infty$.
Since $T^{-1} : T(U) \to \text{Ker}\,(T)^\perp$ we have that each $T^{-1}x_n$ is an element of $\text{Ker}\,(T)^\perp$ which is a closed subspace of U.

Hence we get that $u = \lim_{n\to\infty} T^{-1}x_n \in \operatorname{Ker}(T)^\perp$. In this case there exists $x \in T(U)$ such that $u = T^{-1}x$ and so finally we obtain

$$\|x_n - x\|_{T(U)} = \|T^{-1}x_n - T^{-1}x\|_U = \|T^{-1}x_n - u\|_U \underset{n\to\infty}{\longrightarrow} 0$$

□

Now we want to present a result about the images of linear operators. To this end we need the following lemma.

Lemma C.4. *Let $T \in L(U, H)$. Then the set $\{Tu \mid u \in U, \|u\|_U \le c\}$, $c \ge 0$, is convex and closed.*

Proof. Since T is linear it is obvious that the set is convex.
Let $J : H \to H' = L(H, \mathbb{R},\ x \mapsto \langle x, \cdot\rangle$ be the Riesz isomorphism. Then

$$\begin{aligned}\{Tu \mid u \in U, \|u\|_U \le c\} &= J^{-1}(\{\langle Tu, \cdot\rangle \mid u \in U, \|u\|_U \le c\})\\ &= J^{-1}(\underbrace{\{\langle u, T^*(\cdot)\rangle_U \mid u \in U, \|u\|_U \le c\}}_{=:\,M})\end{aligned}$$

Since J is continuous it is sufficient to prove that M is closed.
To this end let $\langle u_n, T^*(\cdot)\rangle_U$, $n \in \mathbb{N}$, be a sequence in M which is convergent with respect to the operator-norm.
Since U is reflexive $\overline{B_c(0)} \subset U$ is weakly sequentially compact (see [Al 92, 5.7 Satz, p.162]). Therefore there exists a subsequence u_{n_l}, $l \in \mathbb{N}$, and $u \in \overline{B_c(0)} \subset U$ such that

$$\langle u_{n_l}, v\rangle_U \underset{l\to\infty}{\longrightarrow} \langle u, v\rangle_U \quad \text{for all } v \in U.$$

In particular

$$\langle u_{n_l}, T^*x\rangle_U \underset{n\to\infty}{\longrightarrow} \langle u, T^*x\rangle_U \quad \text{for all } x \in H.$$

Since $\langle u_n, T^*(\cdot)\rangle$, $n \in \mathbb{N}$, is convergent in H' with respect to the operator-norm we obtain that

$$\langle u_n, T^*(\cdot)\rangle_U \underset{n\to\infty}{\longrightarrow} \langle u, T^*(\cdot)\rangle_U \quad \text{in } H' \text{ and } \|u\|_U \le c.$$

□

Proposition C.5. *Let $(U_1, \langle\ ,\ \rangle_1)$ and $(U_2, \langle\ ,\ \rangle_2)$ be two Hilbert spaces. In addition, we take $T_1 \in L(U_1, H)$ and $T_2 \in L(U_2, H)$. Then the following statements hold.*

(i) If there exists a constant $c \geq 0$ such that $\|T_1^* x\| \leq c\|T_2^* x\|$ for all $x \in H$ then $\{T_1 u \mid u \in U_1, \|u\|_1 \leq 1\} \subset \{T_2 v \mid v \in U_2, \|v\|_2 \leq c\}$. In particular, this implies that $\operatorname{Im} T_1 \subset \operatorname{Im} T_2$.

(ii) If $\|T_1^* x\|_1 = \|T_2^* x\|_2$ for all $x \in H$ then $\operatorname{Im} T_1 = \operatorname{Im} T_2$ and $\|T_1^{-1} x\|_1 = \|T_2^{-1} x\|_2$ for all $x \in \operatorname{Im} T_1$.

Proof. [DaPrZa 92, Proposition B.1, p.407] (i): Assume that there exists $u_0 \in U_1$ such that

$$\|u_0\|_1 \leq 1 \text{ and } T_1 u_0 \notin \{T_2 v \mid v \in U_2, \|v\|_2 \leq c\}.$$

By Lemma C.4 we know that the set $\{T_2 v \mid v \in U_2, \|v\|_2 \leq c\}$ is closed and convex. Therefore we get by the separation theorem (see [Al 92, 5.11 Trennungssatz, p.166]) there exists $x \in H$, $x \neq 0$, such that

$$1 < \langle x, T_1 u_0 \rangle \text{ and } \langle x, T_2 v \rangle \leq 1 \text{ for all } v \in U_2 \text{ with } \|v\|_2 \leq c$$

Thus $\|T_1^* x\|_1 > 1$ and $c\|T_2^* x\|_2 = \sup_{\|v\|_2 \leq c} |\langle T_2^* x, v\rangle_2| \leq 1$, a contradiction.

(ii): By (i) we know that $\operatorname{Im} T_1 = \operatorname{Im} T_2$. It remains to verify that $\|T_1^{-1} x\| = \|T_2^{-1} x\|$ for all $x \in \operatorname{Im} T_1$.
If $x = 0$ then $\|T_1^{-1} 0\| = 0 = \|T_2^{-1} 0\|$.
If $x \in \operatorname{Im} T_1 \setminus \{0\}$ there exist $u_1 \in (\operatorname{Ker} T_1)^\perp$ and $u_2 \in (\operatorname{Ker} T_2)^\perp$ such that $x = T_1 u_1 = T_2 u_2$. We have to show that $\|u_1\|_1 = \|u_2\|_2$.
Assume that $\|u_1\|_1 > \|u_2\|_2 > 0$. Then (i) implies that

$$\frac{x}{\|u_2\|_2} = T_2\left(\frac{u_2}{\|u_2\|_2}\right) \in \{T_2 v \mid v \in U_2, \|v\|_2 \leq 1\} = \{T_1 u \mid u \in U_1, \|u\|_1 \leq 1\}$$

But $\frac{x}{\|u_2\|_2} = T_1\left(\frac{u_1}{\|u_2\|_2}\right)$ and $\|\frac{u_1}{\|u_2\|_2}\|_1 > 1$, therefore

$$\frac{x}{\|u_2\|_2} \notin \{T_1 u | u \in U_1, \|u\|_1 \leq 1\}.$$

That is a contradiction. □

Corollary C.6. *Let $T \in L(U, H)$ and set $Q := TT^* \in L(H)$. Then we have*

$$\operatorname{Im} Q^{\frac{1}{2}} = \operatorname{Im} T \quad \textit{and} \quad \|Q^{\frac{-1}{2}} x\| = \|T^{-1} x\|_U, \quad \textit{for all } x \in \operatorname{Im} T,$$

where $\tilde{Q}^{\frac{1}{2}}$ is the pseudo inverse of $Q^{\frac{1}{2}}$.

Proof. Since by Lemma 1.24 $Q^{\frac{1}{2}}$ is symmetric we have for all $x \in H$ that

$$\|Q^{\frac{1}{2}} x\|^2 = \langle Qx, x\rangle = \langle TT^* x, x\rangle = \|T^* x\|_U^2.$$

Therefore the assertion follows by Proposition C.5. □

Appendix D

Continuity and Differentiability of Implicit Functions

At the beginning of this part of the appendix we fix four Banach spaces $(E_0, \| \|_0)$, $(E, \| \|)$, $(\Lambda_0, \| \|_{\Lambda_0})$ and $(\Lambda, \| \|_\Lambda)$ where we assume that $(E_0, \| \|_0)$ is continuously embedded in $(E, \| \|)$ and $(\Lambda_0, \| \|_{\Lambda_0})$ is continuously embedded in $(\Lambda, \| \|_\Lambda)$.
For the whole section we consider a mapping $G : \Lambda \times E \to E$ with the following properties:

1. G can be restricted such that $G : \Lambda_0 \times E_0 \to E_0$.

2. There exists an $\alpha \in [0, 1[$ such that

   $$\|G(\lambda, x) - G(\lambda, y)\| \leq \alpha \|x - y\| \qquad \text{for all } \lambda \in \Lambda \text{ and all } x, y \in E$$

   and

   $$\|G(\lambda_0, x_0) - G(\lambda_0, y_0)\|_0 \leq \alpha \|x_0 - y_0\|_0 \qquad \text{for all } \lambda_0 \in \Lambda_0 \text{ and all } x_0, y_0 \in E_0$$

Then we get by the contraction theorem that there exists exactly one mapping $\varphi : \Lambda \to E$ such that
$\varphi(\lambda) = G(\lambda, \varphi(\lambda))$ for all $\lambda \in \Lambda$ and $\varphi : \Lambda_0 \to E_0$.

D.1 Continuity of the implicit function

To prove the continuity of the implicit function φ it is sufficient to consider the case that $E = E_0$ and $\Lambda = \Lambda_0$.

Theorem D.1 (Continuity of the implicit function). *(i) If we assume in addition that the mapping $\lambda \mapsto G(\lambda, x)$ is continuous from Λ to E for all $x \in E$ we get that $\varphi : \Lambda \to E$ is continuous.*

(ii) If the mappings $\lambda \mapsto G(\lambda, x)$ are not only continuous from Λ to E for all $x \in E$ but there even exists a $L \geq 0$ such that
$\|G(\lambda, x) - G(\tilde{\lambda}, x)\|_E \leq L\|\lambda - \tilde{\lambda}\|_\Lambda$ for all $x \in E$
then the mapping $\varphi : \Lambda \to E$ is Lipschitz continuous.

Proof. (i): We fix $\lambda_0 \in \Lambda$. Then for any other $\lambda \in \Lambda$

$$\begin{aligned}\varphi(\lambda) - \varphi(\lambda_0) &= G(\lambda, \varphi(\lambda)) - G(\lambda_0, \varphi(\lambda_0)) \\ &= [G(\lambda, \varphi(\lambda)) - G(\lambda, \varphi(\lambda_0))] + [G(\lambda, \varphi(\lambda_0)) - G(\lambda_0, \varphi(\lambda_0))]\end{aligned}$$

Because of the contraction property we obtain that

$$\|\varphi(\lambda) - \varphi(\lambda_0)\|_E \leq \alpha\|\varphi(\lambda) - \varphi(\lambda_0)\|_E + \|G(\lambda, \varphi(\lambda_0)) - G(\lambda_0, \varphi(\lambda_0))\|_E$$

and for that reason

$$\|\varphi(\lambda) - \varphi(\lambda_0)\|_E \leq \frac{1}{1-\alpha}\|G(\lambda, \varphi(\lambda_0)) - G(\lambda_0, \varphi(\lambda_0))\|_E$$

Therefore we get the result (i).
(ii): In the same way as in (i) we can show that for arbitrary λ and $\tilde{\lambda} \in \Lambda$

$$\|\varphi(\lambda) - \varphi(\tilde{\lambda})\|_E \leq \frac{1}{1-\alpha}\|G(\lambda, \varphi(\tilde{\lambda})) - G(\tilde{\lambda}, \varphi(\tilde{\lambda}))\|_E \leq \frac{L}{1-\alpha}\|\lambda - \tilde{\lambda}\|_\Lambda$$

by the additional Lipschitz property of the mapping G. □

D.2 Different concepts of differentiability in general Banach spaces

Let E_1 and E_2 be two real Banach spaces and let $H : E_1 \to E_2$.

