Thesis for the Degree of Doctor of Philosophy

On Weak Convergence, Malliavin Calculus and Kolmogorov Equations in Infinite Dimensions

Adam Andersson

CHALMERS UNIVERSITY OF GOTHENBURG

Division of Mathematics
Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
Göteborg, Sweden 2015

On Weak Convergence, Malliavin Calculus and Kolmogorov Equations in Infinite Dimensions
Adam Andersson
ISBN 978-91-7597-129-2

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 3810
ISSN 0346-718X

Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
SE–412 96 Göteborg
Sweden
Phone: +46 (0)31-772 10 00

Printed in Göteborg, Sweden, 2014

On Weak Convergence, Malliavin Calculus and Kolmogorov Equations in Infinite Dimensions

Adam Andersson

Abstract

This thesis is focused around weak convergence analysis of approximations of stochastic evolution equations in Hilbert space. This is a class of problems, which is sufficiently challenging to motivate new theoretical developments in stochastic analysis. The first paper of the thesis further develops a known approach to weak convergence based on techniques from the Markov theory for the stochastic heat equation, such as the transition semigroup, Kolmogorov's equation, and also integration by parts from the Malliavin calculus. The thesis then introduces a novel approach to weak convergence analysis, which relies on a duality argument in a Gelfand triple of refined Sobolev-Malliavin spaces. These spaces are introduced and a duality theory is developed for them. The family of refined Sobolev-Malliavin spaces contains the classical Sobolev-Malliavin spaces of Malliavin calculus as a special case. The novel approach is applied to the approximation in space and time of semilinear parabolic stochastic partial differential equations and to stochastic Volterra integro-differential equations. The solutions to the latter type of equations are not Markov processes, and therefore classical proof techniques do not apply. The final part of the thesis concerns further developments of the Markov theory for stochastic evolution equations with multiplicative non-trace class noise, again motivated by weak convergence analysis. An extension of the transition semigroup is introduced and it is shown to provide a solution operator for the Kolmogorov equation in infinite dimensions. Stochastic evolution equations with irregular initial data are used as a technical tool and existence and uniqueness of such equations are established. Application of this theory to weak convergence analysis is not a part of this thesis, but the tools for it are developed.

Keywords: Stochastic evolution equations, stochastic Volterra equations, weak approximation, Kolmogorov equations in infinite dimensions, Malliavin calculus, finite element method, backward Euler method

Preface

This thesis consists of the following papers.

- Adam Andersson and Stig Larsson,
  “*Weak convergence for a spatial approximation of the nonlinear stochastic heat equation*”,
  accepted for publication in *Math. Comp.*

- Adam Andersson, Raphael Kruse and Stig Larsson,
  “*Refined Sobolev-Malliavin spaces and weak approximations of SPDE*”,
  submitted.

- Adam Andersson, Mihály Kovács and Stig Larsson,
  “*Weak error analysis for semilinear stochastic Volterra equations with additive noise*”,
  preprint.

- Adam Andersson and Arnulf Jentzen,
  “*Existence, uniqueness and regularity for stochastic evolution equations with irregular initial values*”,
  preprint.

In all papers I have made major contributions to the development of the ideas, to the proofs, and to the writing.

Acknowledgements

Most of my friends throughout the years were people who I had various creative projects with. As I am writing this acknowledgment it seems like this still holds true at this very moment.

First of all I want to thank Stig Larsson who was my PhD supervisor. I enjoyed very much our discussions and your generosity with your time, deep knowledge, your broad network and your patience during the early time of my PhD before I had learned my subject.

I very much want to thank Raphael Kruse. I claim that good cooperation is based, except for the presence of individual skills and creativity, also on mutual generosity as well as a suitable amount of competition. All these ingredients are present in our cooperation.

I want to thank Arnulf Jentzen for the work with paper IV and our many discussions and for hosting me at ETH, Zürich for two weeks in 2012. Our collaboration has boosted my mathematical skills more than any other collaboration. Thank you.

Next I wish to thank Mihály Kovács for the work with paper III. In other projects most ideas and solutions came during focused moments in solitude. In this project we carried out all proofs together at the blackboard and this was a nice way to do mathematics.

I want to thank Felix Lindner for hosting me in Kaiserslautern for three days in the summer of 2014. I enjoyed our discussions and hope that we will continue with our joint work. I want to thank Boualem Djehiche for hosting me for two weeks at KTH, Stockholm. Unfortunately our work could not be finished for this thesis. I want to thank Xiaojie Wang for interesting conversations related to the work with a new project.

Furthermore, I wish to thank Peter Sjögren for the work with the lecture notes on "Ornstein Uhlenbeck Theory in Finite Dimensions" and for lecturing this nice course. My gratitude goes also to Grigori Rozenblioum, for his kindness, with solving a mathematical problem I presented to him.

I want to thank the members of our SPDE group: Annika Lang, Fredrik Lindgren, Matteo Molteni, and Kristin Kirchner for nice discussions and the work in various study groups we have organized over the years . I want to thank the members of the CAM-group and all the PhD students and staff at the department of mathematical sciences at Chalmers who has contributed to a great working atmosphere. Among the administrative staff I want to thank Lars Almqvist, Lotta Fernström, Marie Kühn, Jeanette Montell-Westerlin, Camilla Nygren for friendliness and efficiency.

I also want to take the opportunity to thank old friends. Thanks Robert Almstrand, Simon Olén and Johan Palm for great times. Johan, you are the most clear exception to the statement in the first paragraph.

I want to thank my parents for all their support and love. They have supported me in my early detours in life and not forced me to follow any straight or predetermined path in life.

Last but not the least I want to thank my beloved wife Fereshteh, my dear doughter Alice and and my dear son Edvin for their love, patience and support.

Adam Andersson
Göteborg, December 2014

Contents

Part I

INTRODUCTION

Introduction

1. A first overview

The main theme of this thesis is the study of approximation, regularity, existence, and uniqueness for the semilinear stochastic evolution equation

$$(1.1)\qquad \mathrm{d}X_t + AX_t\,\mathrm{d}t = F(X_t)\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t, \quad t \in (0,T]; \quad X_0 = \xi,$$

and its transition semigroup $(\mathcal{P}_t)_{t\in[0,T]}$, that is the family of mappings, which act on sufficiently regular functions $\varphi\colon H \to \mathbf{R}$ by

$$(\mathcal{P}_t\varphi)(x) = \mathbf{E}\Big[\varphi(X_t)|X_0 = x\Big].$$

The solution $(X_t)_{t\in[0,T]}$, is a stochastic process, taking values in a separable Hilbert space $(H, \|\cdot\|, \langle\cdot,\cdot\rangle)$. The operator $-A\colon H \subset \mathcal{D}(A) \to H$ is the generator of an analytic semigroup $(S_t)_{t\geq 0} = (e^{-tA})_{t\geq 0}$ of bounded linear operators $H \to H$. The nonlinear drift coefficient $F\colon H \to H$ is assumed to be globally Lipschitz continuous. The driving stochastic process W is a cylindrical id_U-Wiener process, where U is another separable Hilbert space, defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with filtration $(\mathcal{F}_t)_{t\in[0,T]}$. The noise coefficient B maps H into the space of Hilbert-Schmidt operators $U \to \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space with $H \subset \mathcal{H}$ being dense and continuous. The mapping B is assumed to be globally Lipschitz continuous. The initial value $\xi\colon \Omega \to H$ is assumed to satisfy some condition on smoothness and integrability. Further restrictions on F and B are imposed in various parts of the thesis.

By a solution to (1.1) we mean a stochastic process $X \in \mathcal{C}(0,T;L^2(\Omega;H))$, which for all $t \in [0,T]$, satisfies $\mathbf{P}$-almost surely

$$(1.2)\qquad X_t = S_t\xi + \int_0^t S_{t-s}F(X_s)\,\mathrm{d}s + \int_0^t S_{t-s}B(X_s)\,\mathrm{d}W_s.$$

The space $\mathcal{H}$, which is a negative order interpolation space corresponding to the operator A, determines the regularity of the solution. The choice $\mathcal{H} = H$ gives the highest regularity that we consider in this thesis and corresponds to trace class noise. In all papers in this thesis we include space-time white noise as a special case. For sufficiently large spaces $\mathcal{H}$, there is no solution.

Let $(X^h)_{h\in(0,1)} \subset L^\infty(0,T;L^2(\Omega;H))$, be a family of approximations to X. This family is said to converge strongly to X as $h \downarrow 0$, with strong order $\beta > 0$, if there exists C, such that

$$\sup_{t\in[0,T]} \left\|X_t - X_t^h\right\|_{L^2(\Omega;H)} \leq Ch^\beta, \quad h \in (0,1). \tag{1.3}$$

The family $(X^h)_{h\in(0,1)}$ is said to converge weakly to X, with weak rate $\gamma > 0$, if for all sufficiently smooth $\varphi\colon H \to \mathbf{R}$, there exists C, such that

$$\left|\mathbf{E}\Big[\varphi(X_t) - \varphi(X_t^h)\Big]\right| \leq Ch^\gamma, \quad h \in (0,1).$$

In Papers I–III we consider different choices of assumptions for A, F, B, ξ and different approximating families $(X^h)_{h\in(0,1)}$, which converge strongly to X with some rate $\beta > 0$. In all these papers we essentially consider the same goal: show that, for all sufficiently smooth $\varphi\colon H \to \mathbf{R}$, the approximations $(X^h)_{h\in(0,1)}$, converge weakly to X with any weak rate $\gamma \in (0, 2\beta)$, i.e., essentially twice the strong rate.

In probability theory a sequence of probability measures $(\mu_n)_{n\in\mathbf{N}}$ on H is said to converge weakly to a measure μ on H, if for every bounded and continuous function $\varphi\colon H \to \mathbf{R}$ it holds that

$$\int_H \varphi \,\mathrm{d}\mu_n - \int_H \varphi \,\mathrm{d}\mu \to 0, \quad \text{as} \quad n \to \infty,$$

see, e.g., Billingsly [5]. Let $\mathrm{P}_1(H)$ denote the set of all probability measures ν on H, which satisfy $\int_H \|x\| \,\mathrm{d}\nu(x) < \infty$. For two probability measures $\nu_1, \nu_2 \in \mathrm{P}_1(H)$, the Wasserstein distance $\mathcal{W}_1(\nu_1, \nu_2)$ is given by

$$\mathcal{W}_1(\nu_1,\nu_2) = \sup_\varphi \left\{ \int_H \varphi \,\mathrm{d}\nu_1 - \int_H \varphi \,\mathrm{d}\nu_2 : |\varphi(x) - \varphi(y)| \leq \|x - y\| \right\}.$$

The metric $\mathcal{W}_1$ determines weak convergence in the following sense: a family $(\mu_n)_{n\in\mathbf{N}} \subset \mathrm{P}_1(H)$ converges weakly to $\mu \in \mathrm{P}_1(H)$ if and only if $\mathcal{W}_1(\mu_n, \mu) \to 0$ as $n \to \infty$. If $\mu_h = \mathrm{Law}(X_t^h) = \mathbf{P} \circ (X_t^h)^{-1}$, $h \in (0,1)$, are the distributions of X_t^h, $h \in (0,1)$, and $\mu = \mathrm{Law}(X_t) = \mathbf{P} \circ (X_t)^{-1}$ is the distribution of X_t, then it holds that

$$\int_H \varphi \,\mathrm{d}\mu_h = \mathbf{E}[\varphi(X_t^h)] \quad \text{and} \quad \int_H \varphi \,\mathrm{d}\mu = \mathbf{E}[\varphi(X_t)].$$

By (1.3) it follows that

$$\begin{aligned}\mathcal{W}_1(\mu_h, \mu) &= \sup_\varphi \left\{ \left|\mathbf{E}\Big[\varphi(X_t^h) - \varphi(X_t)\Big]\right| : |\varphi(x) - \varphi(y)| \leq \|x - y\| \right\} \\ &\leq \left\|X_t^h - X_t\right\|_{L^2(\Omega;H)} \leq Ch^\beta.\end{aligned}$$

Thus, the rate of weak convergence, measured in the Wasserstein distance, is not less than the strong rate of convergence, but has never been proved to exceed it. However, by increasing the smoothness of the class of test functions, one can often, depending on the problem, determine a weak rate of convergence, which exceeds the strong rate. To formalize this statement we introduce the distances $\mathcal{W}_1^k$, $k \in \mathbf{N}$, on $\mathrm{P}_1(H)$, given by

$$\mathcal{W}_1^k(\nu_1,\nu_2) = \sup_{\varphi}\Big\{ \int_H \varphi \,\mathrm{d}\nu_1 - \int_H \varphi \,\mathrm{d}\nu_2 : \|\varphi^{(1)}\|,\dots,\|\varphi^{(k)}\| \le 1 \Big\},$$

where $\varphi^{(1)},\dots,\varphi^{(k)}$ denote the Fréchet derivatives of φ up to order k, with the relevant norms for the derivatives of different orders. From existing results in the literature and, in particular, from the results in Papers I–III one can write, with $k=2$ or $k=3$, depending on which type of approximation is considered, the weak convergence in the form

$$\mathcal{W}_1^k(\mu_h,\mu) = \mathcal{W}_1^k(\mathrm{Law}(X_t^h),\mathrm{Law}(X_t)) \le C_\gamma h^\gamma, \quad h \in (0,1), \quad \gamma \in (0,2\beta).$$