Definition D.2 (Directional derivatives). H is said to be *differentiable in the point $x_0 \in E_1$ and in the direction $y \in E_1$* if there exists an element $\partial H(x_0; y) \in E_2$ such that

$$\partial H(x_0; y) = \lim_{h \to \infty} \frac{H(x_0 + hy) - H(x_0)}{h}.$$

$\partial H(x_0; y)$ is called the directional derivative of H (in x_0 and direction y).

Definition D.3 (Gâteaux differentiability). H is said to be Gâteaux differentiable in $x_0 \in E_1$ if there exist all directional derivatives $\partial H(x_0; y)$, $y \in E_1$, and if $\partial H(x_0; \cdot) \in L(E_1, E_2)$. Then we write $\partial H(x_0)y$ instead of $\partial H(x_0; y)$ and $\partial H(x_0)$ is called *Gâteaux derivative* of H in x_0.
If $H : E_1 \to E_2$ is Gâteaux differentiable in each $x \in E_1$ we call H simply Gâteaux differentiable.

Lemma D.4. *(i) If $H : E_1 \to E_2$ is differentiable in $x_0 \in E_1$ and in direction $y \in E_1$ then there exist all directional derivatives $\partial H(x_0; \lambda y)$, $\lambda \in \mathbb{R}$, and*

$$\partial H(x_0; \lambda y) = \lambda \partial H(x_0; y)$$

(ii) If there exist all directional derivatives $\partial H(x; y)$, $x, y \in E_1$, such that the mapping $x \mapsto \partial H(x; y)$ is continuous from E_1 to E_2 for each $y \in E_1$ then $\partial H(x; \cdot)$ is additive for all $x \in E_1$, i.e.

$$\partial H(x; y_1 + y_2) = \partial H(x; y_1) + \partial H(x; y_2) \quad \textit{for all } x, y_1, y_2 \in E_1$$

Proof. (i): Because of the definition of the directional derivative the assertion is clear for $\lambda = 0$ and for $\lambda \neq 0$ we get that

$$\begin{aligned}\partial H(x_0; \lambda y) &= \lim_{h \to 0} \frac{H(x_0 + h\lambda y) - H(x_0)}{h} \\ &= \lim_{h \to 0} \frac{\lambda (H(x_0 + h\lambda y) - H(x_0))}{\lambda h} \\ &= \lambda \lim_{h \to 0} \frac{H(x_0 + hy) - H(x_0)}{h} = \lambda \partial H(x_0; y)\end{aligned}$$

(ii): For $x_0, y_1, y_2 \in E_1$ and $h \in \mathbb{R} \setminus \{0\}$

$$\begin{aligned}&\frac{H(x_0 + h(y_1 + y_2)) - H(x_0)}{h} \\ &= \frac{H(x_0 + hy_1 + hy_2) - H(x_0 + hy_1)}{h} + \frac{H(x_0 + hy_1) - H(x_0)}{h}\end{aligned}$$

While it is clear that the second term converges to $\partial H(x_0; y_1)$ as $h \to 0$ we need the fundamental Theorem A.7 to obtain that

$$\|\frac{H(x_0 + hy_1 + hy_2) - H(x_0 + hy_1)}{h} - \partial H(x_0; y_2)\|_2 \longrightarrow 0 \text{ as } h \to 0 \ :$$

Let $\varepsilon > 0$. As the mapping $x \mapsto \partial H(x; y_2)$ is continuous from E_1 to E_2 there exists a $\delta > 0$ such that $\|\partial H(x_0 + hy_1 + sy_2; y_2) - \partial H(x_0; y_2)\|_2 < \varepsilon$ for all $0 \leq h \leq \delta$ and $0 \leq s \leq \delta$. Then we get by the fundamental theorem for Bochner integrals, Theorem A.7, and by the Bochner inequality A.5 that

$$\begin{aligned}
&\|\frac{H(x_0 + hy_1 + hy_2) - H(x_0 + hy_1)}{h} - \partial H(x_0; y_2)\|_2 \\
&= \|\frac{1}{h}\int_0^h \partial H(x_0 + hy_1 + sy_2; y_2) - \partial H(x_0; y_2)\, ds\|_2 \\
&\leq \frac{1}{h}\int_0^h \|\partial H(x_0 + hy_1 + sy_2; y_2) - \partial H(x_0; y_2)\|_2\, ds < \varepsilon
\end{aligned}$$

for all $0 \leq h \leq \delta$. So finally

$$\begin{aligned}
&\partial H(x_0; y_1 + y_2) \\
&= \lim_{h\to 0} \frac{H(x_0 + hy_1 + hy_2) - H(x_0 + hy_1)}{h} + \lim_{h\to 0} \frac{H(x_0 + hy_1) - H(x_0)}{h} \\
&= \partial H(x_0; y_2) + \partial H(x_0; y_1)
\end{aligned}$$

□

Definition D.5 (Fréchet differentiability). Let E_1 and E_2 be two Banach spaces. A mapping $H : E_1 \to E_2$ is said to be *Fréchet differentiable* in $x_0 \in E_1$ if there exists an element $DH(x_0) \in L(E_1, E_2)$ such that

$$H(x_0 + y) = H(x_0) + DH(x_0)y + \circ(x_0, y) \text{ with } \frac{\circ(x_0, y)}{\|y\|_1} \longrightarrow 0 \text{ as } \|y\|_1 \to 0$$

$DH(x_0)$ is called the Féchet derivative of H in x_0.
If $H : E_1 \to E_2$ is Fréchet differentiable in each $x \in E_1$ we call H simply Fréchet differentiable.
H is said to be continuously Fréchet differentiable if $DH : E_1 \to L(E_1, E_2)$ is continuous.

Proposition D.6. *Let E_1 and E_2 be two Banach spaces and $H : E_1 \to E_2$ be a Gâteaux differentiable mapping.*
If the mapping $x \mapsto \partial H(x)$ is continuous from E_1 to $L(E_1, E_2)$ then H is even Fréchet differentiable with $\partial H(x) = DH(x)$ for all $x \in E_1$.

Proof. We have to show that for each $x_0 \in E_1$

$$\frac{\|H(x_0 + y) - H(x_0) - \partial H(x_0)y\|_2}{\|y\|_1} \longrightarrow 0 \quad \text{as } \|y\|_1 \to 0.$$

For this end let $\varepsilon > 0$. Then there exists $\delta > 0$ such that $\|\partial H(x_0 + y) - \partial H(x_0)\|_{L(E_1,E_2)} < \varepsilon$ for all $y \in E_1$ with $\|y\|_1 < \delta$.
Because of the continuity of the mapping $s \mapsto \partial H(x_0 + sy)y \in E_2$ we get by the fundamental theorem for Bochner integrals, Theorem A.7, that

$$H(x_0 + y) - H(x_0) = \int_0^1 \partial H(x_0 + sy)y \, ds$$

Hence we obtain by the Bochner inequality that for all $y \in E_1$ with $\|y\|_1 < \delta$

$$\begin{aligned} &\|H(x_0 + y) - H(x_0) - \partial H(x_0)y\|_2 \\ &= \|\int_0^1 \partial H(x_0 + sy)y - \partial H(x_0)y \, ds\|_2 \\ &\leq \int_0^1 \|\partial H(x_0 + sy)y - \partial H(x_0)y\|_2 \, ds \\ &\leq \int_0^1 \|\partial H(x_0 + sy) - \partial H(x_0)\|_{L(E_1,E_2)} \, ds \, \|y\|_1 \\ &\leq \varepsilon \|y\|_1 \end{aligned}$$

Therefore the assertion is proved. □

Concerning the second order derivatives we introduce the following definitions.

Definition D.7 (Second order derivatives). H is twice differentiable in $x_0 \in E_1$ in the directions $y_1 \in E_1$ and $y_2 \in E_1$ if there exists

$$\lim_{h \to 0} \frac{\partial H(x_0 + hy_2; y_1) - \partial H(x_0; y_1)}{h} =: \partial^2 H(x_0; y_1, y_2) \in E_2$$

H is twice Gâteaux differentiable in $x_0 \in E_1$ if $\partial H : E_1 \to L(E_1, E_2)$ is Gâteaux differentiable in x_0. Analogously to the above notation we write

$$\lim_{h \to 0} \frac{\partial H(x_0 + hy_2) - \partial H(x_0)}{h} =: \underbrace{\partial^2 H(x_0)}_{\in L(E_1, L(E_1, E_2))}(\cdot)y_2 \in L(E_1, E_2)$$

H is twice Fréchet differentiable in $x_0 \in E_1$ if $DH : E_1 \to L(E_1, E_2)$ is Fréchet differentiable in x_0, i.e. there exists $D^2 H(x_0) \in L(E_1, L(E_1, E_2))$ such that

$$DH(x_0 + y) = DH(x_0) + D^2 H(x_0)(\cdot)y + \circ(x_0, y)$$

where $\dfrac{\circ(x_0, y)}{\|y\|_1} \longrightarrow 0$ as $\|y\|_1 \to 0$. H is called twice Fréchet differentiable if it is twice Fréchet differentiable in each point $x \in E_1$. If $D^2 H : E_1 \to L(E_1, L(E_1, E_2))$ is continuous H is said to be twice continuously Fréchet differentiable.

D.3 First order differentiability of the implicit function

To get the Gâteaux differentiability of the implicit function it is still possible to take $\Lambda_0 = \Lambda$ and $E_0 = E$. But to prove the Fréchet differentiability it becomes important to choose smaller spaces Λ_0 and E_0.

Theorem D.8 (First order differentiability).

(i) We assume that the mapping $G : \Lambda \times E \to E$ fulfills the following conditions:

1. *The mapping $\lambda \mapsto G(\lambda, x)$ is continuous from Λ to E for all $x \in E$.*
2. *For all $\lambda, \mu \in \Lambda$ and all $x, y \in E$ there exist the directional derivatives*

$$\partial_1 G(\lambda, x; \mu) = E - \lim_{h \to \infty} \frac{G(\lambda + h\mu, x) - G(\lambda, x)}{h}$$
$$\partial_2 G(\lambda, x; y) = E - \lim_{h \to \infty} \frac{G(\lambda, x + hy) - G(\lambda, x)}{h}$$

of $G : \Lambda \times E \to E$ and
$\partial_1 G : \Lambda \times E \times \Lambda \to E$ and $\partial_2 G : \Lambda \times E \times E \to E$ are continuous.

Then the implicit function $\varphi : \Lambda \to E$ is Gâteaux differentiable such that the mapping $(\lambda, \mu) \mapsto \partial\varphi(\lambda)\mu$ is continuous from $\Lambda \times \Lambda$ to E and

$$\partial\varphi(\lambda)\mu = [I - \partial_2 G(\lambda, \varphi(\lambda))]^{-1} \partial_1 G(\lambda, \varphi(\lambda))\mu$$

for all $\lambda, \mu \in \Lambda$

(ii) We assume now that not only $G : \Lambda \times E \to E$ but also $G : \Lambda_0 \times E_0 \to E_0$ fulfills the conditions of (i), i.e.:

3. *The mapping $\lambda_0 \mapsto G(\lambda_0, x_0)$ is continuous from Λ_0 to E_0 for all $x_0 \in E_0$.*
4. *For all λ_0, $\mu_0 \in \Lambda_0$ and x_0, $y_0 \in E_0$ there exist the directional derivatives*

$$\partial_1 G(\lambda_0, x_0; \mu_0) = E_0 - \lim_{h \to 0} \frac{G(\lambda_0 + h\mu_0, x_0) - G(\lambda_0, x_0)}{h}$$
$$\partial_2 G(\lambda_0, x_0; y_0) = E_0 - \lim_{h \to 0} \frac{G(\lambda_0, x_0 + hy_0) - G(\lambda_0, x_0)}{h}$$

of $G : \Lambda_0 \times E_0 \to E_0$ and
$\partial_1 G : \Lambda_0 \times E_0 \times \Lambda_0 \to E_0$ and $\partial_2 G : \Lambda_0 \times E_0 \times E_0 \to E_0$ are continuous.

In addition we demand

5. *The restricted mappings*

$$\partial_1 G : \Lambda_0 \times E_0 \to L(\Lambda_0, E) \text{ and}$$
$$\partial_2 G : \Lambda_0 \times E_0 \to L(E_0, E)$$

are continuous. (Therefore, by Proposition D.6, it is allowed to write $D_1 G$ instead of $\partial_1 G$ and $D_2 G$ instead of $\partial_2 G$.)

Then $\partial\varphi : \Lambda_0 \to L(\Lambda_0, E)$ is continuous and that means that $\varphi : \Lambda_0 \to E$ is continuously Fréchet differentiable.

Remark D.9. In fact condition 2. and Lemma D.4 especially imply that $\partial_1 G(\lambda, x) \in L(\Lambda, E)$ and $\partial_2 G(\lambda, x) \in L(E)$ for all $\lambda \in \Lambda$ and $x \in E$. This means that the mappings $G(\cdot, x) : \Lambda \to E$, $x \in E$, and $G(\lambda, \cdot) : E \to E$, $\lambda \in \Lambda$, are Gâteaux differentiable. Therefore it is allowed to write $\partial_1 G(\lambda, x)\mu$ and $\partial_2 G(\lambda, x)y$ instead of $\partial_1 G(\lambda, x; \mu)$ and $\partial_2 G(\lambda, x; y)$.
For the same reason it follows from condition 4. of part (ii) that $\partial_1 G : \Lambda_0 \times E_0 \to L(\Lambda_0, E_0)$ and $\partial_2 G : \Lambda_0 \times E_0 \to L(E_0)$. But one has to notice that we only demand the continuity of $\partial_1 G : \Lambda_0 \times E_0 \to L(\Lambda_0, E)$ and $\partial_2 G : \Lambda_0 \times E_0 \to L(E_0, E)$ (see condition 5.) which is weaker.

For the proof of Theorem D.8 we need the following lemma.

Lemma D.10. *Let $(B_n)_{n\in\mathbb{N}}$ be a sequence in $L(E)$ which converges strongly to $B \in L(E)$ (i.e. $B_n x \longrightarrow Bx$ in E as $n \to \infty$ for all $x \in E$). Besides we assume that there exists a constant $\alpha \in [0, 1[$ such that $\|B_n\|_{L(E)}, \|B\|_{L(E)} \leq \alpha$ for all $n \in \mathbb{N}$. Then $(I - B_n)^{-1}, (I - B)^{-1} \in L(E)$ for all $n \in \mathbb{N}$ and*

$$(I - B_n)^{-1} \longrightarrow (I - B)^{-1} \text{ strongly as } n \to \infty.$$

One even gets that

$$(I - B_n)^{-1} x_n \longrightarrow (I - B)^{-1} x \text{ as } n \to \infty$$

for each sequence $(x_n)_{n\in\mathbb{N}}$ in E with $x_n \longrightarrow x$ in E as $n \to \infty$.

Proof. The existence of $(I - B)^{-1} \in L(E)$ and $(I - B_n)^{-1} \in L(E)$, $n \in \mathbb{N}$, follows from [Al 92, 3.6 Neumann-Reihe, p.104]
Hence let $(x_n)_{n\in\mathbb{N}}$ be a sequence in E with $x_n \longrightarrow x$ in E as $n \to \infty$. If we define $y_n := (I - B_n)^{-1} x_n$ and $y := (I - B)^{-1} x$ we obtain that $x_n = y_n - B_n y_n$ and $x = y - By$ such that

$$y_n - y = B_n(y_n - y) + (x_n - x) + (B_n y - By)$$

This implies for the norm of the difference that

$$\|y_n - y\| \leq \alpha\|y_n - y\| + \|x_n - x\| + \|B_n y - By\|$$

and since $\alpha < 1$ we can conclude that

$$\|y_n - y\| \leq \frac{1}{1-\alpha}(\|x_n - x\| + \|B_n y - By\|)$$

So finally the assertion is proved as $\|x_n - x\| \longrightarrow 0$ and $\|B_n y - By\| \longrightarrow 0$ as $n \to \infty$. □

Proof. Proof of (i): We fix λ and $\mu \in \Lambda$. Because of the continuity of the directional derivatives $\partial_1 G$ and $\partial_2 G$ we obtain by the fundamental theorem for Bochner integrals, Theorem A.7, that for all $h \in \mathbb{R}$

$$\begin{aligned}
&\varphi(\lambda + h\mu) - \varphi(\lambda) \\
&= G(\lambda + h\mu, \varphi(\lambda + h\mu)) - G(\lambda, \varphi(\lambda)) \\
&= \big[G(\lambda + h\mu, \varphi(\lambda + h\mu)) - G(\lambda, \varphi(\lambda + h\mu))\big] \\
&\quad + \big[G(\lambda, \varphi(\lambda + h\mu)) - G(\lambda, \varphi(\lambda))\big] \\
&= \int_0^1 \partial_1 G(\lambda + sh\mu, \varphi(\lambda + h\mu))h\mu \, ds \\
&\quad + \int_0^1 \partial_2 G(\lambda, \varphi(\lambda) + s(\varphi(\lambda + h\mu) - \varphi(\lambda)))(\varphi(\lambda + h\mu) - \varphi(\lambda)) \, ds
\end{aligned}$$

For general $x, y \in E$ we define the strong integral

$$\int_0^1 \partial_2 G(\lambda, x + sy) \, ds : E \to E$$

by

$$\int_0^1 \partial_2 G(\lambda, x + sy) \, ds \, z := \int_0^1 \partial_2 G(\lambda, x + sy) z \, ds$$

Then $\displaystyle\int_0^1 \partial_2 G(\lambda, x + sy) \, ds$ describes a linear operator on E.
Furthermore we obtain by the contraction property that

$$\|\int_0^1 \partial_2 G(\lambda, x + sy) z \, ds\| \leq \int_0^1 \|\partial_2 G(\lambda, x + sy) z\| \, ds \leq \alpha\|z\|$$

for all $x, y, z \in E$. Hence for each $h \in \mathbb{R}$ there exists

$$[I - \int_0^1 \partial_2 G(\lambda, \varphi(\lambda) + s(\varphi(\lambda + h\mu) - \varphi(\lambda))) \, ds]^{-1} \in L(E)$$

by Lemma D.10 and we get that

$$\frac{\varphi(\lambda+h\mu)-\varphi(\lambda)}{h} = [I - \int_0^1 \partial_2 G(\lambda, \varphi(\lambda) + s(\varphi(\lambda+h\mu)-\varphi(\lambda)))\, ds]^{-1} \int_0^1 \partial_1 G(\lambda + sh\mu, \varphi(\lambda+h\mu))\mu\, ds$$

So finally by Lemma D.10 we only have to show that for all $x \in E$

$$\int_0^1 \partial_2 G(\lambda, \varphi(\lambda) + s(\varphi(\lambda+h\mu)-\varphi(\lambda)))x\, ds \underset{h\to 0}{\longrightarrow} \int_0^1 \partial_2 G(\lambda, \varphi(\lambda))x\, ds$$

and

$$\int_0^1 \partial_1 G(\lambda + sh\mu, \varphi(\lambda+h\mu))\mu\, ds \underset{h\to 0}{\longrightarrow} \int_0^1 \partial_1 G(\lambda, \varphi(\lambda))\mu\, ds$$

But this assertion follows from Lebesgue's dominated convergence theorem: By Theorem D.1 the mapping $\varphi : \Lambda \to E$ is continuous. As the mappings $\partial_1 G$ and $\partial_2 G$ are also continuous by assumption we obtain

1. For all $s \in [0,1]$ and $x \in E$

$$\partial_1 G(\lambda + sh\mu, \varphi(\lambda+h\mu))\mu \underset{h\to 0}{\longrightarrow} \partial_1 G(\lambda, \varphi(\lambda))\mu$$

and

$$\partial_2 G(\lambda, \varphi(\lambda) + s(\varphi(\lambda+h\mu)-\varphi(\lambda)))x \underset{h\to 0}{\longrightarrow} \partial_2 G(\lambda, \varphi(\lambda))x$$

2. There exists a constant c such that

$$\|\partial_1 G(\lambda + sh\mu, \varphi(\lambda+h\mu))\mu\| \le c$$

and

$$\|\partial_2 G(\lambda, \varphi(\lambda) + s(\varphi(\lambda+h\mu)-\varphi(\lambda)))x\| \le c$$

for all $s, h \in [0,1]$.