As the title of this thesis suggests, we also treat Malliavin calculus and Kolmogorov equations in infinite dimensions. Techniques from both fields are important for weak convergence analysis. In fact we are not aware of any proof of weak convergence, except in the case of linear equations, which does not rely either on Malliavin calculus or on the use of Kolmogorov's equation. In Paper IV we show that under suitable regularity assumptions on F, B, φ, it holds that the function $u\colon [0,T]\times H \to \mathbf{R}$, which for all $t\in[0,T]$, $x \in H$, is given by $u(t,x) = (\mathcal{P}_t\varphi)(x)$, is the solution of the Kolmogorov equation: for $(t,x)\in(0,T]\times H$,

$$\begin{aligned}\frac{\partial u(t,x)}{\partial t} &= \frac{\partial u(t,x)}{\partial x}\Big(-Ax + F(x)\Big) + \frac{1}{2}\sum_{h\in\mathbb{H}} \frac{\partial^2 u(t,x)}{\partial x^2}\Big(B(x)h, B(x)h\Big),\\ u(0,x) &= \varphi(x).\end{aligned}$$

Here $\mathbb{H}\subset H$ is an ON-basis and $\frac{\partial u(t,x)}{\partial x}(\phi_1)$ and $\frac{\partial^2 u(t,x)}{\partial x^2}(\phi_2,\phi_3)$ denote the first and second directional x-derivatives in directions ϕ_1 and ϕ_2,ϕ_3, respectively. In order to make sense of this equation, in the case $\mathcal{H}\neq H$, we must extend $(\mathcal{P}_t)_{t\in[0,T]}$, so that $u(t,x) = (\mathcal{P}_t\varphi)(x)$ is defined on a larger space than H. In order to do this, careful analysis is needed, in particular, stochastic evolution equations with non-smooth initial value and random, time-dependent coefficients. Paper IV contains an existence and uniqueness result for this type of equations.

2. Stochastic integration and Malliavin calculus

In this section we explain both the basic stochastic analysis that is needed to define a solution to (1.1) and elements of the Malliavin calculus, which we use to study weak convergence. The presentation of the stochastic integral follows to a large extent the lecture notes of van Neerven [55], see also Brzeźniak [11], Da Prato & Zabczyk [17], Pezat & Zabczyk [50], Prévôt & Röckner [51]. The presentation of the Malliavin calculus follows Andersson et al. [2], Kruse [40]. For earlier works on Malliavin calculus in the Hilbert space setting, see Grorud & Pardoux [24], León & Nualart [41]. For basic Malliavin calculus we recommend Nualart [47], Privault [52] and for a general exposition of Gaussian analysis see the excellent books by Janson [28] and Bogachev [6].

2.1. The cylindrical Wiener process. Let $(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U)$ be a separable Hilbert space with an ON-basis $\mathbb{U} \subset U$, let $\mathbb{U}^* \subset U^*$ be the dual ON-basis, which is related to $\mathbb{U}$ by $u^* = \langle u, \cdot\rangle_U$ for $u \in \mathbb{U}$. Let $(\beta_t^u)_{t\in[0,T]}$, $u \in \mathbb{U}$, be a sequence of independent standard Brownian motions defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, adapted to a filtration $(\mathcal{F}_t)_{t\in[0,T]}$. We define a cylindrical id_U-Wiener process $W\colon U \to L^2([0,T]\times\Omega;\mathbf{R})$ as the strong operator limit

$$W = \sum_{u\in\mathbb{U}} \beta^u \otimes u^*.$$

Thus, for all $v \in U$, it holds that $Wv = \sum_{u\in\mathbb{U}} \beta^u \langle u, v\rangle_U$. Since for all $u \in \mathbb{U}$ it holds that $\mathbf{E}|\beta_t^u|^2 = t\|u\|_U^2$, and because $(\beta^u)_{u\in\mathbb{U}}$ is an orthogonal system in $L^2(\Omega\times[0,T];\mathbf{R})$ by independence, it holds by Parseval's identity for all $t \in [0,T]$, $v \in U$, that

$$\mathbf{E}\Big|W_t v\Big|^2 = t \sum_{u\in\mathbb{U}} |\langle u, v\rangle_U|^2 = t\|v\|_U^2, \tag{2.1}$$

More generally, one can show, by the polarization identity, that

$$\mathbf{E}\Big[W_t u\, W_s v\Big] = \min(s,t)\langle u, v\rangle_U, \quad s,t \in [0,T],\ u,v \in U. \tag{2.2}$$

As a convergent sum of weighted Brownian motions, for all $v \in U$, it holds that $(W_t v)_{t\in[0,T]}$ is a Brownian motion with covariance

$$\mathrm{Cov}(W_t v, W_s v) = \min(s,t)\|v\|_U^2.$$

This property is often taken together with (2.2) as the definition of the Cylindrical Wiener process, without any explicit construction.

Let Q be a selfadjoint, positive semidefinite, bounded linear operator $H \to H$. Sometimes, in particular, in Papers I–III of this thesis, the spaces H and U are related by $U = Q^{\frac{1}{2}}(H)$, equipped with the inner product

$$\langle u, v\rangle_U = \langle Q^{-\frac{1}{2}} u, Q^{-\frac{1}{2}} v\rangle,$$

where $Q^{-\frac{1}{2}}$ denotes the pseudo inverse of $Q^{\frac{1}{2}}$. In this case it is common to write that W is a cylindrical Q-Wiener process. If Q is of trace class, i.e., if

$\mathrm{Tr}(Q) = \sum_{h\in\mathbb{H}}\langle Qh,h\rangle = \sum_{h\in\mathbb{H}}\|Q^{\frac{1}{2}}h\|^2 < \infty$, where $\mathbb{H} \subset H$ is an arbitrary ON-basis, then the canonical embedding $i\colon U \to H, u \mapsto u$ is a Hilbert-Schmidt operator and the series

$$B_t = \sum_{u\in\mathbb{U}} \beta_t^u \otimes u$$

converges in $L^2(\Omega;H)$, since $(\|i(u)\|)_{u\in\mathbb{U}}$ is a square summable sequence. The process $(B_t)_{t\in[0,T]}$ is called an H-valued Brownian motion. If $\mathrm{Tr}(Q) = \infty$, then B_t converges in any larger Hilbert space $\tilde{H}$, such that the embedding $U \to \tilde{H}$ is Hilbert-Schmidt. This is a common way to define the Q-Wiener process, but we prefer the notion of cylindrical Wiener process, since it is defined the same way regardless what the space U is or, equivalently, what properties Q has.

2.2. The stochastic Wiener integral. The theory for stochastic integration goes back to Wiener [60] and Paley, Wiener & Zygmund [48] for deterministic integrands and to Itō [27] for stochastic integrands. Let $\mathcal{L}_2(U;H)$ denote the space of all Hilbert-Schmidt operators $U \to H$, let $\Phi \in L^2(0,T;\mathcal{L}_2(U;H))$ be a simple, finite-rank integrand, given by

$$\Phi = \sum_{n=1}^{N} \mathbf{1}_{(t_{n-1},t_n]} \otimes \Big(\sum_{j=1}^{k} h_{j,n} \otimes u_j\Big),$$

where $0 = t_1 < \cdots < t_n < \cdots < t_N = T$, $(h_{j,n})_{j=1}^k \subset H$, $n \in \{1,\ldots,N\}$, and $(u_j)_{j=1}^k \subset U$ are orthonormal, $k,N \in \mathbf{N}$. The H-valued Wiener integral $\int_0^T \Phi_t \,\mathrm{d}W_t$ of Φ is the random variable

$$\int_0^T \Phi_t \,\mathrm{d}W_t = \sum_{n=1}^{N}\sum_{j=1}^{k}\Big(W_{t_n}u_j - W_{t_{n-1}}u_j\Big)\otimes h_{j,n}.$$

From the independence of increments and the independence of the Brownian motions $(Wu_j)_{j=1}^k$ it holds that the summands form an orthogonal system in $L^2(\Omega;H)$. Therefore, since $\mathbf{E}[|W_{t_n}u_j - W_{t_{n-1}}u_j|^2] = (t_n - t_{n-1})\|u_j\|_U^2$, and since $\|u \otimes h\|_{U\otimes H} = \|u\|_U\|h\|_H$, it holds that

$$\begin{aligned}
\mathbf{E}\Big[\Big\|\int_0^T \Phi_t \,\mathrm{d}W_s\Big\|^2\Big] &= \sum_{n=1}^{N}(t_n - t_{n-1})\sum_{j=1}^{k}\|u_j\|_U^2\|h_{j,n}\|^2 \\
&= \sum_{n=1}^{N}(t_n - t_{n-1})\sum_{j=1}^{k}\|u_j \otimes h_{j,n}\|_{U\otimes H}^2 \\
&= \int_0^T \|\Phi_t\|_{\mathcal{L}_2(U;H)}^2 \,\mathrm{d}t,
\end{aligned}$$

i.e., we have the Wiener isometry

$$\Big\| \int_0^T \Phi_t \,\mathrm{d}W_s \Big\|_{L^2(\Omega;H)} = \big\|\Phi_t\big\|_{L^2(0,T;\mathcal{L}_2(U;H))}. \tag{2.3}$$

The piecewise constant functions are dense in $L^2(0,T;\mathbf{R})$ and the finite-rank operators are dense in $\mathcal{L}_2(U;H)$. By the completeness of $L^2(0,T;\mathcal{L}_2(U;H))$ it follows it follows that the stochastic integral extends to all $\Phi \in L^2(0,T;\mathcal{L}_2(U;H))$. This integral is called the H-valued Wiener integral, see van Neerven [55]. Moreover, (2.3) holds for all $\Phi \in L^2(0,T;\mathcal{L}_2(U;H))$.

2.3. The stochastic Itō integral. In this section we consider stochastic integration with stochastic integrands. We follow the lecture notes by van Neerven [55], which develops stochastic integration in Banach spaces of UMD-type. This is not the standard way to do it in Hilbert space, but we present this approach since it is elegant.

A stochastic process $\Phi\colon [0,T]\times\Omega \to \mathcal{L}_2(U;H)$ is said to be simple $\mathcal{L}_2(U;H)$-predictable, if it is of the form

$$\Phi = \sum_{n=1}^{N}\sum_{m=1}^{M} \mathbf{1}_{(t_{n-1},t_n]} \otimes \mathbf{1}_{A_{m,n}} \otimes \Big(\sum_{j=1}^{k} h_{j,n} \otimes u_j\Big), \tag{2.4}$$

where $0 = t_1 < \cdots < t_n < \cdots < t_N = T$, $A_{m,n} \in \mathcal{F}_{t_{n-1}}$, $m \in \{1,\dots M\}$, $n \in \{1,\dots,N\}$, $h_{j,n} \in H$, $j \in \{1,\dots,k\}$, $n \in \{1,\dots,N\}$, and $u_1,\dots,u_k \in U$ are orthonormal. It is clear that $\Phi \in L^2([0,T]\times\Omega;\mathcal{L}_2(U;H))$. The Itō integral of Φ is the H-valued random variable

$$\int_0^T \Phi_t \,\mathrm{d}W_t = \sum_{n=1}^{N}\sum_{m=1}^{M} \mathbf{1}_{A_{m,n}} \otimes \sum_{j=1}^{k} \big(W_{t_n}u_j - W_{t_{n-1}}u_j\big) \otimes h_{j,n}.$$

Let $\tilde W\colon U \to L^2([0,T]\times\tilde\Omega)$ be an id_H-Wiener process, which is defined on a probability space $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbf{P}})$. We denote expectation with respect to $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbf{P}})$ by $\tilde{\mathbf{E}}$. By a decoupling, inequality Theorem 13.1 in van Neerven [55], there exist for all $p \in [2,\infty)$, a constant C_p such that

$$\mathbf{E}\Big[\Big\|\int_0^T \Phi_t \,\mathrm{d}W_t\Big\|^p\Big] \le C_p \mathbf{E}\Big[\tilde{\mathbf{E}}\Big[\Big\|\int_0^T \Phi_t \,\mathrm{d}\tilde W_t\Big\|^p\Big]\Big]. \tag{2.5}$$

The constant C_p is uniform with respect to k,M,N. In this situation results for the Wiener integral apply since the integrand can be considered deterministic with respect to $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbf{P}})$. First, the Kahane-Khintchine inequality, in van Neerven [55, Corollary 4.13], states in particular that the $L^p(\Omega;H)$-norms are all equivalent on the space consisting of all Gaussian H-valued random variables. Therefore, there exists a new constant C_p', such that

$$\begin{aligned}\mathbf{E}\Big[\Big\|\int_0^T \Phi_t \, \mathrm{d}W_t\Big\|^p\Big] &\le C_p'\mathbf{E}\Big[\Big(\tilde{\mathbf{E}}\Big[\Big\|\int_0^T \Phi_t \, \mathrm{d}\tilde{W}_t\Big\|^2\Big]\Big)^{\frac{p}{2}}\Big] \\ &= C_p'\mathbf{E}\Big[\Big(\int_0^T \|\Phi\|^2_{\mathcal{L}_2(U;H)} \, \mathrm{d}t\Big)^{\frac{p}{2}}\Big].\end{aligned} \tag{2.6}$$

For the equality we used the Wiener isometry (2.3). Since H is a Hilbert space it holds that $C_2 = 1$, for $p = 2$, and equality holds in (2.5). This holds by (12.1), Definition 12.3 of a UMD-space, and the proof of Theorem 13.1 in van Neerven [55]. In this way we obtain the Itō isometry

$$\Big\|\int_0^T \Phi_t \, \mathrm{d}W_t\Big\|_{L^2(\Omega;H)} = \|\Phi\|_{L^2([0,T]\times\Omega;\mathcal{L}_2(U;H))}. \tag{2.7}$$

Let $L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$ denote the closure in $L^2([0,T]\times\Omega;\mathcal{L}_2(U;H))$ of all simple $\mathcal{L}_2(U;H)$-predictable processes. We say that $\Phi \in L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$ is an $\mathcal{L}_2(U;H)$-predictable process. By (2.7) the stochastic integral extends to all of $L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$. The constant C_p' in (2.6) is known to be bounded from above by $C_p' \le (p(p-1)/2)^{p/2}$, see Lemma 7.7 in Da Prato & Zabczyk [17]. We restate it: for all $\Phi \in L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$, $p \in [2,\infty)$, it holds that

$$\Big\|\int_0^T \Phi_t \, \mathrm{d}W_t\Big\|_{L^p(\Omega;H)} \le \sqrt{\frac{p(p-1)}{2}}\,\|\Phi\|_{L^p(\Omega;L^2(0,T;\mathcal{L}_2(U;H)))}. \tag{2.8}$$

2.4. Malliavin calculus. It is safe to say that integration by parts is a very powerful tool in mathematical analysis. Malliavin calculus offers a way to integrate by parts in stochastic analysis, which turns out to be very powerful indeed. It is a natural part of stochastic analysis. Malliavin calculus was introduced by Malliavin in [46], to give a probabilistic proof of Hörmander's Theorem on hypoelliptic partial differential operators.

To explain its power let us state a very simple question, which has no satisfactory answer without Malliavin calculus. By the polarization identity and (2.7) it holds for all $\Phi, \Psi \in L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$ that

$$\Big\langle\int_0^T \Psi_t \, \mathrm{d}W_t, \int_0^T \Phi_t \, \mathrm{d}W_t\Big\rangle_{L^2(\Omega;H)} = \big\langle\Psi,\Phi\big\rangle_{L^2([0,T]\times\Omega;\mathcal{L}_2(U;H))}. \tag{2.9}$$

This is the Itō isometry and it is the most basic result in stochastic analysis. From this basic result it is natural to ask: is there a useful result which applies if $\int_0^T \Psi \, \mathrm{d}W$ is replaced by a random variable $F \in L^2(\Omega;H)$? The answer is positive, if F has the proper regularity, and the formula reads:

$$\left\langle F, \int_0^T \Phi_t \,\mathrm{d}W_t \right\rangle_{L^2(\Omega;H)} = \left\langle DF, \Phi \right\rangle_{L^2([0,T]\times\Omega;\mathcal{L}_2(U;H))}. \tag{2.10}$$

Here $DF = (D_tF)_{t\in[0,T]}$ is an $\mathcal{L}_2(U;H)$-valued stochastic process and the unbounded operator $D\colon L^2(\Omega;H) \to L^2([0,T]\times\Omega;\mathcal{L}_2(U;H))$ is called the Malliavin derivative. We refer to (2.10) as the Malliavin integration by parts formula. It remains to understand the operator D, in order for (2.10) to be useful. Papers I and III contain brief introductions to Malliavin calculus and Paper II provides a theoretical account of Malliavin calculus. We use this section to complement these papers with some of the ideas behind Malliavin calculus and refer to Paper II for a more rigorous introduction.

Below we define the directional Malliavin derivative as a limit of difference quotients. In order to define a difference quotient, we need some notion of translation. The type of translation that we now introduce was first studied by Cameron & Martin [12], [13] for real-valued integrals By identifying $L^2(0,T;U) \cong L^2(0,T;\mathcal{L}_2(U;\mathbf{R}))$, it is clear that the mapping

$$I\colon L^2(0,T;U) \to L^2(\Omega;\mathbf{R}), \quad I(\phi) = \int_0^T \phi_t \,\mathrm{d}W_t,$$

is well defined. Moreover, for $\theta \in L^2(0,T;U)$, let

$$I^\theta\colon L^2(0,T;U) \to L^2(\Omega;\mathbf{R}), \quad I^\theta(\phi) = I(\phi) + \langle \phi, \theta\rangle_{L^2(0,T;U)}.$$

The Cameron-Martin Theorem in this setting states that for all $\theta \in L^2(0,T;U)$, the family $I^\theta(\phi)$, $\phi \in L^2(0,T;U)$, has the same distribution as the family $I(\phi)$, $\phi \in L^2(0,T;U)$, under the measure $\mathbf{Q}$, which is determined by

$$\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}} = \exp\Big(I(\theta) - \frac{1}{2}\|\theta\|^2_{L^2(0,T;U)}\Big),$$

see Bogachev [6, Theorem 1.4.2]. In particular, for all $n \in \mathbf{N}$, measurable functions $f\colon \mathbf{R}^n \to \mathbf{R}$, and $(\phi_i)_{i=1}^n \subset L^2(0,T;U)$, it holds that

$$\begin{aligned}\mathbf{E}\Big[f\Big(I^\theta(\phi_1),\dots,I^\theta(\phi_n)\Big)\Big]& \\ = \mathbf{E}\Big[f\Big(I(\phi_1),\dots,I(\phi_n)\Big)\exp\Big(I(\theta) - \frac{1}{2}\|\theta\|_{L^2(0,T;U)}\Big)\Big].&\end{aligned} \tag{2.11}$$

REMARK 2.1. Recall that we define the Cylindrical Wiener process as an operator $W\colon U \to L^2([0,T]\times\Omega;\mathbf{R})$. For $\theta \in L^2(0,T;U)$, we define $\theta^* \in L^2(0,T;U^*)$ by $\theta^*_t = \langle \theta_t, \cdot\rangle_U$, $t \in [0,T]$. With this notation we get that

$$I^\theta(\phi) = \int_0^T \phi_t\Big(\mathrm{d}W_t + \theta^*_t \,\mathrm{d}t\Big).$$

We think of $W^\theta := W + \int_0^\cdot \theta_s^* \, \mathrm{d}s$, as a translated Cylindrical Wiener process in the direction $\int_0^\cdot \theta_s^* \, \mathrm{d}s \colon U \to L^2(0,T;U^*)$, and $I^\theta(\phi)$ as the corresponding translation of $I(\phi)$.

In order to define the Malliavin derivative we introduce a suitable class of smooth random variables. For $q \in [2,\infty]$, let $\mathcal{S}^q$ denote the space of all random variables of the form

$$F = f(I(\phi_1),\dots,I(\phi_n)), \quad f \in \mathcal{C}_{\mathrm{p}}^1(\mathbf{R}^n;\mathbf{R}),\ (\phi_i)_{i=1}^n \subset L^q(0,T;U),\ n \in \mathbf{N}.$$

If $F \in \mathcal{S}^q$, then for $\theta \in L^2(0,T;U)$, we write $F^\theta = f(I^\theta(\phi_1),\dots,I^\theta(\phi_n))$. We define the directional Malliavin derivative of $F \in \mathcal{S}^q$, in direction $\theta \in L^2(0,T;U)$, by

$$D^\theta F = \lim_{\epsilon \to 0} \frac{F^{\epsilon\theta} - F}{\epsilon}.$$

First, $D^\theta I(\phi) = I(\phi) + \langle \theta, \phi \rangle_{L^2(0,T;U)} - I(\phi) = \langle \theta, \phi \rangle_{L^2(0,T;U)}$ and by the usual chain rule it holds that

$$D^\theta F = \sum_{i=1}^n \partial_i f(I(\phi_1),\dots,I(\phi_n)) \langle \theta, \phi_i \rangle_{L^2(0,T;U)}.$$

The Malliavin derivative is therefore the operator $D \colon \mathcal{S}^q \to L^2(\Omega; L^q(0,T;U))$, which is given by

$$DF = \sum_{i=1}^n \partial_i f(I(\phi_1),\dots,I(\phi_n)) \otimes \phi_i.$$

We now sketch how to prove a first version of the integration by parts formula. The formula states that for all $F \in \mathcal{S}^2$, $\phi \in L^2(0,T;U)$, it holds that

$$\Big\langle DF, \theta \Big\rangle_{L^2([0,T]\times\Omega;U)} = \Big\langle F, I(\theta) \Big\rangle_{L^2(\Omega;\mathbf{R})}. \tag{2.12}$$

This is proved by the dominated convergence theorem, the Cameron-Martin formula (2.11), and a first order Taylor expansion:

$$\begin{aligned}
\Big\langle DF, \theta \Big\rangle_{L^2([0,T]\times\Omega;U)} &= \mathbf{E}\Big[D^\theta F\Big] = \lim_{\epsilon\to 0} \epsilon^{-1} \mathbf{E}\Big[F^{\epsilon\theta} - F\Big] \\
&= \lim_{\epsilon\to 0} \epsilon^{-1} \mathbf{E}\Big[\Big(\exp\Big(I(\epsilon\theta) - \frac{1}{2}\|\epsilon\theta\|^2_{L^2(0,T;U)}\Big) - 1\Big)F\Big] \\
&= \lim_{\epsilon\to 0} \mathbf{E}\Big[\Big(I(\theta) + \mathcal{O}(\epsilon)\Big)F\Big] \\
&= \mathbf{E}\Big[FI(\theta)\Big] = \Big\langle F, I(\theta) \Big\rangle_{L^2(\Omega;\mathbf{R})}.
\end{aligned}$$

The use of the dominated convergence theorem must be justified, but we refrain from presenting the details.

For $q \in [0,\infty]$, let $\mathcal{S}^q(H)$ denote the space of random variables of the form $X = \sum_{j=1}^m F_j \otimes h_j$, for $(F_j)_{j=1}^m \subset \mathcal{S}^q$, $(h_j)_{j=1}^m \subset H$, $m \in \mathbf{N}$. The Malliavin derivative of $X \in \mathcal{S}^q(H)$ is the operator

$$D\colon \mathcal{S}^q(H) \to L^2(\Omega; L^q(0,T;\mathcal{L}_2(U;H))), \quad DX = \sum_{j=1}^m h_j \otimes DF_j.$$

The integration by parts formula (2.12) is the main tool in proving that the operator $D\colon \mathcal{S}^q(H) \to L^2(\Omega; L^q(0,T;\mathcal{L}_2(U;H)))$ is closable. For $p \in [2,\infty)$, $q \in [2,\infty]$, let $\mathbf{M}^{1,p,q}(H)$ denote the closure of $\mathcal{S}^q(H)$ under the norm

$$\|X\|_{\mathbf{M}^{1,p,q}(H)} = \Big(\|X\|^p_{L^p(\Omega;H)} + \|DX\|^p_{L^p(\Omega;L^q(0,T;\mathcal{L}_2(U;H)))} \Big)^{\frac{1}{p}}.$$

These spaces are Banach spaces and the space $\mathbf{M}^{1,2,2}(H)$ is a Hilbert space. In the literature the spaces $\mathbf{M}^{1,p,2}(H)$, $p \in [2,\infty)$, are often denoted $\mathbf{D}^{1,p}(H)$, see, e.g., Nualart [47]. We refer to the former as refined Sobolev-Malliavin spaces and the latter as classical Sobolev-Malliavin spaces. The refined Sobolev-Malliavin spaces were introduced in Paper II, and also used in Paper III. In Paper II we introduce a duality theory based on the Gelfand triple

$$\mathbf{M}^{1,p,q}(H) \subset L^2(\Omega;H) \subset \mathbf{M}^{1,p,q}(H)^*.$$

One of the main results in that paper is the following inequality: for all $p \in [2,\infty)$, $q \in [2,\infty]$, $\Phi \in L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$ and $\frac{1}{p} + \frac{1}{p'} = 1$, $\frac{1}{q} + \frac{1}{q'} = 1$, it holds that

$$\Big\| \int_0^T \Phi_t \,\mathrm{d}W_t \Big\|_{\mathbf{M}^{1,p,q}(H)^*} \leq \|\Phi\|_{L^{p'}(\Omega;L^{q'}(0,T;\mathcal{L}_2(U;H)))}. \tag{2.13}$$

This should be compared with (2.8), in which the integrability in time is L^2. Here we can take $q > 2$ to get $1 \leq q' < 2$.