So far we know that there exist the directional derivatives of φ and that they are described by the given formula. This representation already provides that $\partial\varphi(\lambda;\cdot) \in L(\Lambda, E)$ and therefore the Gâteaux differentiability of φ. The continuity of the mapping

$$(\lambda, \mu) \mapsto \partial\varphi(\lambda;\mu) = [I - \partial_2 G(\lambda, \varphi(\lambda))]^{-1} \partial_1 G(\lambda, \varphi(\lambda))\mu$$

is again a consequence of Lemma D.10.
Proof of (ii): We consider $G : \Lambda_0 \times E_0 \to E_0$. Then we get by part (i) that $\varphi : \Lambda_0 \to E_0$ is Gâteaux differentiable and for $\lambda_0,\ \mu_0 \in \Lambda_0$ the derivative is given by

$$\partial\varphi(\lambda_0)\mu_0 = [I - \partial_2 G(\lambda_0, \varphi(\lambda_0))]^{-1}\partial_1 G(\lambda_0, \varphi(\lambda_0))\mu_0$$

where $\partial_2 G(\lambda_0, \varphi(\lambda_0)) \in L(E_0)$ and $\partial_1 G(\lambda_0, \varphi(\lambda_0)) \in L(\Lambda_0, E_0) \subset L(\Lambda_0, E)$. But by condition 1. $\partial_2 G(\lambda_0, \varphi(\lambda_0))$ can also be considered as an element of $L(E)$ with $\|\partial_2 G(\lambda_0, \varphi(\lambda_0))\|_{L(E)} \leq \alpha < 1$ for all $\lambda_0 \in \Lambda_0$.
Now let $(\lambda_{0,n})_{n\in\mathbb{N}}$ be a sequence in Λ_0 with $\lambda_{0,n} \longrightarrow \lambda_0$ in Λ_0 as $n \to \infty$. Besides we fix $\mu_0 \in \Lambda_0$ and define

$$\begin{aligned} y_n &:= \partial\varphi(\lambda_{0,n})\mu_0 \in E_0 \\ y &:= \partial\varphi(\lambda_0)\mu_0 \in E_0 \end{aligned}$$

Applying the formula for the derivative of φ we obtain that

$$\begin{aligned} y_n &= \partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))y_n + \partial_1 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))\mu_0 \\ y &= \partial_2 G(\lambda_0, \varphi(\lambda_0))y + \partial_1 G(\lambda_0, \varphi(\lambda_0))\mu_0 \end{aligned}$$

In this way we get for the norm of the difference

$$\begin{aligned} &\|\partial\varphi(\lambda_{0,n})\mu_0 - \partial\varphi(\lambda_0)\mu_0\| \\ &= \|y_n - y\| \\ &\leq \|\partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))y_n - \partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))y\| \\ &\quad + \|\partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))y - \partial_2 G(\lambda_0, \varphi(\lambda_0))y\| \\ &\quad + \|\partial_1 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))\mu_0 - \partial_1 G(\lambda_0, \varphi(\lambda_0))\mu_0\| \\ &\leq \|\partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n}))\|_{L(E)}\|y_n - y\| \\ &\quad + \|\partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n})) - \partial_2 G(\lambda_0, \varphi(\lambda_0))\|_{L(E_0,E)}\|y\|_0 \\ &\quad + \|\partial_1 G(\lambda_{0,n}, \varphi(\lambda_{0,n})) - \partial_1 G(\lambda_0, \varphi(\lambda_0))\|_{L(\Lambda_0,E)}\|\mu_0\|_{\Lambda_0} \\ &\leq \alpha\|y_n - y\| \\ &\quad + \|\partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n})) - \partial_2 G(\lambda_0, \varphi(\lambda_0))\|_{L(E_0,E)}\|\partial\varphi(\lambda_0)\|_{L(\Lambda_0,E_0)}\|\mu_0\|_{\Lambda_0} \\ &\quad + \|\partial_1 G(\lambda_{0,n}, \varphi(\lambda_{0,n})) - \partial_1 G(\lambda_0, \varphi(\lambda_0))\|_{L(\Lambda_0,E)}\|\mu_0\|_{\Lambda_0} \end{aligned}$$

Since $\alpha < 1$ and $\varphi(\lambda_{0,n}) \longrightarrow \varphi(\lambda_0)$ in E_0 as $n \to \infty$ we can finally conclude that

$$\begin{aligned} &\|\partial\varphi(\lambda_{0,n})\mu_0 - \partial\varphi(\lambda_0)\mu_0\| \\ &\leq \frac{1}{1-\alpha}\big[\|\partial_2 G(\lambda_{0,n}, \varphi(\lambda_{0,n})) - \partial_2 G(\lambda_0, \varphi(\lambda_0))\|_{L(E_0,E)}\|\partial\varphi(\lambda_0)\|_{L(\Lambda_0,E_0)} \\ &\qquad + \|\partial_1 G(\lambda_{0,n}, \varphi(\lambda_{0,n})) - \partial_1 G(\lambda_0, \varphi(\lambda_0))\|_{L(\Lambda_0,E)}\big]\|\mu_0\|_{\Lambda_0} \end{aligned}$$

$= c_n \|\mu_0\|_{\Lambda_0}$ where $c_n \longrightarrow 0$ as $n \to \infty$.

by condition 5. Hence the Fréchet differentiability follows from Proposition D.6. □

Corollary D.11. *If the assumptions of Theorem D.8 (i) are fulfilled and if we have in addition that* $\|\partial_1 G(\lambda, x)\|_{L(\Lambda,E)} \leq C$ *for all* $\lambda \in \Lambda$ *and* $x \in E$ *then the mapping* $\partial\varphi : \Lambda \to L(\Lambda, E)$ *is also bounded.*

Proof. Let $\lambda, \mu \in \Lambda$ and $x \in E$. Since $G : \Lambda \times E \to E$ fulfills the assumptions of Theorem D.8 (i) we get for $\varphi : \Lambda \to E$ that

$$\partial\varphi(\lambda)\mu = [I - \partial_2 G(\lambda, \varphi(\lambda))]^{-1} \partial_1 G(\lambda, \varphi(\lambda))\mu$$

This implies that

$$\partial\varphi(\lambda)\mu = \partial_2 G(\lambda, \varphi(\lambda))\partial\varphi(\lambda)\mu + \partial_1 G(\lambda, \varphi(\lambda))\mu$$

where $\partial_2 G(\lambda, \varphi(\lambda)) \in L(E)$ with $\|\partial_2 G(\lambda, \varphi(\lambda))\|_{L(E)} \leq \alpha < 1$. Therefore

$$\|\partial\varphi(\lambda)\mu\| \leq \alpha\|\partial\varphi(\lambda)\mu\| + C\|\mu\|_\Lambda$$

Since $\alpha < 1$ it follows that

$$\|\partial\varphi(\lambda)\mu\| \leq \frac{C}{1-\alpha}\|\mu\|_\Lambda$$

□

Corollary D.12. *We assume that the assumptions of Theorem D.8 (ii) are fulfilled. If we have in addition that the mappings*

$$\partial_1 G : \Lambda_0 \times E_0 \to L(\Lambda_0, E)$$
$$\partial_2 G : \Lambda_0 \times E_0 \to L(E_0, E)$$

are uniformly continuous and that $\partial\varphi : \Lambda_0 \to L(\Lambda_0, E_0)$ *is bounded then the Fréchet derivative* $D\varphi : \Lambda_0 \to L(\Lambda_0, E)$ *is even uniformly continuous.*

Proof. As $\partial\varphi : \Lambda_0 \to L(\Lambda_0, E_0)$ is bounded we have that $\varphi : \Lambda_0 \to E_0$ is Lipschitz continuous by the fundamental Theorem A.7. Then the result easily follows from the proof of Theorem D.8 (ii). □

D.4 Second order differentiability of the implicit function

Theorem D.13 (Second order differentiability).

(i) We assume that

1. *Both mappings $G : \Lambda \times E \to E$ and $G : \Lambda_0 \times E_0 \to E_0$ fulfill the conditions of Theorem D.8 (i).*
2. *There exist the directional derivatives of second order from $G : \Lambda_0 \times E_0 \to E$ which means that for all λ_0, μ_0, $\nu_0 \in \Lambda_0$ and x_0, y_0, $z_0 \in E_0$ the following limits exist in E.*

$$\partial_1^2 G(\lambda_0, x_0; \mu_0, \nu_0) := \lim_{h\to 0} \frac{\partial_1 G(\lambda_0 + h\nu_0, x_0)\mu_0 - \partial_1 G(\lambda_0, x_0)\mu_0}{h}$$

$$\partial_1\partial_2 G(\lambda_0, x_0; y_0, \mu_0) := \lim_{h\to 0} \frac{\partial_2 G(\lambda_0 + h\mu_0, x_0)y_0 - \partial_2 G(\lambda_0, x_0)y_0}{h}$$

$$\partial_2\partial_1 G(\lambda_0, x_0; \mu_0, y_0) := \lim_{h\to 0} \frac{\partial_1 G(\lambda_0, x_0 + hy_0)\mu_0 - \partial_1 G(\lambda_0, x_0)\mu_0}{h}$$

$$\partial_2^2 G(\lambda_0, x_0; y_0, z_0) := \lim_{h\to 0} \frac{\partial_2 G(\lambda_0, x_0 + hz_0)y_0 - \partial_2 G(\lambda_0, x_0)y_0}{h}$$

Then the mapping $\varphi : \Lambda_0 \to E$ is twice differentiable in all points $\lambda_0 \in \Lambda_0$ and all directions $\mu_0, \nu_0 \in \Lambda_0$ such that the mapping

$$\partial^2\varphi : \Lambda_0 \times \Lambda_0 \times \Lambda_0 \to E$$

$$(\lambda_0, \mu_0, \nu_0) \mapsto \partial^2\varphi(\lambda_0; \mu_0, \nu_0) = \lim_{h\to 0} \frac{\partial\varphi(\lambda_0 + h\nu_0)\mu_0 - \varphi(\lambda_0)\mu_0}{h}$$

is continuous. There even is an explicit formula describing the directional derivatives:

$$\begin{aligned}\partial^2\varphi(\lambda_0; \mu_0, \nu_0) = [I - \partial_2 G(\lambda_0, \varphi(\lambda_0))]^{-1} \\ \{\partial_1^2 G(\lambda_0, \varphi(\lambda_0); \mu_0, \nu_0) \\ + \partial_1\partial_2 G(\lambda_0, \varphi(\lambda_0); \partial\varphi(\lambda_0)\mu_0, \nu_0) \\ + \partial_2\partial_1 G(\lambda_0, \varphi(\lambda_0); \mu_0, \partial\varphi(\lambda_0)\nu_0) \\ + \partial_2^2 G(\lambda_0, \varphi(\lambda_0); \partial\varphi(\lambda_0)\mu_0, \partial\varphi(\lambda_0)\nu_0)\}\end{aligned}$$

(ii) Let $(\Lambda_1, \| \|_{\Lambda_1})$ be a further Banach space which is continuously embedded in $(\Lambda_0, \| \|_{\Lambda_0})$.
In addition to the hypotheses of part (i) we assume that the following conditions hold:

3. The mapping $G : \Lambda_0 \times E_0 \to E$ is even twice Fréchet differentiable in each variable such that the derivatives

$$\begin{aligned} D_1^2 G &: \Lambda_0 \times E_0 \to L(\Lambda_0, L(\Lambda_0, E)) \\ D_1 D_2 G &: \Lambda_0 \times E_0 \to L(\Lambda_0, L(E_0, E)) \\ D_2 D_1 G &: \Lambda_0 \times E_0 \to L(E_0, L(\Lambda_0, E)) \\ D_2^2 G &: \Lambda_0 \times E_0 \to L(E_0, L(E_0, E)) \end{aligned}$$

are continuous.