Finally, we introduce the adjoint operator

$$\delta\colon L^2(\Omega\times[0,T];\mathcal{L}_2(U;H)) \supset \mathcal{D}(\delta) \to L^2(\Omega;H),$$

of the unbounded operator $D\colon L^2(\Omega;H) \to L^2(\Omega\times[0,T];\mathcal{L}_2(U;H))$. It is defined by

$$\big\langle DY, \Phi \big\rangle_{L^2(\Omega\times[0,T];\mathcal{L}_2(U;H))} = \big\langle Y, \delta\Phi \big\rangle_{L^2(\Omega;H)}. \tag{2.14}$$

Theorem 4.13 in Kruse [40] states that $L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H)) \subset \mathcal{D}(\delta)$ and that for all $\Phi \in L^2_{\mathcal{F}}([0,T]\times\Omega;\mathcal{L}_2(U;H))$ it coincides with the Itō integral

$$\delta(\Phi) = \int_0^T \Phi_t \,\mathrm{d}W_t.$$

With this knowledge, the duality between D and δ in (2.14) is precisely the integration by parts formula (2.10). The operator δ is also called the Skorohod integral.

3. Deterministic evolution equations

Semigroup theory allows us to consider many parabolic and hyperbolic partial differential equations as infinite dimensional ordinary differential equations. Different equations require different types of semigroups. Throughout this thesis we consider parabolic equations, which require analytic semigroups, and also Volterra integro-differential equations, which essentially are treated within the same framework, but with a solution operator family which is not a semigroup. In Papers I–II, we consider, from a semigroup theoretical point of view, a simple setting, where the semigroup can be defined via a spectral decomposition. In Paper III, Volterra integro-differential equations are considered and in Paper IV we allow general analytic semigroups. We limit the presentation in this introduction to the setting of the Papers I–III. For semigroup theory we recommend Pazy [49], Lunardi [45] and for Volterra equations Prüss [53].

3.1. Analytic semigroups generated by selfadjoint operators. Let H be the Hilbert space from the previous sections, and let $\mathcal{L}(H)$ denote the space of all bounded linear operators on H. We consider an operator $A\colon H \supset \mathcal{D}(A) \to H$, which is selfadjoint, positive definite, and with compact inverse. These conditions ensure that there exists eigenpairs $(\lambda_n, \phi_n)_{n\in\mathbf{N}}$, such that $A\phi_n = \lambda_n\phi_n$, $n \in \mathbf{N}$, and such that $(\phi_n)_{n\in\mathbf{N}} \subset H$ forms an ON-basis, and such that $\lambda_n \to \infty$. We order the eigenvalues in increasing order, i.e., $0 < \lambda_1 \le \dots \le \lambda_n \le \lambda_{n+1} \le \dots$, $n \in \mathbf{N}$.

The analytic semigroup $S_t = e^{-tA}$, generated by $-A$, is defined as the strong operator limit

$$S_t = \sum_{n\in\mathbf{N}} e^{-\lambda_n t}\phi_n \otimes \phi_n, \quad t \ge 0.$$

It has the semigroup property

$$S_s \circ S_t = S_{s+t}, \quad s,t \ge 0, \tag{3.1}$$

$$S_0 = \mathrm{id}_H, \tag{3.2}$$

$$t \mapsto S_t \text{ is strongly continuous.} \tag{3.3}$$

Any operator family $(S_t)_{t\ge0} \subset \mathcal{L}(H)$, which satisfies properties (3.1)–(3.3) is called an operator semigroup. The particular semigroup $(S_t)_{t\ge0}$ has an additional very good property, namely it is analytic. This means that it extends to an analytic function, in a sector of the complex plane, containing the positive real line. From our point of view the most important properties of analytic semigroups are the smoothing property and Hölder estimate in (3.5) below.

In order to proceed we define fractional powers of the operator A. For $r \in \mathbf{R}$, let $A^r \colon H \subset \mathcal{D}(A^r) \to H$, be the operator which is given by the strong operator limit

$$A^r = \sum_{n\in\mathbf{N}} \lambda_n^r \phi_n \otimes \phi_n,$$

with

$$\mathcal{D}(A^r) = \begin{cases} h \in H : \sum_{n\in\mathbf{N}} \lambda_n^{2r} |\langle \phi_n, h\rangle|^2 < \infty, & r > 0, \\ H, & r \le 0. \end{cases}$$

For $r \geq 0$ let H_r denote the space $H_r = \mathcal{D}(A^r)$, equipped with the norm

$$\|h\|_{H_r} = \left\|A^r h\right\|, \quad h \in H_r. \tag{3.4}$$

For $r < 0$, let H_r be the closure or H under the norm (3.4). Since $\mathbf{R}^+ \ni x \mapsto x^{2r}e^{-2x}$ is a bounded function for all $r \geq 0$ it holds by Parseval's identity that

$$\begin{aligned} \left\|A^r S_t h\right\|^2 &= \sum_{n\in\mathbf{N}} \lambda_n^{2r} e^{-2\lambda_n t} |\langle \phi_n, h\rangle|^2 = t^{-2r} \sum_{n\in\mathbf{N}} (t\lambda_n)^{2r} e^{-2\lambda_n t} |\langle \phi_n, h\rangle|^2 \\ &\le C_r t^{-2r} \sum_{n\in\mathbf{N}} |\langle \phi_n, h\rangle|^2 = C_r t^{-2r} \|h\|^2. \end{aligned}$$

It also holds, since $\mathbf{R}^+ \ni x \mapsto x^{-r}(e^{-x} - 1)$ is bounded, that for all $r \in [0,1]$ and $t > 0$,

$$\begin{aligned} \left\|A^{-r}(S_t - \mathrm{id}_H)h\right\|^2 &= \sum_{n\in\mathbf{N}} \lambda_n^{-2r} (e^{-\lambda_n t} - 1)^2 |\langle \phi_n, h\rangle|^2 \\ &= t^{2r} \sum_{n\in\mathbf{N}} (\lambda_n t)^{-2r} (e^{-\lambda_n t} - 1)^2 |\langle \phi_n, h\rangle|^2 \\ &\le C_r t^{2r} \sum_{n\in\mathbf{N}} |\langle \phi_n, h\rangle|^2 = C t^{2r} \|h\|^2. \end{aligned}$$

We restate these two assertions:

$$\begin{aligned} \left\|A^r S_t\right\|_{\mathcal{L}(H)} &\le C_r t^{-r}, \quad t > 0,\ r \ge 0, \\ \left\|A^{-r}(S_t - \mathrm{id}_H)\right\|_{\mathcal{L}(H)} &\le C_r t^r,\ t > 0,\ r \in [0,1]. \end{aligned} \tag{3.5}$$

It is clear that any power A^r, $r \in \mathbf{R}$, commutes with the semigroup S, i.e., for all $h \in H_r$ it holds $S_t A^r h = A^r S_t h$. These are essentially the properties of S, which will be used in this thesis.

3.2. Cauchy problems. One property of S, which we did not mention above, is that $t \mapsto S_t$ is strongly differentiable and that

$$\frac{\mathrm{d}}{\mathrm{d}t} S_t h + A S_t h = 0, \quad t > 0; \quad h \in H.$$

Since $S_0 = \mathrm{id}_H$, it is clear that $u(t,x) = S_t x$ is the solution to the homogenous Cauchy problem

$$\dot{u} + Au = 0, \quad t > 0; \quad u_0 = x.$$

The solution u to the inhomogeneous Cauchy problem

$$\dot{u} + Au = f, \quad t > 0; \quad u_0 = x,$$

where $f\colon [0,T] \to H$ is sufficiently regular, is given by the variation of constants formula, or Duhamel's principle, which reads

$$u_t = S_t x + \int_0^t S_{t-s} f_s \,\mathrm{d}s, \quad t \in [0,T]. \tag{3.6}$$

In this thesis we will consider this type of problems with f depending in a nonlinear way on the solution, and with an additional stochastic term in the right hand side of the equation. The solution in (3.6) called a mild solution.

3.3. Volterra integro-differential equations. Let $b\colon (0,\infty) \to \mathbf{R}$ be the Riesz kernel $b_t = t^{\rho-2}/\Gamma(\rho-1)$, where $\rho \in (1,2)$ is some fixed number. We consider first the linear homogenous equation

$$\dot{u} + \int_0^t b_{t-s} A u_s \,\mathrm{d}s = 0,\ t > 0; \quad u_0 = x.$$

The solution operator $(\mathcal{S}_t)_{t\geq 0} \subset \mathcal{L}(H)$ to this equation, is given by the strong operator limit

$$\mathcal{S}_t = \sum_{n\in\mathbf{N}} s_{n,t}\,(\phi_n \otimes \phi_n), \quad t \geq 0,$$

where $s_{n,t}$, is the solution to the scalar equation

$$\dot{s}_{n,t} + \lambda_n \int_0^t b(t-r) s_{n,r} \,\mathrm{d}r = 0,\ t > 0; \quad s_{n,0} = 1.$$

This operator family does not satisfy $\mathcal{S}_s \circ \mathcal{S}_t = \mathcal{S}_{s+t}$ and is therefore no semigroup. Nevertheless, the solution of the inhomogeneous equation

$$\dot{u} + \int_0^t b_{t-s} A u_s \,\mathrm{d}s = f,\ t > 0; \quad u_0 = x,$$

is given by the mild solution

$$u_t = \mathcal{S}_t + \int_0^t \mathcal{S}_{t-s} f_s \,\mathrm{d}s, \quad t \geq 0,$$

which looks formally the same as (3.6).

Moreover, the family satisfies bounds analogous to (3.5) but modified by the parameter ρ. For example, we have the smoothing property

$$\left\| A^{\frac{r}{\rho}} \mathcal{S}_t \right\|_{\mathcal{L}(H)} \leq C_r t^{-r}, \quad t > 0,\ r \in [0,1],$$

and other bounds which are used in the analysis.

3.4. Finite element approximation. Here we treat a concrete partial differential equation. There is a rich literature on the finite element method, see Brenner & Scott [10] for elliptic problems and Thomée [54] for parabolic problems. In this thesis we apply existing results, for the most basic finite element approximation, and the only new results we use are obtained by interpolation between known results, see Papers I–III.

We consider $\mathrm{D} \subset \mathbf{R}^d$, $d = 1,2,3$, a convex, polygonal domain, and $H = L^2(\mathrm{D})$. Let $A = -\Delta$, where $\Delta = \sum_{i=1}^d \partial^2/\partial\xi_i^2$ is the Laplace operator with homogeneous Dirichlet boundary condition, i.e., $\mathcal{D}(A) = H_0^1(\mathrm{D}) \cap H^2(\mathrm{D})$. The operator A satisfies all assumptions of Section 3.1 and generates therefore an analytic semigroup $(S_t)_{t\geq 0}$. Let $(\mathcal{T}_h)_{h\in(0,1)}$ denote a regular family of triangulations of D. Here h is a refinement parameter which is the diameter of the largest triangle in the mesh. Let $(V_h)_{h\in(0,1)}$ denote the corresponding family of spaces $V_h \subset H$, which consists of continuous functions on D being affine linear on each triangle. We define $P_h\colon H \to V_h$ to be the orthogonal projector onto V_h. In finite element theory the Ritz projector $R_h\colon H_{1/2} \to V_h$ is also important.

Let $A_h\colon V_h \to V_h$ denote the discrete Laplacian, which is the operator on V_h satisfying

$$\langle A_h \phi_h, \psi_h \rangle = \langle \nabla\phi_h, \nabla\psi_h \rangle, \quad \forall \phi_h, \psi_h \in V_h.$$

The operator A_h is selfadjoint and positive definite. It therefore generates an analytic semigroup $(S_t^h)_{t\geq 0} \subset \mathcal{L}(V_h)$, which is the solution operator to the Cauchy problem

$$\dot{u}_h + A_h u_h = 0,\ t > 0; \quad u_{h,0} = P_h x.$$

The semigroup is analytic uniformly in h in the sense that the characteristic smoothing property analogous to (3.5) holds uniformly in h, namely

$$\left\| A_h^r S_{h,t} \right\|_{\mathcal{L}(H)} \leq C_r t^{-r}, \quad t > 0, \quad h \in (0,1), \quad r \geq 0.$$

The following error estimates holds for the projectors and for the approximation of the semigroup:

$$\begin{aligned}
\left\|A^{\frac{s}{2}}(\mathrm{id}_H - P_h)\phi\right\| &\leq \ Ch^{r-s}\left\|A^{\frac{r}{2}}\phi\right\|, \quad s \in [0,1],\ r \in [s,2],\\
\left\|A^{\frac{s}{2}}(\mathrm{id}_H - R_h)\phi\right\| &\leq \ Ch^{r-s}\left\|A^{\frac{r}{2}}\phi\right\|, \quad s \in [0,1],\ r \in [1,2],\\
\left\|\left(S_t - S_t^h\right)\phi\right\| &\leq \ Ct^{-\frac{s-r}{2}}h^s\|A^{\frac{r}{2}}\phi\|, \quad s \in [0,2],\ r \in [0,s].
\end{aligned}$$

Recall that $h = h_{\max}$ is the largest diameter of any triangle in $\mathcal{T}_h$. Let $h_{\min}$ be the diameter of the smallest triangle in $\mathcal{T}_h$. The family $(\mathcal{T}_h)_{h\in(0,1)}$ is said to be quasi-uniform, if there exists a number ρ, such that

$$\frac{h_{\max}}{h_{\min}} \leq \rho, \quad \forall \mathcal{T}_h \in (\mathcal{T}_h)_{h\in(0,1)}.$$

If the mesh family is quasi-uniform, then the following estimates hold

$$\left\|A_h^{\frac{1}{2}}P_h\phi\right\| \leq C\left\|A^{\frac{1}{2}}\phi\right\|,\ \phi \in H_{1/2}; \quad \left\|A_hP_h\right\|_{\mathcal{L}(H)} \leq Ch^{-2}.$$

In Paper I these estimates are used, enforcing us to assume quasi-uniformity. In Papers II–III this restriction is removed.

3.5. Full approximation. Above we described two ways to discretize space. We now consider full discretization with finite element approximation in space and the Backward Euler method for approximation in time. Let $N \in \mathbf{N}$, $k = T/N$, and $0 = t_0 < t_1 < \cdots < t_N = T$ be a uniform grid with $t_j = jk$, $j \in \{0,\dots,N\}$. The fully discrete scheme reads, in abstract form,

$$\frac{U_n^{h,k} - U_{n-1}^{h,k}}{k} + A_hU_n^{h,k} = 0, \quad n \in \{1,\dots,N\}; \quad U_0^{h,k} = P_hx,$$

or rewritten and iterated

$$U_n^{h,k} = (\mathrm{id}_H + kA_h)^{-1}U_{n-1}^{h,k} = \cdots = (\mathrm{id}_H + kA_h)^{-n}P_nx =: S_n^{h,k}x.$$

The family $(S_n^{h,k})_{n\in\mathbf{N}}$ is a fully discrete approximation of the semigroup $(S_t)_{t\geq0}$. The error and stability estimates holds for $s \in [0,2]$, $r \in [0,s]$,

$$\begin{aligned}
\left\|\left(S_{t_n} - S_n^{h,k}\right)\phi\right\| &\leq Ct_n^{-\frac{s-r}{2}}\left(h^s + k^{\frac{s}{2}}\right)\left\|A^{\frac{r}{2}}\phi\right\|, \quad n \in \{1,2,\dots\},\\
\left\|A_h^{\frac{s}{2}}S_n^{h,k}\phi\right\| &\leq Ct_n^{-\frac{s}{2}}\|\phi\|, \quad n \in \{1,2,\dots\}.
\end{aligned}$$

4. Stochastic evolution equations

The main topic of this thesis is the study of stochastic evolution equations (SEE) in Hilbert space, treated within the semigroup framework. In Papers I–III we consider a well established setting, in which we can rely on existing results on existence, uniqueness and regularity, see Baeumer et al. [4], Brzeźniak [11], Da Prato & Zabczyk [17], Jentzen & Röckner [31], van Neerven [55]. For regularity in the Malliavin sense we rely on Fuhrman & Tessitore [22], but in all of Papers I–III we prove refined results, which we need. In Paper IV we study Markov theory for SEE, in particular, we study smoothness properties of the transition semigroup and the Kolmogorov equation. For this purpose we need, as a technical tool, to consider SEE with initial values in spaces $H_{-\delta}$, for $\delta \in [0,1/2)$. This was studied in Chen & Dalang [15], [14] for the heat equation, on the real line in the framework of Walsh [56]. In the semigroup framework on the hand no such results were previously available in the literature, and establishing existence and uniqueness is one of the purposes of this thesis.

4.1. SEE with irregular initial value. In Paper IV, Section 2, we consider consider equations of the following type

$$\mathbf{X}_t = S_t \xi + \int_0^t S_{t-s}\mathbf{F}(s,\mathbf{X}_s)\,\mathrm{d}s + \int_0^t S_{t-s}\mathbf{B}(s,\mathbf{X}_s)\,\mathrm{d}W_s, \quad t \in [0,T]. \tag{4.1}$$

Here $(S_t)_{t\geq 0}$ is an analytic semigroup and W is a cylindrical id_U-Wiener process. We assume that $\mathbf{F}\colon (0,T]\times H\times\Omega \to \mathcal{H}_1$, and $\mathbf{B}\colon (0,T]\times H\times\Omega \to \mathcal{L}_2(U;\mathcal{H}_2)$ are predictable and globally Lipschitz continuous in a suitable sense. Here $\mathcal{H}_1 \supset H$ and $\mathcal{H}_2 \supset H$ are continuous, and, unless $\mathcal{H}_2 = H$, the noise is not of trace class. We allow initial singularities in $\mathbf{F}$ and $\mathbf{B}$, which is captured by the following assumptions,

$$\|\mathbf{F}(t,0)\|_{L^p(\Omega;\mathcal{H}_1)} \leq Ct^{-\hat{\alpha}}, \quad \|\mathbf{B}(t,0)\|_{L^p(\Omega;\mathcal{H}_2)} \leq Ct^{-\hat{\beta}}, \quad t \in (0,T],$$

for some $\hat{\alpha} \in [0,1)$, and $\hat{\beta} \in [0,1/2)$. What is most interesting is the assumption on ξ. We assume that, for some $p \in [2,\infty)$,

$$\xi \in L^p(\Omega;H_{-\delta}) \quad \text{with} \quad \begin{cases} \delta \in [0,1), & \text{if the noise is additive,} \\ \delta \in [0,1/2), & \text{otherwise.} \end{cases}$$

In Theorem 2.7 in Paper IV, we show the following

THEOREM 4.1. *Under the above assumptions, there exist an up to modification unique stochastic process* $\mathbf{X}\colon [0,T]\times\Omega \to H_{-\delta}$*, which satisfy* (4.1)*, and* $\mathbf{X}_t \in H$*,* $t \in (0,T]$ $\mathbf{P}$*-a.s., and moreover*

$$\sup_{t\in(0,T]} t^{\lambda}\|\mathbf{X}_t\|_{L^p(\Omega;H)} \leq C\Big(1 + \|\xi\|_{L^p(\Omega;H_{-\delta})}\Big),$$

where $\lambda \geq 0$, depends on δ, the strengths of the singularities of $\mathbf{F}$ *and* $\mathbf{B}$*, and on* $\mathcal{H}_1$*,* $\mathcal{H}_2$*.*

The proof is performed by a classical contraction argument, using Banach's fixed point theorem. More precisely, $\mathbf{X}$ is shown to be the unique fixed point of the mapping

$$\Phi(\mathbf{Y}) = \left(S_t \xi + \int_0^t S_{t-s} \mathbf{F}(s, \mathbf{X}_s) \, \mathrm{d}s + \int_0^t S_{t-s} \mathbf{B}(s, \mathbf{X}_s) \, \mathrm{d}W_s \right)_{t \in [0,T]}, \tag{4.2}$$

defined on the Banach space $\mathbb{L}^p_{\delta,\lambda}$ of predictable stochastic processes $\mathbf{Y}\colon [0,T] \times \Omega \to H_{-\delta}$, such that

$$\|\mathbf{Y}\|_{\mathbb{L}^p_{\lambda,r}} := \sup_{t \in (0,T]} t^\lambda e^{rt} \|\mathbf{Y}_t\|_{L^p(\Omega;H)} < \infty.$$

For $r \in (-\infty, 0)$ with $|r|$ sufficiently large, this map is shown to be a contraction.

4.2. SEE with smooth coefficients. Here we consider the following equations

$$X_t^x = S_t x + \int_0^t S_{t-s} F(X_s^x) \, \mathrm{d}s + \int_0^t S_{t-s} B(X_s^x) \, \mathrm{d}W_t, \quad t \in [0,T], \tag{4.3}$$

being indexed over the initial value $x \in \Xi$, where Ξ is the union of all spaces $H_{-\delta}$, $\delta \geq 0$, for which (4.3) has a solution. For fixed $n \in \mathbf{N}$ we assume that $F \in \mathcal{C}^n_{\mathrm{b}}(H; \mathcal{H}_1)$ and $B \in \mathcal{C}^n_{\mathrm{b}}(H; \mathcal{L}_2(U; \mathcal{H}_2))$.

In Paper IV, Theorem 3.1, we prove that $x \mapsto X^x$ is Fréchet differentiable from negative order spaces. One feature of this result is that that there exists $\delta > 0$, such that $H_{-\delta/k} \ni x \mapsto X^x$ is k times Fréchet differentiable, for $k \in \{1, \ldots, n\}$. Thus for higher order derivatives, Fréchet differentiability holds only on smaller and smaller spaces.

Let $(\mathcal{P}_t)_{t \in (0,T]}$ denote the family of mappings which, for $t \in (0,T]$ act on $\varphi \in \mathcal{C}^1_{\mathrm{b}}(H; \mathbf{R})$ by

$$(\mathcal{P}_t \varphi)(x) := \mathbf{E}\Big[\varphi(X_t^x)\Big].$$

Since X^x, is well defined for irregular $x \in \Xi$, and since $X_t^x \in H$, for $t \in (0,T]$, $x \in \Xi$, it holds that $\Xi \ni x \mapsto (\mathcal{P}_t \varphi)(x) \in \mathbf{R}$ is well defined. We call $(\mathcal{P}_t)_{t \in (0,T]}$ the extended transition semigroup. In Paper IV, Theorem 3.2, we show, in particular, that there exists $\delta > 0$, such that $H_{-\delta/k} \ni x \mapsto (\mathcal{P}_t \varphi)(x) \in \mathbf{R}$ is k times Fréchet differentiable, for $k \in \{1, \ldots, n\}$, and moreover that for all $\delta_1, \ldots, \delta_n \in [0, \delta)$ with $\delta_1 + \cdots + \delta_n < \delta$ it holds

$$\Big| \big((\mathcal{P}_t \varphi) \big)^{(n)}(x) \big(u_1, \ldots, u_n \big) \Big| \leq C t^{-(\delta_1 + \cdots + \delta_n)} \|u_1\|_{H_{-\delta_1}} \cdots \|u_n\|_{H_{-\delta_n}}.$$

This is a useful result, which allows one to distribute smoothness onto $u_1, \ldots, u_n$ in an asymmetric way. This is one of the main results in Paper IV. This result should be compared with [3, (4.2)–(4.3)], [8, Lemma 5.3], [20, Lemma 4.4–4.6], [32, Chapter 5, Proposition 7.1], [59, Lemma 3.3], but all these results restrict to a finite dimensional setting. To the best of our knowledge Debussche [20] is the first paper containing this kind of bounds, and [20] was the inspiration for Paper IV.

For $\varphi \in \mathcal{C}_{\mathrm{b}}^2(H;\mathbf{R})$, $F \in \mathcal{C}_{\mathrm{b}}^2(H;\mathcal{H}_1)$, $B \in \mathcal{C}_{\mathrm{b}}^2(H;\mathcal{L}_2(U;\mathcal{H}_2))$ consider the Kolmogorov equation

$$\begin{aligned}\frac{\partial}{\partial t}u(t,x) &= \frac{\partial}{\partial x}u(t,x)\Big(-Ax+F(x)\Big)+\frac{1}{2}\sum_{v\in\mathbb{U}}\frac{\partial^2}{\partial x^2}u(t,x)\Big(B(x)v,B(x)v\Big),\\ u(0,x) &= \varphi(x).\end{aligned}$$

In Paper IV, Theorem 4.1 we prove that for all $\varphi \in \mathcal{C}_{\mathrm{b}}^2(H;\mathbf{R})$, $t \in (0,T]$, $x \in H_1$, the function $u(t,x) := (\mathcal{P}_t\varphi)(x)$ satisfies the Kolmogorov equation.