4. The mapping $\varphi : \Lambda_1 \to E_0$ is Fréchet differentiable with continuous derivative $D\varphi : \Lambda_1 \to L(\Lambda_1, E_0) \subset L(\Lambda_1, E)$.

Then the Fréchet derivative $D\varphi : \Lambda_1 \to L(\Lambda_1, E)$ is once Gâteaux differentiable.

(iii) If it is even possible to verify that

5. $\partial D\varphi : \Lambda_1 \to L(\Lambda_1, L(\Lambda_1, E_0)) \subset L(\Lambda_1, L(\Lambda_1, E))$

then $\varphi : \Lambda_1 \to E$ is twice Fréchet differentiable.

Proof. (i): [Za 83, Theorem 10.4, p.208]
(ii): Let $\lambda_1, \mu_1, \nu_1 \in \Lambda_1$. At first we make use of condition 1. which provides that φ is Gâteaux differentiable as a mapping from Λ to E (see Theorem D.8). Using the formula which describes the derivative we get that for all $h \in \mathbb{R}$

$$\begin{aligned} & D\varphi(\lambda_1 + h\nu_1)\mu_1 - D\varphi(\lambda_1)\mu_1 \\ &= \partial\varphi(\lambda_1 + h\nu_1)\mu_1 - \partial\varphi(\lambda_1)\mu_1 \\ &= \partial_2 G(\lambda_1, \varphi(\lambda_1))(\partial\varphi(\lambda_1 + h\nu_1)\mu_1 - \partial\varphi(\lambda_1)\mu_1) \\ &\quad + [\partial_1 G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - \partial_1 G(\lambda_1, \varphi(\lambda_1))]\mu_1 \\ &\quad + [\partial_2 G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - \partial_2 G(\lambda_1, \varphi(\lambda_1))]\partial\varphi(\lambda_1 + h\nu_1)\mu_1 \end{aligned}$$

and therefore

$$\begin{aligned} & \frac{1}{h}(D\varphi(\lambda_1 + h\nu_1)\mu_1 - D\varphi(\lambda_1)\mu_1) \\ &= [I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1}\{\frac{1}{h}J_1(h)\mu_1 + \frac{1}{h}J_2(h)\mu_1\} \end{aligned}$$

where by condition 3.

$$J_1(h) := \partial_1 G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - \partial_1 G(\lambda_1, \varphi(\lambda_1))$$

$$= D_1G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_1G(\lambda_1, \varphi(\lambda_1))$$
$$\in L(\Lambda_0, E) \subset L(\Lambda_1, E)$$

and by condition 3. in addition to 4.

$$J_2(h) := [\partial_2 G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - \partial_2 G(\lambda_1, \varphi(\lambda_1))]\partial\varphi(\lambda_1 + h\nu_1)$$
$$= \underbrace{[D_2G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_2G(\lambda_1, \varphi(\lambda_1))]}_{\in L(E_0,E)} \underbrace{D\varphi(\lambda_1 + h\nu_1)}_{\in L(\Lambda_1,E_0)}$$
$$\in L(\Lambda_1, E)$$

Claim.

1. $$\frac{1}{h}J_1(h) \underset{h\to 0}{\longrightarrow} D_1^2G(\lambda_1, \varphi(\lambda_1))(\cdot)\nu_1 + D_2D_1G(\lambda_1, \varphi(\lambda_1))(\cdot)D\varphi(\lambda_1)\nu_1 \quad \text{in } L(\Lambda_1, E)$$
2. $$\frac{1}{h}J_2(h) \underset{h\to 0}{\longrightarrow} D_1D_2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)(\cdot))\nu_1 + D_2^2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)(\cdot))D\varphi(\lambda_1)\nu_1 \quad \text{in } L(\Lambda_1, E)$$

1. We prove the first part of the claim:

$$\|\frac{1}{h}J_1(h)\mu_1 - D_1^2G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1 - D_2D_1G(\lambda_1, \varphi(\lambda_1))(\mu_1)D\varphi(\lambda_1)\nu_1\|$$
$$\leq \|\frac{[D_1G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_1G(\lambda_1, \varphi(\lambda_1 + h\nu_1))]\mu_1}{h} - D_1^2G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1\|$$
$$+ \|\frac{[D_1G(\lambda_1, \varphi(\lambda_1 + h\nu_1)) - D_1G(\lambda_1, \varphi(\lambda_1))]\mu_1}{h} - D_2D_1G(\lambda_1, \varphi(\lambda_1))(\mu_1)D\varphi(\lambda_1)\nu_1\|$$

(i) Using the continuity of the Fréchet derivative $D_1^2G : \Lambda_0 \times E_0 \to L(\Lambda_0, L(\Lambda_0, E))$ we obtain by the fundamental theorem for Bochner integrals that the first summand can be estimated in the following way:

$$\|\frac{[D_1G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_1G(\lambda_1, \varphi(\lambda_1 + h\nu_1))]\mu_1}{h} - D_1^2G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1\|$$
$$= \|\int_0^1 D_1^2G(\lambda_1 + \sigma h\nu_1, \varphi(\lambda_1 + h\nu_1))(\mu_1)\nu_1 - D_1^2G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1 \, d\sigma\|$$
$$\leq \int_0^1 \|D_1^2G(\lambda_1 + \sigma h\nu_1, \varphi(\lambda_1 + h\nu_1))(\mu_1)\nu_1 - D_1^2G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1\| \, d\sigma$$

$$\leq \int_0^1 \|D_1^2 G(\lambda_1 + \sigma h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_1^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(\Lambda_0, L(\Lambda_0, E))}\, d\sigma$$
$$\|\mu_1\|_{\Lambda_0} \|\nu_1\|_{\Lambda_0}$$
$$\leq c_h \|\mu_1\|_{\Lambda_0} \leq c_h c \|\mu_1\|_{\Lambda_1}$$

where the last estimate follows from the fact that Λ_1 is continuously embedded in Λ_0. Besides $c_h \longrightarrow 0$ as $h \to 0$ because of Lebesgue's dominated convergence theorem :
Since $D_1^2 G : \Lambda_0 \times E_0 \to L(\Lambda_0, L(\Lambda_0, E))$ and $\varphi : \Lambda_0 \to E_0$ are continuous we get that for all $\sigma \in [0,1]$

$$\|D_1^2 G(\lambda_1 + \sigma h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_1^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(\Lambda_0, L(\Lambda_0, E))} \longrightarrow 0$$

as $h \to 0$ and of course there exists a constant C such that

$$\|D_1^2 G(\lambda_1 + \sigma h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_1^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(\Lambda_0, L(\Lambda_0, E))} \leq \tilde{c}$$

for all $(\sigma, h) \in [0,1] \times [0,1]$
(ii) With respect to the second summand we obtain in a similar way that

$$\| \frac{[D_1 G(\lambda_1, \varphi(\lambda_1 + h\nu_1)) - D_1 G(\lambda_1, \varphi(\lambda_1))]\mu_1}{h}$$
$$- D_2 D_1 G(\lambda_1, \varphi(\lambda_1))(\mu_1) D\varphi(\lambda_1)\nu_1 \|$$
$$\leq \int_0^1 \|D_2 D_1 G(\lambda_1, \varphi(\lambda_1) + \sigma(\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)))(\mu_1)(\frac{\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)}{h})$$
$$- D_2 D_1 G(\lambda_1, \varphi(\lambda_1))(\mu_1) D\varphi(\lambda_1)\nu_1 \|\, d\sigma$$
$$\leq \int_0^1 \|D_2 D_1 G(\lambda_1, \varphi(\lambda_1) + \sigma(\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)))\|_{L(E_0, L(\Lambda_0, E))}\, d\sigma$$
$$\|\mu_1\|_{\Lambda_0} \| \frac{\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)}{h} - D\varphi(\lambda_1)\nu_1 \|_0$$
$$+ \int_0^1 \|D_2 D_1 G(\lambda_1, \varphi(\lambda_1) + \sigma(\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)))$$
$$- D_2 D_1 G(\lambda_1, \varphi(\lambda_1))\|_{L(E_0, L(\Lambda_0, E))}\, d\sigma$$
$$\|\mu_1\|_{\Lambda_0} \|D\varphi(\lambda_1)\nu_1\|_0$$
$$= c_h \|\mu_1\|_{\Lambda_0} \leq c_h c \|\mu_1\|_{\Lambda_1}$$

where $c_h \longrightarrow 0$ as $h \to 0$:
Since $D_2 D_1 G : \Lambda_0 \times E_0 \to L(E_0, L(\Lambda_0, E_0))$ and $\varphi : \Lambda_0 \to E_0$ are continuous we can use Lebesgue's dominated convergence theorem like above to get that

$$
\begin{aligned}
&|\int_0^1 \|D_2D_1G(\lambda_1, \varphi(\lambda_1) + \sigma(\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)))\|_{L(E_0, L(\Lambda_0, E))}\, d\sigma \\
&\quad - \|D_2D_1G(\lambda_1, \varphi(\lambda_1)\|_{L(E_0, L(\Lambda_0, E))}| \\
&\leq \int_0^1 \|D_2D_1G(\lambda_1, \varphi(\lambda_1) + \sigma(\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1))) \\
&\quad - D_2D_1G(\lambda_1, \varphi(\lambda_1))\|_{L(E_0, L(\Lambda_0, E))}\, d\sigma \underset{h\to 0}{\longrightarrow} 0.
\end{aligned}
$$

Besides assumption 4. provides that

$$
\|\frac{\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)}{h} - D\varphi(\lambda_1)\nu_1\|_0 \underset{h\to 0}{\longrightarrow} 0.
$$

In this way we proved that

$$
\begin{aligned}
\frac{1}{h}J_1(h) \underset{h\to 0}{\longrightarrow} & D_1^2G(\lambda_1, \varphi(\lambda_1))(\cdot)\nu_1 \\
& + D_2D_1G(\lambda_1, \varphi(\lambda_1))(\cdot)D\varphi(\lambda_1)\nu_1 \quad \text{in } L(\Lambda_1, E)
\end{aligned}
$$

In fact it is even possible to get this convergence in $L(\Lambda_0, E)$ but we will see that it is really necessary to introduce the smaller space Λ_1 to prove the second part of the claim.