This result extends [18, Theorem 7.5.1] in the case when $-A$ generates an analytic semigroup, which in fact is required in order to have a solution of the stochastic equation for $\mathcal{H}_1 \supsetneq H$ or $\mathcal{H}_2 \supsetneq H$. While we assume $F \in \mathcal{C}_{\mathrm{b}}^2(H;\mathcal{H}_1)$, $B \in \mathcal{C}_{\mathrm{b}}^2(H;\mathcal{L}_2(U;\mathcal{H}_2))$, they assume $F \in \mathcal{C}_{\mathrm{b}}^3(H;H)$, $B \in \mathcal{C}_{\mathrm{b}}^3(H;\mathcal{L}_2(U;H))$. We also remark that our result in fact does not require $x \in H_1$, in order for $(t,x) \mapsto (\mathcal{P}_t\varphi)(x)$ to satisfy the Kolmogorov equation, but less regular x are allowed. In all other works we are aware of, $x \in H_1$ is assumed.

5. SPDE and stochastic Volterra equations

Here we consider concrete settings, to which the results of the previous section apply. First we discuss stochastic partial differential equations and second we discuss stochastic Volterra integro-differential equations. For more about concrete settings see Jentzen & Kloeden [30], Jentzen & Röckner [31], Jentzen [29], van Neerven [55].

5.1. Stochastic reaction-diffusion equations. Let $\mathrm{D} \subset \mathbf{R}^d$, $d = 1,2,3$, be a convex polygonal domain and let $H = L^2(\mathrm{D})$. The linear operator $A\colon H \supset \mathcal{D}(A) \to H$ is chosen to be $A = -\Delta$, where $= \sum_{i=1}^d \partial^2/\partial\xi^2$ is the Laplace operator with homogeneous Dirichlet boundary condition, i.e., $\mathcal{D}(A) = H^2(\mathrm{D}) \cap H_0^1(\mathrm{D})$. Due to the concrete setting we prefer to work with the notation $\dot{H}^r = H_{r/2}$, where $(H_r)_{r\in\mathbf{R}}$, are the spaces introduced in Section 3 corresponding to the operator A. With this notation $\dot{H}^r$ coincides with the classical Sobolev spaces $W^{r,2}(\mathrm{D})$ with certain boundary conditions depending on r.

The nonlinear drift $F\colon H \to H$ is a Nemytskii operator, defined by $(F(x))(\xi) = f(x(\xi))$, for $x \in H$, $\xi \in \mathrm{D}$, and some $f\colon \mathbf{R} \to \mathbf{R}$, which is globally Lipschitz continuous or more regular. Under this assumption the mapping F is globally Lipschitz continuous as well.

Let $Q \in \mathcal{L}(H)$ be selfadjoint, positive definite, not necessarily of finite trace. The Hilbert space U is here given as the image $U = Q^{\frac{1}{2}}(H)$ of H under the unique positive square root $Q^{\frac{1}{2}}$ of Q. It is equipped with the scalar product $\langle u,v\rangle = \langle Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}v\rangle$, where $Q^{-\frac{1}{2}}$ is the pseudoinverse of $Q^{\frac{1}{2}}$. Let $\beta \in (0,1]$ be a regularity parameter. The multiplicative noise coefficient $B\colon H \to \mathcal{L}_2(U;\dot{H}_{\beta-1})$ is a Nemytskii operator, defined by $(B(x)u)(\xi) = b(x(\xi))u(\xi)$, for $x \in H$, $u \in U$, $\xi \in \mathrm{D}$, and some $b\colon \mathbf{R} \to \mathbf{R}$, being globally Lipschitz continuous. Under these assumptions it is not clear that B is well defined, but for different choices of U, b, β, one has to check if $B(x) \in \mathcal{L}_2(U;\dot{H}_{\beta-1})$, for all $x \in H$, and moreover if $x \mapsto B(x)$ is Lipschitz continuous.

Example 5.1 (Linear multiplicative noise). Assume that $d = 1$, $\mathrm{D} = [0,1]$, $Q = \mathrm{id}_H$, $U = H$, $\beta \in (0,1/2)$, and that $(b(x))(\xi) = x(\xi)$, for $\xi \in [0,1]$. Let $(\phi_i,\lambda_i)_{i\in\mathbf{N}}$ denote the eigenpairs of A. We get

$$\begin{aligned}\big\|B(x)\big\|^2_{\mathcal{L}_2(U;\dot{H}^{\beta-1})} &= \sum_{i\in\mathbf{N}} \big\|B(x)\phi_i\big\|^2_{\dot{H}^{\beta-1}} = \sum_{i,j\in\mathbf{N}} \Big|\Big\langle A^{\frac{\beta-1}{2}}B(x)\phi_i,\phi_j\Big\rangle\Big|^2\\ &= \sum_{i,j\in\mathbf{N}} \Big|\Big\langle B(x)\phi_i, A^{\frac{\beta-1}{2}}\phi_j\Big\rangle\Big|^2 = \sum_{i,j\in\mathbf{N}} \lambda_j^{\beta-1}\Big|\Big\langle B(x)\phi_i,\phi_j\Big\rangle\Big|^2\\ &\le \sup_{n\in\mathbf{N}}\sup_{\xi\in[0,1]} |\phi_n(\xi)|^2 \sum_{i,j\in\mathbf{N}} \lambda_j^{\beta-1}\Big|\Big\langle x,\phi_i\Big\rangle\Big|^2 = C\big\|A^{\frac{\beta-1}{2}}\big\|^2_{\mathcal{L}_2(H)}\|x\|^2.\end{aligned}$$

Since B is linear the same calculation with $B(x)$ replaced by $B(x)-B(y)$, shows that $H \ni x \mapsto B(x) \in \mathcal{L}_2(H;\dot{H}^{\beta-1})$ is Lipschitz continuous. This calculation is taken from Jentzen [29, § 5.2.1]

Example 5.2 (Additive space-time white noise). Assume that $d = 1$, $\mathrm{D} = [0,1]$, $Q = \mathrm{id}_H$, $U = H$, $\beta \in (0,1/2)$, and that $b = 1$. Since for all $\gamma > 1/2$, it holds that $\|A^{-\gamma/2}\|_{\mathcal{L}_2(H)} < \infty$, we have

$$\big\|B\big\|_{\mathcal{L}_2(H;\dot{H}^{\beta-1})} = \big\|A^{\frac{\beta-1}{2}}\big\|_{\mathcal{L}_2(H)} < \infty.$$

Example 5.3 (Additive trace class noise). Assume $d = 1,2,3$, $\mathrm{Tr}(Q) < \infty$, $\beta = 1$, and $b = 1$. Then

$$\big\|B\big\|_{\mathcal{L}_2(U;H)} = \big\|BQ^{\frac{1}{2}}\big\|_{\mathcal{L}_2(H)} = \big\|Q^{\frac{1}{2}}\big\|_{\mathcal{L}_2(H)} = \sqrt{\mathrm{Tr}(Q)} < \infty.$$

Example 5.4. In Paper I and in Debussche [20] it is assumed, up to a unnatural nonlinear perturbation term, that $B(x) = B_1 x + B_2$, where $B_1 \in \mathcal{L}(H;\mathcal{L}(H))$ and $B_2 \in \mathcal{L}(H)$. This assumption is not satisfactory if we want to consider Nemytskii operators. Let $d = 1$, $\mathrm{D} = [0,1]$, $Q = \mathrm{id}_H$, $U = H$, $\beta \in (0,1/2)$, $b = 1$. Then it holds

$$\big\|B(x)\phi\big\| = \Big(\int_0^1 \big|x(\xi)\phi(\xi)\big|^2 \,\mathrm{d}\xi \Big)^{\frac{1}{2}},$$

and taking, for instance $x, \phi \in H$ given by $x(\xi) = \phi(\xi) = \xi^{-3/8}$ yields $\|B(x)\phi\| = \infty$ and this shows that $B \notin \mathcal{L}(H;\mathcal{L}(H))$.

We end this subsection with a discussion about derivatives of F and B, given smooth f, b. If f is continuously differentiable, then $(F'(x)\phi)(\xi) = f'(x(\xi))\phi(\xi)$, for $\xi \in \mathrm{D}$, $x, \phi \in H$, and since f' is bounded it holds that

$$\|F'(x)\phi\| \leq \sup_{y\in\mathbf{R}} |f'(y)|\,\|\phi\|.$$

This means that F is Fréchet differentiable. On the other hand, if f is twice continuously differentiable, then the second derivative

$$(F''(x)(\phi,\psi))(\xi) = f''(x(\xi))\phi(\xi)\psi(\xi), \quad \xi \in \mathrm{D}, \quad x, \phi, \psi \in H,$$

is not a Fréchet derivative since by the Cauchy-Schwarz inequality we get no better estimate than

$$\begin{aligned} \|F''(x)(\phi,\psi)\| &= \Big(\int_{\mathrm{D}} |f''(x(\xi))\phi(\xi)\psi(\xi)|^2 \,\mathrm{d}\xi \Big)^{\frac{1}{2}} \\ &\leq \sup_{y\in\mathbf{R}} |f''(y)| \|\phi\|_{L^4(\mathrm{D})} \|\psi\|_{L^4(\mathrm{D})}. \end{aligned}$$

But, by using the Sobolev embedding theorem one can show that for all $\gamma > d/2$, the embedding $L^1(\mathrm{D}) \subset \dot{H}^{-\gamma}$ is continuous. Therefore

$$\begin{aligned} \|F''(x)(\phi,\psi)\|_{\dot{H}^{-\gamma}} \leq C\|F''(x)(\phi,\psi)\|_{L^1(\mathrm{D})} &= \int_{\mathrm{D}} |f''(x(\xi))\phi(\xi)\psi(\xi)| \,\mathrm{d}\xi \\ &\leq \sup_{y\in\mathbf{R}} |f''(y)|\, \|\phi\|_{L^2(\mathrm{D})} \|\psi\|_{L^2(\mathrm{D})}. \end{aligned}$$

This means that $F\colon H \to \dot{H}^{-\gamma}$ is twice Fréchet differentiable for all $\gamma > d/2$. Therefore, in order to include this type of drift terms, in Papers II–III we consider the assumption that $F\colon H \to H$ is once Fréchet differentiable and $F\colon H \to \dot{H}^{-\gamma}$ is twice Fréchet differentiable, for some γ. In Paper I we assume $F\colon H \to H$ to be twice Fréchet differentiable, and this forces $F'' = 0$, or otherwise that F is something more abstract, and less interesting, than a reaction term.

For the mapping B, we proceed with an example.

EXAMPLE 5.5. Consider the setting of Example 5.2. Then $((B'(x)\phi)u)(\xi) = u(\xi)\phi(\xi) = (B(\phi)u)(\xi)$, for $\xi \in \mathrm{D}$, $u, x, \phi \in H$. Thus by Example 5.1 we get that

$$\|B'(x)\phi\|_{\mathcal{L}_2(U;\dot{H}^{\beta-1})} = \|B(\phi)\|_{\mathcal{L}_2(U;\dot{H}^{\beta-1})} \leq C\|\phi\|.$$

This proves that $B\colon H \to \mathcal{L}_2(U;\dot{H}^{\beta-1})$ is Fréchet differentiable.

If B, in the example, instead was defined by a continuously differentiable $b\colon \mathbf{R} \to \mathbf{R}$, then also this B would be Fréchet differentiable. The reason why we consider linear or constant B in the examples is that we need the second derivative B'' in our analysis. No use of the Sobolev embedding theorem can prove such B to be twice Fréchet differentiable. Therefore we need $B'' = 0$.

5.2. Stochastic Volterra integro-differential equations. Here we continue with the setting of the previous subsection and let B be defined by $b = 1$, i.e., we consider additive noise. We consider the equation

$$X_t = \mathcal{S}_t x + \int_0^t \mathcal{S}_{t-s} F(X_s)\,\mathrm{d}s + \int_0^t \mathcal{S}_{t-s}\,\mathrm{d}W_s$$

where we recall from Subsection 3.3 that $(\mathcal{S}_t)_{t\geq 0}$ is the solution operator to the linear deterministic equation

$$u_t + \int_0^t b_{t-s} A u_s\,\mathrm{d}s = 0, \quad t > 0; \quad u_0 = x,$$

in the sense $u_t = \mathcal{S}_t x$, $t \geq 0$. Existence and uniqueness of this type of equations is proved in [4]. Malliavin regularity is proved in Paper III.

6. Approximation by the finite element method

In this section approximation schemes for stochastic partial differential equations and stochastic Volterra integro-differential equations are introduced. We consider the concrete setting of the previous section, but we do not discuss the weak formulations of the equations, which would be the starting point for implementation. We therefore keep the presentations, still, on a rather abstract level, as we do in Papers I–III.

6.1. Stochastic partial differential equations. Consider the setting of Subsection 3.4. Let X be the solution to (4.3) under the setting of Section 5.1. We first consider semidiscretization in space. The finite element approximations $(X^h)_{h\in(0,1)}$, corresponding to the family $(\mathcal{T}_h)_{h\in(0,1)}$, are the solutions to the equations

$$X_t^h = S_{h,t} P_h x + \int_0^t S_{h,t-s} P_h F(X_s^h)\,\mathrm{d}s + \int_0^t S_{h,t-s} P_h B(X_s^h)\,\mathrm{d}W_s.$$

Recall that $B\colon H \to \mathcal{L}_2(U;\dot{H}^{\beta-1})$, for some $\beta \in (0,1]$. It is well known, that for $\gamma \in [0,\beta)$, and $x \in L^p(\Omega;\dot{H}^\gamma)$, it holds

$$\sup_{t\in[0,T]} \left\| X_t - X_t^h \right\|_{L^p(\Omega;H)} \leq C h^\gamma, \quad h \in (0,1).$$

For $\beta = 1$, and in fact $\gamma = \beta$, this is proved in Kruse [39], and for $\beta \in (0,1)$, to the best of our knowledge, no proof is available in the literature, except for linear equations, see Kovács et al. [33].

We continue with full discretization and recall the notation of Subsection (3.5). We approximate X by a semi-implicit Euler-Maruyama method and finite element approximation in space:

$$\begin{aligned}&\frac{X_n^{h,k} - X_{n-1}^{h,k}}{k} + A_h X_n^{h,k} = kP_h F(X_{n-1}^{h,k}) + \int_{t_{n-1}}^{t_n} P_h B(X_{n-1}^{h,k})\,\mathrm{d}W_s, n \in \{1,\dots,N\},\\ &X_0^{h,k} = P_h x.\end{aligned}$$

Recalling $S_n^{h,k} = (\mathrm{id}_H + kA_h)^{-n}$ and $S^{h,k} = S_1^{h,k}$, one can rewrite this as

$$X_n^{h,k} = S^{h,k} X_{n-1}^{h,k} + kS^{h,k} P_h F(X_{n-1}^{h,k}) + \int_{t_{n-1}}^{t_n} S^{h,k} P_h B(X_{n-1}^{h,k})\,\mathrm{d}W_s.$$

Iteration of this equation yields

$$X_n^{h,k} = S_n^{h,k} X_{n-1}^{h,k} + k\sum_{j=0}^{n-1} S_{n-j}^{h,k} P_h F(X_j^{h,k}) + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} S_{n-j}^{h,k} B(X_j^{h,k})\,\mathrm{d}W_s. \tag{6.1}$$

Also for full discretization it is well known, that for $\gamma \in (0,\beta)$, and $x \in \dot{H}^{-\gamma}$ it holds

$$\sup_{n\in\{0,\dots,N_h\}} \left\|X_{t_n} - X_n^{h,k}\right\|_{L^p(\Omega;H)} \le C\left(h^\gamma + k^{\frac{\gamma}{2}}\right), \quad h,k \in (0,1).$$

For $\beta = 1$, and $\gamma = \beta$, this is proved in Kruse [39]. For $\beta \in (0,1)$, it is proved in Paper III, under the case of additive noise, i.e., for the case when B is constant.

6.2. Stochastic Volterra integro-differential equations. Consider the setting of Subsections 3.3 and 5.2. Recall that $b_t = t^{\rho-2}/\Gamma(\rho-1)$, $t > 0$, and that $\rho \in (0,1)$. Let $\hat{b}$ denote the Laplace transform of b and let $(\omega_j)_{j\in\mathbf{N}}$, be the weights which are determined by

$$\hat{b}\left(\frac{1-z}{k}\right) = \sum_{j=0}^{\infty} \omega_j^k z^j, \quad |z| < 1.$$

For the convolution we use the following approximation

$$\sum_{j=1}^{n} \omega_{n-j}^k f(t_j) \sim \int_0^{t_n} b(t_n - s) f(s)\,\mathrm{d}s, \quad f \in \mathcal{C}(0,T;\mathbf{R}),$$

see Lubich [43], [44]. To discretize the time derivative we use a backward Euler method, which is explicit in the semilinear term F. Our fully discrete scheme then reads:

$$\begin{aligned}
&X_{n+1}^{h,k} - X_n^{h,k} + k\sum_{j=1}^{n+1}\omega_{n+1-j}^{k} A_h X_j^{h,k} = kP_hF(X_n^{h,k}) + \int_{t_n}^{t_{n+1}} P_h\,\mathrm{d}W_t, \quad n \geq 0,\\
&X_0^{h,k} = P_h x_0.
\end{aligned}$$

It is possible to write $(X_n^{h,k})_{n=0}^{N}$ as a variation of constants formula (6.1). Indeed, it is shown in [37] that one has the explicit representation

$$B_n^{h,k} = \int_0^\infty S_{ks}^h P_h \frac{e^{-s}s^{n-1}}{(n-1)!}\,\mathrm{d}s, \quad n \geq 1,$$

where

$$S_t^h = \sum_{j=1}^{N_h} s_{j,t}^h\,(e_j^h \otimes e_j^h)P_h; \quad \dot{s}_{j,t}^h + \lambda_j^h \int_0^t b(t-r)s_{j,r}^h\,\mathrm{d}r = 0,\ t > 0; \quad s_{j,0}^h = 1,$$

and $(\lambda_j^h, e_j^h)_{j=1}^{N_h}$ are the eigenpairs corresponding to A_h.

7. Weak convergence

Weak convergence analysis for numerical approximation of equations with values in infinite-dimensional spaces is a rather young subject. The early papers but also subsequent papers have treated linear equations, see Debussche [21], Geissert et al. [23], Kovács et al. [34], [35], [36], Kovaćs & Printems [38], Kruse [40], Lindner & Schilling [42]. For linear parabolic and hyperbolic equations driven by Gaussian noise in Hilbert space, this theory is rather complete. New progress concerns linear equations driven by non-Gaussian noise, [36], [42], or linear Volterra type equations, see [38]. Much of the groundwork for treating more complicated equations is to be found in these papers, in particular concerning the finite element theory needed. Often, the required error estimates for solutions with low regularity are not available in the classical finite element literature.

Adding a nonlinear drift term increases the difficulty. Semilinear equations driven by additive noise are considered in Andersson et al. [1] (Paper III), [2] (Paper II), Andersson & Larsson [3] (Paper I), Bréhier [8], [7], Bréhier & Kopec [9], Hausenblas [25], [26], Kopec [32, Chapt. 5], Wang [57], [58], Wang & Gan [59]. Also for this type of equation the theory is almost complete for parabolic, hyperbolic and for Volterra type equations driven by additive Gaussian noise.

It is considerably more challenging to consider equations with multiplicative noise, i.e., equations with a noise coefficient which depends on the solution. This has been done in Andersson & Larsson [3] (Paper I), Conus et al. [16], de Bouard & Debussche [19], Debussche [20], but the results are still not satisfactory. The multiplicative noise considered in [20], and later in [3], restricts the dependence on the solution to be affine linear. In [16] this restriction is removed, but other restrictive assumptions are imposed, which are not met by any nonlinear Nemytskii operator.

7.1. Our weak convergence results. Papers I–III all treat weak convergence analysis of numerical approximations to stochastic evolution equations. Here we discuss the results of these papers and try to extract what our main achievements are in this field and also put these in relation to other works. We focus on Paper III, which contains, from a weak convergence perspective, our most important results.

This paper treats additive noise, which is regular enough so that for all $t \in (0,T]$ and $\gamma \in [0,\beta)$ it holds $\mathbf{P}$-a.s. that $X_t \in \dot{H}^\gamma$, where β is a fixed regularity parameter. The process X is either a solution to the stochastic reaction-diffusion equation of Subsection 5.1 or the stochastic Volterra integro-differential equation of Subsection 5.2. Our assumptions include any drift F, which is a nonlinear Nemytskii operator defined by a function $f \in \mathcal{C}^2_{\mathrm{b}}(\mathbf{R};\mathbf{R})$, see Section 5.1. For $d=3$ we only allow mildly singular kernels b in the case of Volterra equations, together with Nemytskii drift F. Thus, for equations driven by additive noise, we impose very natural assumptions on the drift and on the noise.

Let $(X^{h,k})_{n=0}^N$, $h,k \in (0,1)$, be a family of approximations to X, discretized in space by the finite element method with refinement parameter h, and discretized in time by the backward Euler method, with time step k. If X is a solution to a stochastic Volterra equation, then convolution quadrature is used for the convolution, see Section 6. Let $(\tilde{X}^{h,k}_t)_{t\in[0,T]}$, $h,k \in (0,1)$, denote piecewise constant interpolations.

We consider weak convergence of certain functionals of the path, more precisely, we show that for all $\gamma \in [0,\beta)$, test functions $\varphi\colon H \to \mathbf{R}$ having two continuous Fréchet derivatives with polynomial growth, and all finite Borel measures ν defined on $[0,T]$ it holds that

$$\left|\mathbf{E}\Big[\varphi\Big(\int_0^T X_t \,\mathrm{d}\nu_t\Big) - \varphi\Big(\int_0^T X^{h,k}_t \,\mathrm{d}\nu_t\Big)\Big]\right| \le C\big(h^{2\gamma} + k^{\rho\gamma}\big), \quad h,k \in (0,1). \tag{7.1}$$

If we take $\nu = \delta_T$, being the Dirac measure concentrated at T, then we get the classical type of weak convergence estimate

$$\left|\mathbf{E}\Big[\varphi\big(X_T\big) - \varphi\big(X^{h,k}_T\big)\Big]\right| \le C\big(h^{2\gamma} + k^{\rho\gamma}\big), \quad h,k \in (0,1). \tag{7.2}$$

Paper II treats weak approximation of reaction-diffusion equations, but with a technical restriction, which only allows the nonlinear drift F to be a Nemytskii operator for $d = 1$. In Paper II we found a way to remove this restriction. The type of convergence considered is of the type (7.2).

Paper I considers finite element approximation of the stochastic heat equation introduced in Section 5. It follows the same setting as in the seminal paper [20] by Debussche, which considers discretization in time by the backward Euler method. The importance of the paper [20] for subsequent works [3], [8], [7], [9], [16], [32, Chapt. 5], [57], [58], [59] can not be underestimated, but the setting is not useful for stochastic partial differential equations, and unfortunately, for us, this is also true for our Paper I. We assume that the multiplicative noise is of form $B \in \mathcal{L}(H;\mathcal{L}(H))$, with an additional additive term. This assumption excludes nontrivial linear Nemytskii operators, see example 5.4. Furthermore we assume $F \in \mathcal{C}_{\mathrm{b}}^2(H;H)$, which is a space which excludes all nonlinear Nemytskii operators, see the discussion in Subsection 5.1.

7.2. A new weak convergence analysis. Here we explain the main ideas of the weak convergence analysis introduced in Paper II, and whose advantages were utilized to a larger extent in Paper III. In this presentation we consider reaction diffusion equations, which corresponds to $\rho = 1$ in Paper III. The argument is based on the following linearization

$$\Big|\mathbf{E}\Big[\varphi(X_T) - \varphi(X_N^{h,k})\Big]\Big| = \Big|\Big\langle \Phi^{h,k}, X_T - X_N^{h,k}\Big\rangle\Big|,$$

based on the mean value theorem, where

$$\Phi^{h,k} = \int_0^1 \varphi'\Big(X_N^{h,k} + \lambda\Big(X_T - X_N^{h,k}\Big)\Big)\,\mathrm{d}\lambda.$$

In a next step we consider a Gelfand triple

$$V \subset L^2(\Omega;H) \subset V^*,$$

where V is a Banach space to be chosen. By duality in this Gelfand triple it holds that

$$\Big|\mathbf{E}\Big[\varphi(X_T) - \varphi(X_N^{h,k})\Big]\Big| \le \Big(\sup_{h,k\in(0,1)} \big\|\Phi^{h,k}\big\|_V\Big)\big\|X_T - X_N^{h,k}\big\|_{V^*}. \tag{7.3}$$

If V has the good property that for all $\gamma \in [0,\beta)$ it holds

$$\sup_{h,k\in(0,1)} \big\|\Phi^{h,k}\big\|_V < \infty, \quad \big\|X_T - X_N^{h,k}\big\|_{V^*} \le C_\gamma\Big(h^{2\gamma} + k^\gamma\Big), \quad h,k \in (0,1),$$

then this solves the weak convergence problem. Thus, we reduce the weak convergence problem into one regularity problem of bounding $\Phi^{h,k}$ in the V-norm, and one strong convergence problem in the V^*-norm.