2. We prove the second part of the claim:
At first we split the term in the following way

$$
\begin{aligned}
&\|\frac{1}{h}J_2(h)\mu_1 \\
&- D_1D_2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1 - D_2^2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1\| \\
\leq\ &\|\frac{D_2G(\lambda_1 + h\nu_1, \varphi(\lambda_1 + h\nu_1)) - D_2G(\lambda_1, \varphi(\lambda_1 + h\nu_1))}{h}D\varphi(\lambda_1 + h\nu_1)\mu_1 \\
&- D_1D_2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1\| \\
&+ \|\frac{D_2G(\lambda_1, \varphi(\lambda_1 + h\nu_1)) - D_2G(\lambda_1, \varphi(\lambda_1))}{h}D\varphi(\lambda_1 + h\nu_1)\mu_1 \\
&\quad - D_2^2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)D\varphi(\lambda_1)\nu_1\|
\end{aligned}
$$

(i) By assumption the Fréchet derivative $D_1D_2G : \Lambda_0 \times E_0 \to L(\Lambda_0, L(E_0, E))$ is continuous. Therefore we can use the fundamental theorem for Bochner integrals to obtain for the first summand that

$$
\begin{aligned}
&\|\frac{D_2G(\lambda_1+h\nu_1,\varphi(\lambda_1+h\nu_1))-D_2G(\lambda_1,\varphi(\lambda_1+h\nu_1))}{h}D\varphi(\lambda_1+h\nu_1)\mu_1\\
&-D_1D_2G(\lambda_1,\varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1\|\\
=&\|\int_0^1 D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))(D\varphi(\lambda_1+h\nu_1)\mu_1)\nu_1\\
&-D_1D_2G(\lambda_1,\varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1\; d\sigma\|\\
\le&\int_0^1 \|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))(D\varphi(\lambda_1+h\nu_1)\mu_1)\nu_1\\
&\quad -D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))(D\varphi(\lambda_1)\mu_1)\nu_1\|\; d\sigma\\
+&\int_0^1 \|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))(D\varphi(\lambda_1)\mu_1)\nu_1\\
&\quad -D_1D_2G(\lambda_1,\varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1\|\; d\sigma\\
\le&\int_0^1 \|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))\|_{L(\Lambda_0,L(E_0,E))}\; d\sigma\\
&\|D\varphi(\lambda_1+h\nu_1)-D\varphi(\lambda_1)\|\|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_0}\\
&+\int_0^1 \|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))-D_1D_2G(\lambda_1,\varphi(\lambda_1))\|\; d\sigma\\
&\quad \|D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_0}\le c_h\|\mu_1\|_{\Lambda_1}
\end{aligned}
$$

where $c_h \longrightarrow 0$ as $h \to 0$ for the following reasons :
Because of the continuity of $D_1D_2G : \Lambda_0 \times E_0 \to L(\Lambda_0, L(E_0,E))$ and $\varphi : \Lambda_0 \to E_0$ it is clear that for all $\sigma \in [0,1]$

$$\|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))-D_1D_2G(\lambda_1,\varphi(\lambda_1))\|_{L(\Lambda_0,L(E_0,E))}\underset{h\to 0}{\longrightarrow} 0.$$

Besides there exists a constant c such that

$$\|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))-D_1D_2G(\lambda_1,\varphi(\lambda_1))\|_{L(\Lambda_0,L(E_0,E))}\le c$$

for all $(\sigma,h) \in [0,1]\times[0,1]$.Therefore it is clear by Lebesgue's dominated convergence theorem that

$$
\begin{aligned}
&|\int_0^1 \|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))\|_{L(\Lambda_0,L(E_0,E))}\; d\sigma\\
-&\int_0^1 \|D_1D_2G(\lambda_1,\varphi(\lambda_1))\|\; d\sigma|\\
\le&\int_0^1 \|D_1D_2G(\lambda_1+\sigma h\nu_1,\varphi(\lambda_1+h\nu_1))-D_1D_2G(\lambda_1,\varphi(\lambda_1))\|_{L(\Lambda_0,L(E_0,E))}\; d\sigma\\
&\longrightarrow\; 0 \text{ as } h\to 0.
\end{aligned}
$$

Besides we obtain by assumption 4. that

$$\|D\varphi(\lambda_1 + h\nu_1) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)} \longrightarrow 0 \text{ as } h \to 0$$

(ii) For the second summand we get in a similar way that

$$\begin{aligned}
&\|\frac{D_2G(\lambda_1,\varphi(\lambda_1+h\nu_1)) - D_2G(\lambda_1,\varphi(\lambda_1))}{h} D\varphi(\lambda_1+h\nu_1)\mu_1 \\
&\quad - D_2^2G(\lambda_1,\varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)D\varphi(\lambda_1)\nu_1\| \\
&\le \int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1)+\sigma(\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1)))(D\varphi(\lambda_1+h\nu_1)\mu_1) \\
&\qquad (\frac{\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1)}{h}) \\
&\qquad - D_2^2G(\lambda_1,\varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)D\varphi(\lambda_1)\nu_1\| \, d\sigma \\
&\le \int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1)+\sigma(\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1)))\|_{L(E_0,L(E_0,E))} \, d\sigma \\
&\quad \|D\varphi(\lambda_1+h\nu_1)\|_{L(\Lambda_1,E_0)}\|\mu_1\|_{\Lambda_1}\|\frac{\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1)}{h} - D\varphi(\lambda_1)\nu_1\|_0 \\
&\quad + \int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1)+\sigma(\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1)))\|_{L(E_0,L(E_0,E))} \, d\sigma \\
&\quad \|D\varphi(\lambda_1+h\nu_1) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\mu_1\|_{\Lambda_1}\|D\varphi(\lambda_1)\nu_1\|_0 \\
&\quad + \int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1)+\sigma(\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1))) \\
&\qquad - D_2^2G(\lambda_1,\varphi(\lambda_1))\|_{L(E_0,L(E_0,E))} \, d\sigma \\
&\quad \|D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\mu_1\|_{\Lambda_1}\|D\varphi(\lambda_1)\nu_1\|_0 \\
&\le c_h\|\mu_1\|_{\Lambda_1}
\end{aligned}$$

where $c_h \longrightarrow 0$ as $h \to 0$:
Using the continuity of the derivative $D_2^2G : \Lambda_0 \times E_0 \to L(E_0, L(E_0,E))$ and the continuity of $\varphi : \Lambda_0 \to E_0$ we get that

$$\begin{aligned}
&|\int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1)+\sigma(\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1)))\|_{L(E_0,L(E_0,E))} \, d\sigma \\
&\quad - \int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1))\| \, d\sigma| \\
&\le \int_0^1 \|D_2^2G(\lambda_1,\varphi(\lambda_1)+\sigma(\varphi(\lambda_1+h\nu_1)-\varphi(\lambda_1))) \\
&\qquad - D_2^2G(\lambda_1,\varphi(\lambda_1))\|_{L(E_0,L(E_0,E))} \, d\sigma \\
&\quad \longrightarrow 0 \text{ as } h \to 0
\end{aligned}$$

In addition assumption 4. provides that

$$
\begin{aligned}
&|\|D\varphi(\lambda_1 + h\nu_1)\|_{L(\Lambda_1,E_0)} - \|D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}| \\
\leq &\|D\varphi(\lambda_1 + h\nu_1) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)} \longrightarrow 0 \text{ as } h \to 0
\end{aligned}
$$

and

$$
\|\frac{\varphi(\lambda_1 + h\nu_1) - \varphi(\lambda_1)}{h} - D\varphi(\lambda_1)\nu_1\|_0 \longrightarrow 0 \text{ as } h \to 0.
$$

Now we define

$$
\begin{aligned}
\tilde{J}_1 :=& D_1^2 G(\lambda_1, \varphi(\lambda_1))(\cdot)\nu_1 \\
&+ D_2 D_1 G(\lambda_1, \varphi(\lambda_1))(\cdot) D\varphi(\lambda_1)\nu_1 \in L(\Lambda_1, E)
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{J}_2 :=& D_1 D_2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)(\cdot))\nu_1 \\
&+ D_2^2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)(\cdot)) D\varphi(\lambda_1)\nu_1
\end{aligned}
$$

By assumption 1. we have that $\partial_2 G(\lambda, x) \in L(E)$ for all $\lambda \in \Lambda$ and $x \in E$. Thus we finally get that

$$
\begin{aligned}
&\|\frac{1}{h}\big(D\varphi(\lambda_1 + h\nu_1)\mu_1 - D\varphi(\lambda_1)\mu_1\big) - [I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1}\{\tilde{J}_1\mu_1 + \tilde{J}_2\mu_1\}\| \\
= &\|[I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1}\{(\frac{1}{h}J_1(h) - \tilde{J}_1)\mu_1 + (\frac{1}{h}J_2(h) - \tilde{J}_2)\mu_1\}\| \\
\leq &\|[I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1}\|_{L(E)} \\
&(\|[\frac{1}{h}J_1(h) - \tilde{J}_1]\mu_1\| + \|[\frac{1}{h}J_2(h) - \tilde{J}_2]\mu_1\|) \\
\leq &\|[I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1}\|_{L(E)} \\
&(\|\frac{1}{h}J_1(h) - \tilde{J}_1\|_{L(\Lambda_1,E)} + \|\frac{1}{h}J_2(h) - \tilde{J}_2\|_{L(\Lambda_1,E)})\|\mu_1\|_{\Lambda_1} \\
\leq &c_h\|\mu_1\|_{\Lambda_1} \text{ with } c_h \longrightarrow 0 \text{ as } h \to 0
\end{aligned}
$$

Thus it is verified that there exist the directional derivatives of $D\varphi : \Lambda_1 \to L(\Lambda_1, E)$ and we have the following representation

$$
\begin{aligned}
(\partial D\varphi(\lambda_1; \nu_1))\mu_1 = &[I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1} \\
&\{D_1^2 G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1 \\
&\quad + D_1 D_2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1 \\
&\quad + D_2 D_1 G(\lambda_1, \varphi(\lambda_1))(\mu_1, D\varphi(\lambda_1))\nu_1 \\
&\quad + D_2^2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1) D\varphi(\lambda_1)\nu_1\}
\end{aligned}
$$

Therefore it is easy to see that the mapping $\partial D\varphi(\lambda_1; \cdot) : \Lambda_1 \to L(\Lambda_1, E_1)$ is linear and continuous. That means that $D\varphi : \Lambda_1 \to L(\Lambda_1, E)$ is Gâteaux

differentiable and we write $\partial D\varphi(\lambda_1)(\mu_1)\nu_1$ instead of $(\partial D\varphi(\lambda_1;\nu_1))\mu_1$.
Proof of (iii): By Proposition D.6 we only have to show that $\partial D\varphi : \Lambda_1 \to L(\Lambda_1, L(\Lambda_1, E))$ is continuous.
Hence we fix λ_1, μ_1, $\nu_1 \in \Lambda_1$ and a sequence $(\lambda_{1,n})_{n\in\mathbb{N}}$ such that $\lambda_{1,n} \longrightarrow \lambda_1$ as $n \to \infty$. Then we set

$$\begin{aligned} y_n &:= \partial D\varphi(\lambda_{1,n})(\mu_1)\nu_1 \\ &= \partial_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\partial D\varphi(\lambda_{1,n})(\mu_1)\nu_1 \\ &\quad + D_1^2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))(\mu_1)\nu_1 \\ &\quad + D_1 D_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))(D\varphi(\lambda_{1,n})\mu_1)\nu_1 \\ &\quad + D_2 D_1 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))(\mu_1)D\varphi(\lambda_{1,n})\nu_1 \\ &\quad + D_2^2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))(D\varphi(\lambda_{1,n})\mu_1)D\varphi(\lambda_{1,n})\nu_1 \end{aligned}$$

and

$$\begin{aligned} y &:= \partial D\varphi(\lambda_1)(\mu_1)\nu_1 \\ &= \partial_2 G(\lambda_1, \varphi(\lambda_1))\partial D\varphi(\lambda_1)(\mu_1)\nu_1 \\ &\quad + D_1^2 G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1 \\ &\quad + D_1 D_2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1 \\ &\quad + D_2 D_1 G(\lambda_1, \varphi(\lambda_1))(\mu_1)D\varphi(\lambda_1)\nu_1 \\ &\quad + D_2^2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)D\varphi(\lambda_1)\nu_1 \end{aligned}$$

In this way we obtain that

$$\begin{aligned} &\|\partial D\varphi(\lambda_{1,n})(\mu_1)\nu_1 - \partial D\varphi(\lambda_1)(\mu_1)\nu_1\| \\ &= \|y_n - y\| \\ &\le \|\partial_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\|_{L(E)}\|y_n - y\| \\ &\quad + \|D_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n})) - D_2 G(\lambda_1, \varphi(\lambda_1))\|_{L(E_0,E)}\|\partial D\varphi(\lambda_1)(\mu_1)\nu_1\|_0 \\ &\quad + \|D_2^2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\|_{L(E_0,L(E_0,E))}\,\|D\varphi(\lambda_{1,n})\|_{L(\Lambda_1,E_0)}\|\mu_1\|_{\Lambda_1} \\ &\qquad\qquad \|D\varphi(\lambda_{1,n}) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_2^2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\|_{L(E_0,L(E_0,E))}\,\|D\varphi(\lambda_{1,n}) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\mu_1\|_{\Lambda_1} \\ &\qquad\qquad \|D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_2^2 G(\lambda_{1,n}, \varphi(\lambda_{1,n})) - D_2^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(E_0,L(E_0,E))}\,\|D\varphi(\lambda_1)\|^2_{L(\Lambda_1,E_0)} \\ &\qquad\qquad \|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_2 D_1 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\|_{L(E_0,L(\Lambda_0,E))}\|\mu_1\|_{\Lambda_0}\,\|D\varphi(\lambda_{1,n}) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)} \\ &\qquad\qquad \|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_2 D_1 G(\lambda_{1,n}, \varphi(\lambda_{1,n})) - D_2 D_1 G(\lambda_1, \varphi(\lambda_1))\,\|_{L(E_0,L(\Lambda_0,E))}\|\mu_1\|_{\Lambda_0} \\ &\qquad\qquad \|D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)}\|\nu_1\|_{\Lambda_1} \end{aligned}$$

$$+ \|D_1 D_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\|_{L(\Lambda_0, L(E_0,E))} \|D\varphi(\lambda_{1,n}) - D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)} \|\mu_1\|_{\Lambda_1} \|\nu_1\|_{\Lambda_0}$$
$$+ \|D_1 D_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n})) - D_1 D_2 G(\lambda_1 , \varphi(\lambda_1))\|_{L(\Lambda_0, L(E_0,E))} \|D\varphi(\lambda_1)\|_{L(\Lambda_1,E_0)} \|\mu_1\|_{\Lambda_1} \|\nu_1\|_{\Lambda_0}$$
$$+ \|D_1^2 G(\lambda_{1,n}, \varphi(\lambda_{1,n})) - D_1^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(\Lambda_0, L(\Lambda_0,E))} \|\mu_1\|_{\Lambda_0} \|\nu_1\|_{\Lambda_0}$$

Now we use assumption 1. which provides that

$$\|\partial_2 G(\lambda_{1,n}, \varphi(\lambda_{1,n}))\|_{L(E)} \leq \alpha < 1$$

Further we get by assumption 5. that

$$\|\partial D\varphi(\lambda_1)(\mu_1)\nu_1\|_0 \leq \|\partial D\varphi(\lambda_1)\|_{L(\Lambda_1, L(\Lambda_1,E_0))} \|\nu_1\|_{\Lambda_1} \|\mu_1\|_{\Lambda_1}$$

Since Λ_1 is continuously embedded in Λ_0 there exists a constant $c < 0$ such that

$$\|\eta_1\|_{\Lambda_0} \leq c\|\eta_1\|_{\Lambda_1} \quad \text{for all } \eta_1 \in \Lambda_1.$$

In this way we get that

$$\|\partial D\varphi(\lambda_{1,n})(\mu_1)\nu_1 - \partial D\varphi(\lambda_1)(\mu_1)\nu_1\| \leq \frac{1}{1-\alpha} c_n \|\mu_1\|_{\Lambda_1} \|\nu_1\|_{\Lambda_1}$$

where $c_n \longrightarrow 0$ as $n \to \infty$ because of the continuity of the second Fréchet derivatives of $G : \Lambda_0 \times E_0 \to E$ (assumption 3.) and the continuity of $D\varphi : \Lambda_1 \to L(\Lambda_1, E_0)$ (assumption 4.). That is all we had to show. □

Corollary D.14. *We assume that the assumptions 1.-4. of Theorem D.13 are fulfilled.*

(i) In addition, we demand that

7. $D\varphi : \Lambda_1 \to L(\Lambda_1, E_0)$ *is bounded*

8. the second order derivatives

$$\begin{aligned} D_1^2 G &: \Lambda_0 \times E_0 \to L(\Lambda_0, L(\Lambda_0, E)) \\ D_1 D_2 G &: \Lambda_0 \times E_0 \to L(\Lambda_0, L(E_0, E)) \\ D_2 D_1 G &: \Lambda_0 \times E_0 \to L(E_0, L(\Lambda_0, E)) \\ D_2^2 G &: \Lambda_0 \times E_0 \to L(E_0, L(E_0, E)) \end{aligned}$$

are bounded.

Then it follows that $\partial D\varphi : \Lambda_1 \to L(\Lambda_1, L(\Lambda_1, E))$ *is bounded.*

(ii) In addition, we assume that condition 5. of Theorem D.13 is fulfilled and that

9. the above functions appearing in 7. and 8. not only bounded but also uniformly continuous

10. $D_2G : \Lambda_0 \times E_0 \to L(E_0, E)$ *is uniformly continuous*

11. $\partial D\varphi : \Lambda_1 \to L(\Lambda_1, L(\Lambda_1, E_0))$ *is bounded.*
Then it follows that $\partial D\varphi = D^2\varphi : \Lambda_1 \to L(\Lambda_1, L(\Lambda_1, E))$ *is uniformly continuous.*

Proof. (i): Let $\lambda_1, \mu_1, \nu_1 \in \Lambda_1$. Since $\partial D\varphi : \Lambda_1 \to L(\Lambda_1, L(\Lambda_1, E))$ is given by

$$\begin{aligned}\partial D\varphi(\lambda_1)(\mu_1)\nu_1 = [I - \partial_2 G(\lambda_1, \varphi(\lambda_1))]^{-1} \\ \{D_1^2 G(\lambda_1, \varphi(\lambda_1))(\mu_1)\nu_1 \\ + D_1D_2G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)\nu_1 \\ + D_2D_1G(\lambda_1, \varphi(\lambda_1))(\mu_1)D\varphi(\lambda_1)\nu_1 \\ + D_2^2 G(\lambda_1, \varphi(\lambda_1))(D\varphi(\lambda_1)\mu_1)D\varphi(\lambda_1)\nu_1\}\end{aligned}$$

we get that

$$\begin{aligned}&\|\partial D\varphi(\lambda_1)(\mu_1)\nu_1\| \\ \leq\ &\frac{1}{1-\alpha}\big(\|D_1^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(\Lambda_0, L(\Lambda_0, E))} c^2 \|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_1D_2G(\lambda_1, \varphi(\lambda_1))\|_{L(\Lambda_0, L(E_0, E))}\|D\varphi(\lambda_1)\|_{L(\Lambda_1, E_0)}\ c\ \|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_2D_1G(\lambda_1, \varphi(\lambda_1))\|_{L(E_0, L(\Lambda_0, E))}\|D\varphi(\lambda_1)\|_{L(\Lambda_1, E_0)}\ c\ \|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1} \\ &\quad + \|D_2^2 G(\lambda_1, \varphi(\lambda_1))\|_{L(E_0, L(E_0, E))}\|D\varphi(\lambda_1)\|^2_{L(\Lambda_1, E_0)}\|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1}\big) \\ \leq\ &\frac{\tilde{c}}{1-\alpha}\|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1}\end{aligned}$$