For linear equations, and under additional, too strong, assumptions on φ it is possible to take $V = L^2(\Omega;\dot{H}^\gamma)$, see Paper II for more details. Our new approach is as follows. For linear equations, without additional assumptions on φ one can take $V = \mathbf{M}^{1,p,q}(H)$, for suitable choices of $p \in [2,\infty)$ and $q \in [2,\infty]$. To present this assume that $\mathrm{Tr}(Q) < \infty$, $\beta = 1$ and consider approximation of the stochastic convolution. The difference of the stochastic convolution and its approximation in time and space can be written in the form

$$\Delta_T^{h,k} = \int_0^T E_t^{h,k} \,\mathrm{d}W_t,$$

where $(E_t^{h,k})_{t\in[0,T]} \subset \mathcal{L}(H)$, is a piecewise constant in time interpolation of the error operator $S_{t_n} - S_n^{h,k}$. It satisfies the error bound

$$\|E_t^{h,k}\|_{\mathcal{L}(H)} \le C_\theta\Big(h^\theta + k^{\frac{\theta}{2}}\Big)(T-t)^{-\frac{\theta}{2}}, \quad t \in (0,T], \quad h,k \in (0,1), \quad \theta \in [0,2].$$

Fix $p = 2$, $q = \infty$, i.e., let $V = \mathbf{M}^{1,2,\infty}(H)$. Inequality (2.13) ensures that, for all $\epsilon > 0$ it holds

$$\begin{aligned}\Big\| \int_0^T E_t^{h,k} \,\mathrm{d}W_t \Big\|_{\mathbf{M}^{1,2,\infty}(H)^*} &\le \big\| E^{h,k} \big\|_{L^1(0,T;\mathcal{L}_2(U;H))} = \int_0^T \big\| E_t^{h,k} \big\|_{\mathcal{L}_2(U;H)} \,\mathrm{d}t \\ &\le C_{2-2\epsilon}\Big(h^{2-2\epsilon} + k^{1-\epsilon}\Big) \int_0^T (T-t)^{-1+\epsilon} \,\mathrm{d}t \\ &\le C\Big(h^{2-2\epsilon} + k^{1-\epsilon}\Big).\end{aligned}$$

This should be compared with the strong error, measured in the $L^2(\Omega;H)$-norm. While (2.13) offers an L^1-estimate in time, for the stochastic integral, the Itō isometry (2.7) offers only an L^2-estimate in time, and therefore the strong rate of convergence is only half the weak rate. More precisely, for all $\epsilon > 0$, it holds

$$\begin{aligned}\Big\| \int_0^T E_t^{h,h} \,\mathrm{d}W_t \Big\|_{L^2(\Omega;H)} &= \big\| E^{h,k} \big\|_{L^2(0,T;\mathcal{L}_2(U;H))} = \Big(\int_0^T \big\| E_t^{h,k} \big\|^2_{\mathcal{L}_2(U;H)} \,\mathrm{d}t \Big)^{\frac{1}{2}} \\ &\le C_{\frac{1-\epsilon}{2}}\Big(h^{1-\epsilon} + k^{\frac{1}{2}-\frac{\epsilon}{2}}\Big)\Big(\int_0^T (T-t)^{-1+\epsilon} \,\mathrm{d}t \Big)^{\frac{1}{2}} \\ &\le C\Big(h^{1-\epsilon} + k^{\frac{1}{2}-\frac{\epsilon}{2}}\Big).\end{aligned}$$

In strong error analysis for semilinear equations it is classical that a Gronwall argument is used. In our situation we need, in order for Gronwall's Lemma to apply, to prove that for some $\alpha \in (0,1]$ it holds

$$\big\| X_{t_n} - X_n^{h,k} \big\|_{V^*} \le C\left(h^{2\gamma} + k^\gamma + \sum_{j=0}^{n-1} t_{n-j}^{-1+\alpha} \big\| X_{t_j} - X_j^{h,k} \big\|_{V^*} \right).$$

In order to prove this, a bound is required, of the form

$$\left\| E^{h,k}_{t_j}\Big(F(X_{t_j}) - F(X^{h,k}_j)\Big)\right\|_{V^*} \leq C t_{n-j}^{-1+\alpha} \left\| X_{t_j} - X^{h,k}_j \right\|_{V^*}. \tag{7.4}$$

To obtain such a bound we introduce the spaces $\mathbf{G}^{1,p}(H) = \mathbf{M}^{1,p,p}(H) \cap L^{2p}(\Omega;H)$, $p \in [2,\infty)$, equipped with the norm

$$\|Y\|_{\mathbf{G}^{1,p}(H)} = \max\Big(\|Y\|_{\mathbf{M}^{1,p,p}(H)}, \|Y\|_{L^{2p}(\Omega;H)}\Big).$$

With $V = \mathbf{G}^{1,p}(H)$, we can show (7.4). The singularity comes from the fact that F is a Nemytskii operator.

This approach has advantages and disadvantages. One advantage is that it does not require tools from Markov theory, such as the transition semigroup or the Kolmogorov equation. Stochastic Volterra integro-differential equations are non-Markovian and our approach is to the best of our knowledge the only established approach which applies to this type of equations. Another advantage is that the more general type of weak convergence in (7.1) can be considered. A disadvantage, it seems, is that the bound (7.3) is too crude, in order to treat equations with multiplicative noise, see Paper II, Subsection 4.3, for a discussion about this.

REFERENCES

[1] A. Andersson, M. Kovács, and S. Larsson, *Weak and strong error analysis for semilinear stochastic Volterra equations*, ArXiv Preprint, arXiv:1411.6476 (2014).

[2] A. Andersson, R. Kruse, and S. Larsson, *Duality in refined Sobolev-Malliavin spaces and weak approximation of SPDE*, ArXiv Preprint, arXiv:1312.5893 (2013).

[3] A. Andersson and S. Larsson, *Weak convergence for a spatial approximation of the nonlinear stochastic heat equation*, ArXiv Preprint, arXiv:1212.5564 (2012). To appear in Math. Comp.

[4] B. Baeumer, M. Geissert, and M. Kovács, *Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise*, J. Differential Equations **258** (2015), no. 2.

[5] P. Billingsley, *Convergence of Probability Measures*, Second, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999.

[6] V. Bogachev, *Differentiable Measures and the Malliavin Calculus*, Mathematical Surveys and Monographs, vol. 164, American Mathematical Society, Providence, RI, 2010.

[7] C.-E. Bréhier, *Strong and weak orders in averaging for SPDEs*, Stochastic Process. Appl. **122** (2012), no. 7.

[8] C.-E. Bréhier, *Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise*, Potential Analysis **40** (2014), 1–40 (English).

[9] C. E. Bréhier and M. Kopec, *Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme*, ArXiv Preprint, arXiv:1311.7030 (2013).

[10] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Third, Texts in Applied Mathematics, vol. 15, Springer, New York, 2008.

[11] Z. Brzeźniak, *On stochastic convolution in Banach spaces and applications*, Stochastics Stochastics Rep. **61** (1997), no. 3-4, 245–295. MR1488138

[12] R. H. Cameron and W. T. Martin, *Transformations of Wiener integrals under translations*, Ann. of Math. (2) **45** (1944), 386–396.

[13] ______, *Transformations of Wiener integrals under a general class of linear transformations*, Trans. Amer. Math. Soc. **58** (1945), 184–219.

[14] L. Chen and R. Dalang, *Hölder continuity for the nonlinear stochastic heat equation with rough initial conditions*, arXiv:1307.0600 (2013).

[15] ______, *Moments and growth indices for the nonlinear stochastic heat equation with rough initial value*, arXiv:1307.0600 (2013). To appear in Ann. Probab.

[16] D. Conus, A. Jentzen, and R. Kurniawan, *Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients*, ArXiv Preprint, arXiv:1408.1108 (2014).

[17] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.

[18] ______, *Second Order Partial Differential Equations in Hilbert Spaces*, London Mathematical Society Lecture Note Series, vol. 293, Cambridge University Press, Cambridge, 2002. MR1985790 (2004e:47058)

[19] A. de Bouard and A. Debussche, *Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation*, Appl. Math. Optim. **54** (2006), 369–399.

[20] A. Debussche, *Weak approximation of stochastic partial differential equations: the nonlinear case*, Math. Comp. **80** (2011), no. 273, 89–117.

[21] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comp. **78** (2009), no. 266, 845–863.

[22] M. Fuhrman and G. Tessitore, *Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control*, Ann. Probab. **30** (2002), 1397–1465.

[23] M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT **49** (2009), 343–356.

[24] A. Grorud and É. Pardoux, *Intégrales Hilbertiennes anticipantes par rapport à un processus de Wiener cylindrique et calcul stochastique associé*, Appl. Math. Optim. **25** (1992), no. 1, 31–49.

[25] E. Hausenblas, *Weak approximation for semilinear stochastic evolution equations*, Stochastic analysis and related topics VIII, 2003, pp. 111–128. MR2189620 (2006k:60114)

[26] ______, *Weak approximation of the stochastic wave equation*, J. Comput. Appl. Math. **235** (2010), no. 1, 33–58.

[27] K. Itô, *Stochastic integral*, Proc. Imp. Acad. Tokyo **20** (1944), 519–524.

[28] S. Janson, *Gaussian Hilbert Spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997.

[29] A. Jentzen, *Numerical Analysis of Stochastic Partial Differential Equations*, 2014. Lecture notes, ETH Zurich, summer semester 2014, available online at http://www.math.ethz.ch/education/bachelor/lectures/fs2014/math/numsol.

[30] A. Jentzen and P. E. Kloeden, *Taylor Approximations for Stochastic Partial Differential Equations*, CBMS-NSF regional conference series in applied mathematics ; 83, SIAM, 2011 (eng).

[31] A. Jentzen and M. Röckner, *Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise*, J. Differential Equations **252** (2012), no. 1, 114–136.

[32] M. Kopec, *Quelques contributions à l'analyse numérique d'équations stochastiques*, Ph.D. Thesis, 2014.

[33] M. Kovács, S. Larsson, and F. Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise*, Numer. Algorithms **53** (2010), no. 2-3, 309–320.

[34] ______, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise*, BIT Numer. Math. **52** (2012), no. 1, 85–108.

[35] M. Kovács, S. Larsson, and F. Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes*, BIT Numer. Math. **53** (2013), no. 2, 497–525.

[36] M. Kovács, F. Lindner, and R. Schilling, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive Lévy noise*, ArXiv Preprint, arXiv:1411.1051 (2014).

[37] M. Kovács and J. Printems, *Strong order of convergence of a fully discrete approximation of a linear stochastic Volterra type evolution equation*, Math. Comp. **83** (2014), no. 289.

[38] ———, *Weak convergence of a fully discrete approximation of a linear stochastic evolution equation with a positive-type memory term*, J. Math. Anal. Appl. **413** (2014), no. 2.

[39] R. Kruse, *Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise*, IMA J. Numer. Anal. **34** (2014), no. 1, 217–251.

[40] ———, *Strong and weak approximation of stochastic evolution equations*, Lecture Notes in Mathematics, vol. 2093, Springer, 2014.

[41] J. A. León and D. Nualart, *Stochastic evolution equations with random generators*, Ann. Probab. **26** (1998), no. 1, 149–186.

[42] F. Lindner and R. L. Schilling, *Weak order for the discretization of the stochastic heat equation driven by impulsive noise*, Potential Anal. **38** (2012), no. 2, 345–179.

[43] C. Lubich, *Convolution quadrature and discretized operational calculus. I*, Numer. Math. **52** (1988), no. 2, 129–145. MR923707 (89g:65018)

[44] ———, *Convolution quadrature and discretized operational calculus. II*, Numer. Math. **52** (1988), no. 4, 413–425. MR932708 (89g:65019)

[45] A. Lunardi, *Interpolation Theory*, Second, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie). [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], Edizioni della Normale, Pisa, 2009.

[46] P. Malliavin, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), 1978, pp. 195–263.

[47] D. Nualart, *The Malliavin Calculus and Related Topics*, Second, Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.

[48] R.E.A.C. Paley, N. Wiener, and A. Zygmund, *Notes on random functions*, Math. Z. **37** (1933), no. 1, 647–668.

[49] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer, New York, 1983.

[50] S. Peszat and J. Zabczyk, *Stochastic Partial Differential Equations with Lévy Noise*, Encyclopedia of Mathematics and its Applications, vol. 113, Cambridge University Press, Cambridge, 2007.

[51] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.

[52] N. Privault, *Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales*, Lecture Notes in Mathematics, vol. 1982, Springer-Verlag, Berlin, 2009.

[53] J. Prüss, *Evolutionary integral equations and applications*, Modern Birkhäuser Classics, Birkhäuser/Springer Basel AG, Basel, 1993. [2012] reprint of the 1993 edition. MR2964432

[54] V. Thomée, *Galerkin finite element methods for parabolic problems*, Second, Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 2006.

[55] J. M. A. M. van Neerven, *Stochastic Evolution Equations*, 2008. ISEM lecture notes.

[56] J. B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, XIV—1984, 1986, pp. 265–439.

[57] X. Wang, *An exponential integrator scheme for time discretization of nonlinear stochastic wave equation*, ArXiv Preprint, arXiv:1312.5185.

[58] ———, *Weak error estimates of the exponential Euler scheme for semi-linear SPDEs without Malliavin calculus*, ArXiv Preprint, arXiv:1408.0713 (2014).

[59] X. Wang and S. Gan, *Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise*, J. Math. Anal. Appl. **398** (2013), no. 1, 151–169.

[60] N. Wiener, *Generalized harmonic analysis*, Acta Math. **55** (1930), no. 1, 117–258.