because of the boundedness of the second order derivatives and of $D\varphi : \Lambda_1 \to L(\Lambda_1, E_0)$.
(ii): To prove the uniformly continuity let $\lambda_1, \tilde{\lambda}_1, \mu_1, \nu_1 \in \Lambda_1$. Using the estimate as in the proof of Theorem D.13 (iii) we obtain that

$$
\begin{aligned}
&\|\partial D\varphi(\lambda_1)(\mu_1)\nu_1 - \partial D\varphi(\tilde{\lambda}_1)(\mu_1)\nu_1\| \\
\leq\ & \frac{c}{1-\alpha}\big(\|D_2G(\lambda_1,\varphi(\lambda_1)) - D_2G(\tilde{\lambda}_1,\varphi(\tilde{\lambda}_1))\|_{L(E_0,E)} \\
&\qquad + \|D\varphi(\lambda_1) - D\varphi(\tilde{\lambda}_1)\|_{L(\Lambda_1,E_0)} \\
&\qquad + \|D_2^2G(\lambda_1,\varphi(\lambda_1)) - D_2^2G(\tilde{\lambda}_1,\varphi(\tilde{\lambda}_1))\|_{L(E_0,L(E_0,E))} \\
&\qquad + \|D_2D_1G(\lambda_1,\varphi(\lambda_1)) - D_2D_1G(\tilde{\lambda}_1,\varphi(\tilde{\lambda}_1))\|_{L(E_0,L(\Lambda_0,E))} \\
&\qquad + \|D_1D_2G(\lambda_1,\varphi(\lambda_1)) - D_1D_2G(\tilde{\lambda}_1,\varphi(\tilde{\lambda}_1))\|_{L(\Lambda_0,L(E_0,E))} \\
&\qquad + \|D_1^2G(\lambda_1,\varphi(\lambda_1)) - D_1^2G(\tilde{\lambda}_1,\varphi(\tilde{\lambda}_1))\|_{L(\Lambda_0,L(\Lambda_0,E))}\big)\|\mu_1\|_{\Lambda_1}\|\nu_1\|_{\Lambda_1}
\end{aligned}
$$

for some constant $c > 0$. Therefore the assertion follows since $\varphi : \Lambda_1 \to E_0$ is Lipschitz continuous if $D\varphi : \Lambda_1 \to L(\Lambda_1, E_0)$ is bounded by the fundamental theorem for Bochner integrals. □

Symbols

$L(U,H)$	space of all bounded and linear operators from U to H	p.153
$L(U)$	$L(U,U)$	
tr Q	trace of Q	p.154
$L_1(U,H)$	space of all nuclear operators from U to H	p.153
$N(m,Q)$	Gaussin measure with mean m and covarince Q	p.10
$W(t)$, $t \in [0,T]$	(standard) Wiener process	p.13
	cylindrical Wiener process	p.44
$E(X\|\mathcal{G})$	conditional expectation of X given $\mathcal{G}$	p.17
$\mathcal{M}_T^2(E)$	space of all continuous E-valued, square integrable martingales	p.19
$\mathcal{E}$	class of all $L(U,H)$-valued elementary processes	p.21
$L_2(U,H)$	space of all Hilbert-Schmidt operators from U to H	p.155
A^*	adjoint operator of $A \in L(U,H)$	p.153
$Q^{\frac{1}{2}}$	square root of $Q \in L(U)$	p.23
T^{-1}	(pseudo) inverse of $T \in L(U,H)$	p.159
U_0	$Q^{\frac{1}{2}}(U)$	p.25
L_2^0	$L_2(U_0,H)$	p.25
Ω_T	$[0,T] \times \Omega$	p.20
dx	Lebesgue measure	p.20
P_T	$dx_{\|[0,T]} \otimes P$	p.20
$\mathcal{P}_T$	predictable σ-field on Ω_T	p.26
$L^p(\Omega,\mathcal{F},\mu;X)$	set of all with respect to μ p-integrable mappings from Ω to X	p.150
$L^p(\Omega,\mathcal{F},\mu)$	$L^p(\Omega,\mathcal{F},\mu;\mathbb{R})$	
L_0^p	$L^p(\Omega,\mathcal{F}_0,P;H)$	
$L^p([0,T];H)$	$L^p([0,T],\mathcal{B}([0,T]),dx;H)$	
$L^p([0,T],dx)$	$L^p([0,T];\mathbb{R})$	
$\\| \\|_T$	L^2-norm on $L^2(\Omega_T,\mathcal{P}_T,P_T;L_2^0)$	p.24

$\mathcal{N}_W^2(0,T;H)$	$L^2(\Omega_T,\mathcal{P}_T,P_T;L_2^0)$	p.26
$\mathcal{N}_W^2(0,T)$	$\mathcal{N}_W^2(0,T;H)$	
$\mathcal{N}_W^2$	$\mathcal{N}_W^2(0,T)$	
$\mathcal{N}_W(0,T;H)$	space of all stochastically integrable processes	p.29
$\mathcal{N}_W(0,T)$	$\mathcal{N}_W(0,T;H)$	
$\mathcal{N}_W$	$\mathcal{N}_W(0,T)$	
$\partial F(x;y)$	directional derivative of F in x in the direction y	p.164
∂F	Gâteaux derivative of F	p.165
DF	Fréchet derivative of F	p.166
$\partial_k F(x;y_k)$	directional derivative of F in the kth variable x_k of $x=(x_1,\dots,x_n)$ in the direction y_k	
$\partial_k F$	Gâteaux derivative of F in the kth variable	
$D_k F$	Fréchet derivative of F in the kth variable	
$\frac{\partial}{\partial t}F$	derivative of $F(\cdot,x):[0,T]\to\mathbb{R}$	p.37
$C^k([0,T],E)$	space of all k times continuously differentiable functions from $[0,T]$ to E	p.54
$C^k([0,T])$	$C^k([0,T],\mathbb{R})$	
$C_b^k(H)$	space of all k times Fréchet differentiable functions from H to $\mathbb{R}$ with bounded derivatives up to order k	p.129
$UC_b^k(H)$	subspace of all functions in $C_b^k(H)$ with uniformly continuous derivatives	p.129
M_T	$\sup_{t\in[0,T]}\|S(t)\|_{L(H)}$	p.66
$\|Y\|_{\mathcal{H}^p}$	$\sup_{t\in[0,T]}\|Y(t)\|_{L^p}$	p.67
$\|Y\|_{p,\lambda,T}$	$\sup_{t\in[0,T]} e^{-\lambda t}\|Y(t)\|_{L^p}$	p.67
$\mathcal{H}^p(T,H)$	$\{Y:\Omega_T\to H \mid Y$ predictable and $\|Y(t)\|_{\mathcal{H}^p}<\infty\}$	p.67
$W_A^\Phi(t),\ t\in[0,T]$	stochastic convolution	p.84

Bibliography

[Al 92] Alt, H.W.: Lineare Funktionalanalysis. Berlin-Heidelberg-New York: Springer Verlag 92

[AlKoRö 95] Albeverio, S., Kondratiev, Y.G., Röckner, M.: Dirichlet operators via stochastic analysis. Journal of Functional Analysis **128**, p.102-138, (1995)

[BeDal 80] Belopolskaja, J.I., Daletsky, J.L.: Diffusion processes in smooth Banach spaces and manifolds. Trans Moscow Math. Soc. **37**, p.113-150, (1980)

[Bi 68] Billingsley, P.: Convergence of probability measures. New York-London-Sydney-Toronto: John Wiley & Sons, Inc. 68

[Ce 98] Cerrai, S.: Differentiability with respect to initial datum for solutions of SPDE's with no Fréchet differentiable drift term. Comm. in Applied Analysis **2**, p.249-270, (1998)

[Ce 2000] Cerrai, S.: Second order PDE's in finite and infinite dimensions. A probabilistic approach. Lecture Notes in Mathematics **1762**, Springer Verlag 2001

[Co 80] Cohn, D.L.: Measure theory. Boston: Birkhäuser 1980

[Dal 66] Daletsky, J.L.: Differential equations with functional derivatives and stochastic equations for generalized random processes. Dokl. Akad. Nauk SSSR **166**, p.1035-38, (1966)

[Dal 67] Daletsky, J.L.: Infinite-dimensional elliptic operators and parabolic equations connected with them. Russian Math. Surveys **22**, p.1-53 (1967)

[DaPrElZa 95] Da Prato, G., Elworthy, K.D., Zabczyk, J.: Strong Feller property for stochastic semilinear equations. Stochastic Analysis Appl. **13**, no.1, p.35-45 (1995)

[DaPrZa 92] Da Prato, G., Zabczyk, J.: Stochastic equations in infinite dimensions. Cambridge: Cambridge University Press 1992

[DaPrZa 96] Da Prato, G., Zabczyk, J.: Ergodicity for infinite dimensional systems. Cambridge: Cambridge University Press 1996

[ElLi 93] Elworthy, K.D., Li, X.-M.: Differentiation of heat semigroups and applications. Probability Theory and mathematical statistics, Vilnius 1993, p.239-251

[Gr 67] Gross, L.: Potential theory on Hilbert spaces. Journal of Functional Analysis **1**, p.123-181 (1967)

[Ku 90] Kunita, H.: Stochastic flows and stochastic differential equations. Cambridge: Cambridge University Press 1990

[MasSei 2000] Maslowski, B., Seidler, J.: Probabilistic approach to the strong Feller property. Probab. Theory Related Fields **118**, no.2, p.187-210 (2000)

[MaRö 92] Ma, Z.-M., Röckner, M.: Introduction to the theory of (non-symmetric) Dirichlet forms. Berlin-Heidelberg-New York: Springer Verlag 1992

[MeVo 92] Meise, R., Vogt, D.: Einführung in die Funktionalanalysis. Braunschweig-Wiesbaden: Vieweg Verlag 1992

[Pa 83] Pazy, A.: Semigroups of linear operators. Berlin-Heidelberg-New York-Tokyo: Springer Verlag 1983

[ReSi 72] Reed, M., Simon, B.: Methods of modern mathematical physics. New York-London: Academic Press 1972

[Za 98] Zabczyk, J.: Parabolic equations on Hilbert spaces. In: Stochastic PDE's and Kolmogorov Equations in Infinite Dimensions, p.117-213, Lecture Notes in Mathematics, Cetraro, Italy 1998, Editor: Da Prato, G., Berlin-Heidelberg-New York: Springer Verlag 1998

[Zü 95] Zühlsdorff, C.: Zur Anfangswertabhängigkeit von Lösungen stochastischer Differential-Gleichungen in Hilbert-Räumen. Diploma-Thesis, Bonn 1995